SIAM J. NUMER. ANAL.
Vol. 39, No. 2, pp. 463–479

AN ALGEBRAIC CONVERGENCE THEORY FOR RESTRICTED ADDITIVE SCHWARZ METHODS USING WEIGHTED MAX NORMS*

ANDREAS FROMMER† AND DANIEL B. SZYLD‡

Abstract. Convergence results for the restrictive additive Schwarz (RAS) method of Cai and Sarkis [*SIAM J. Sci. Comput.*, 21 (1999), pp. 792–797] for the solution of linear systems of the form $Ax = b$ are provided using an algebraic view of additive Schwarz methods and the theory of multisplittings. The linear systems studied are usually discretizations of partial differential equations in two or three dimensions. It is shown that in the case of A symmetric positive definite, the projections defined by the methods are not orthogonal with respect to the inner product defined by A, and therefore the standard analysis cannot be used here. The convergence results presented are for the class of M-matrices (and more generally for H-matrices) using weighted max norms. Comparison between different versions of the RAS method are given in terms of these norms. A comparison theorem with respect to the classical additive Schwarz method makes it possible to indirectly get quantitative results on rates of convergence which otherwise cannot be obtained by the theory. Several RAS variants are considered, including new ones and two-level schemes.

Key words. restricted additive Schwarz methods, domain decomposition, multisplittings, nonnegative matrices, parallel algorithms

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

PII. S0036142900370824

1. Introduction.

In this paper we bring forth results from multisplitting theory to study the class of *restrictive additive Schwarz (RAS) methods* which were recently introduced [11] as an efficient alternative to the classical additive Schwarz preconditioner. Practical experiments have proven RAS to be particularly attractive because it reduces communication time while maintaining the most desirable properties of the classical Schwarz methods [9, 11]. RAS preconditioners are widely used in practice and are the default preconditioner in the PETSc software package [1].

Up until now there is no general convergence theory for RAS. The purpose of this paper is to provide such a theory for a general class of matrices including M-matrices. Our convergence theory is developed in terms of weighted max norms and, in some cases, in terms of the spectral radius of the iteration matrix. A disk in the complex plane containing the spectrum of the preconditioned matrix is thus obtained. Some of the tools we use were developed in [23] for the analysis of classical additive Schwarz methods; see also [2]. Our results provide the theoretical underpinnings for the behavior of the preconditioners as observed in [11]. The theory we develop is not complete in the sense that we do not get quantitative results (like mesh independence in the presence of a coarse grid, for example). However, as we shall stress later, such results can be obtained indirectly by using some of the comparison results of section 6 and classical results for the usual additive Schwarz method.

*Received by the editors April 13, 2000; accepted for publication (in revised form) November 19, 2000; published electronically April 6, 2001.
http://www.siam.org/journals/sinum/39-2/37082.html
†Fachbereich Mathematik, Bergische Universität GH Wuppertal, Gauss-Strasse 20, D-42097 Wuppertal, Germany (frommer@math.uni-wuppertal.de).
‡Department of Mathematics, Temple University (038-16), 1805 N. Broad Street, Philadelphia, PA 19122-6094 (szyld@math.temple.edu). This author's work was supported by National Science Foundation grant DMS-9973219.

Our approach is purely algebraic, and therefore our results apply to discretization of differential equations as well as to algebraic additive Schwarz. We believe that the algebraic tools used here and in [2, 23] complement the usual analytic tools used for the analysis of Schwarz methods; see, e.g., the books [31, 36] and the extensive bibliography therein.

We formulate all our results in terms of a quantity $r \in [0, 1)$ which represents the spectral radius of the respective iteration matrix or an upper bound for that spectral radius. If we view the methods as preconditioners, our results may be restated as saying that the spectrum of the preconditioned system is contained in the circle with center 1 and radius r in the right half plane. Note that such a half plane condition, together with the assumption of diagonalizability, guarantees restarted GMRES to converge for any restart value [34]. Moreover, we may expect GMRES to converge faster when r becomes smaller, although the precise convergence behavior of GMRES does not depend only on the eigenvalues (and even less on circles containing those).

This paper is organized as follows. We start by giving algebraic representations of the usual and the restricted additive Schwarz method and we introduce the splittings associated with each of the methods. In section 3 we show that the RAS preconditioner is a sum of oblique projections. The algebraic representation and the analysis in section 3 holds for general nonsingular matrices, while in the rest of the paper we generally assume that the matrix associated with the linear system is an M-matrix (or more generally an H-matrix). Section 4 contains our central convergence theorem for RAS, whereas in section 5 we study the effect of overlap on the quality of the preconditioner. We then deal with several variants of the RAS preconditioner in section 6. Both sections 5 and 6 also identify situations where one can compare the rate of convergence of different RAS variants. In section 7 we indicate how to study the case of inexact local solutions. We finish with some brief observations on coarse grid corrections, i.e., two-level schemes (section 8), and on nonstationary and asynchronous variants (section 9).

2. The algebraic representation.

The linear system in $\mathbb{R}^n$ is given as

$$Ax = b. \tag{2.1}$$

As in [11] we consider p nonoverlapping subspaces $W_{i,0},\ i = 1, \dots, p$, which are spanned by columns of the identity I over $\mathbb{R}^n$ and which are then augmented to produce overlap. For a precise definition, let $S = \{1, \dots, n\}$ and let

$$S = \bigcup_{i=1}^{p} S_{i,0}$$

be a partition of S into p disjoint, nonempty subsets. For each of these sets $S_{i,0}$ we consider a nested sequence of larger sets $S_{i,\delta}$ with

$$S_{i,0} \subseteq S_{i,1} \subseteq S_{i,2} \cdots \subseteq S = \{1, \dots, n\}, \tag{2.2}$$

so that we again have $S = \cup_{i=1}^{p} S_{i,\delta}$ for all values of δ, but for $\delta > 0$ the sets $S_{i,\delta}$ are not necessarily pairwise disjoint, i.e., we have introduced *overlap*. A common way to obtain the sets $S_{i,\delta}$ is to add those indices to $S_{i,0}$ which correspond to nodes lying at distance δ or less from those nodes corresponding to $S_{i,0}$ in the (undirected) graph of A. This approach is particularly adequate in discretizations of partial differential equations where the indices correspond to the nodes of the discretization mesh; see [6, 9, 10, 11, 16, 36].

Let $n_{i,\delta} = |S_{i,\delta}|$ denote the cardinality of the set $S_{i,\delta}$. For each nested sequence from (2.2) we can find a permutation π_i on $\{1,\dots,n\}$ with the property that for all $\delta \geq 0$ we have $\pi_i(S_{i,\delta}) = \{1,\dots,n_{i,\delta}\}$.

We now build matrices $R_{i,\delta} \in \mathbb{R}^{n_{i,\delta}\times n}$ whose rows are precisely those rows j of the identity for which $j \in S_{i,\delta}$. Formally, such a matrix $R_{i,\delta}$ can be expressed as

$$R_{i,\delta} = [I_{i,\delta}|O]\,\pi_i \tag{2.3}$$

with $I_{i,\delta}$ the identity on $\mathbb{R}^{n_{i,\delta}}$. Finally, we define the weighting matrices

$$E_{i,\delta} = R_{i,\delta}^T R_{i,\delta} \quad \left(= \pi_i^T \begin{bmatrix} I_{i,\delta} & O \\ O & O \end{bmatrix} \pi_i \right) \in \mathbb{R}^{n\times n}$$

and the subspaces

$$W_{i,\delta} = \operatorname{range}(E_{i,\delta}), \quad i = 1,\dots,p.$$

Note the inclusion $W_{i,\delta} \supseteq W_{i,\delta'}$ for $\delta \geq \delta'$, and in particular $W_{i,\delta} \supseteq W_{i,0}$ for all $\delta \geq 0$.

We view the matrices $R_{i,\delta}$ as restriction operators and $R_{i,\delta}^T$ as prolongations. We can identify the image of $R_{i,\delta}^T$ with the subspace $W_{i,\delta}$. For each subspace $W_{i,\delta}$ we define a restriction of the operator A on $W_{i,\delta}$ as

$$A_{i,\delta} = R_{i,\delta} A R_{i,\delta}^T.$$

The classical additive Schwarz method consists of the following algorithm for solving (2.1).

ALGORITHM 2.1 (additive Schwarz). *Solve the equation*

$$M_{AS,\delta}^{-1} A x = M_{AS,\delta}^{-1} b$$

by a Krylov subspace method, where the preconditioner is defined by

$$M_{AS,\delta}^{-1} = \sum_{i=1}^{p} R_{i,\delta}^T A_{i,\delta}^{-1} R_{i,\delta}. \tag{2.4}$$

In order to describe the *restricted* additive Schwarz method we introduce "restricted" operators $\widetilde{R}_{i,\delta}$ as

$$\widetilde{R}_{i,\delta} = R_{i,\delta} E_{i,0} \in \mathbb{R}^{n_{i,\delta}\times n}. \tag{2.5}$$

The image of $\widetilde{R}_{i,\delta}^T = E_{i,0} R_{i,\delta}^T$ can be identified with $W_{i,0}$, so $\widetilde{R}_{i,\delta}^T$ "restricts" $R_{i,\delta}^T$ in the sense that the image of the latter, $W_{i,\delta}$, is restricted to its subspace $W_{i,0}$, the space from the nonoverlapping decomposition. The restricted additive Schwarz method from [9, 11] replaces the prolongation operator $R_{i,\delta}^T$ by $\widetilde{R}_{i,\delta}^T$ in Algorithm 2.1, i.e., one uses

$$M_{RAS,\delta}^{-1} = \sum_{i=1}^{p} \widetilde{R}_{i,\delta}^T A_{i,\delta}^{-1} R_{i,\delta} \tag{2.6}$$

instead of (2.4).[1] For practical parallel implementations, replacing $R_{i,\delta}^T$ by $\widetilde{R}_{i,\delta}^T$ means that the corresponding part of the computation will not require any communication,

[1] We note that the representations (2.4) and (2.6) using rectangular matrices $R_{i,\delta}$ and matrices $A_{i,\delta}$ of smaller size are consistent with the standard literature [12, 31, 36] and different than that of [11], where $n \times n$ matrices are used.

since the images of the $\widetilde{R}_{i,\delta}^T$ do not overlap. In addition, the numerical results in [11] indicate that the restrictive additive Schwarz method is at least as fast (in terms of number of iterations and/or CPU time) as the classical one. Note that we lose symmetry, however, since if A is symmetric, $M_{AS,\delta}^{-1}$ will be symmetric as well, whereas $M_{RAS,\delta}^{-1}$ will usually be nonsymmetric.

As in [2, 23], the key to our analysis is a different representation for $M_{AS,\delta}^{-1}$ and $M_{RAS,\delta}^{-1}$ in the spirit of multisplitting methods [8, 19, 20, 21, 29]. In fact, the RAS preconditioning is *the same* as the overlapping block Jacobi multisplitting introduced in [19, Definition 2.1].

LEMMA 2.2. *The following identities hold:*

$$M_{AS,\delta}^{-1} = \sum_{i=1}^{p} E_{i,\delta} M_{i,\delta}^{-1} \tag{2.7}$$

and

$$M_{RAS,\delta}^{-1} = \sum_{i=1}^{p} E_{i,0} M_{i,\delta}^{-1}, \tag{2.8}$$

where the matrices $M_{i,\delta}$ are defined as

$$M_{i,\delta} = \pi_i^T \begin{bmatrix} A_{i,\delta} & O \\ O & D_{\neg i,\delta} \end{bmatrix} \pi_i \tag{2.9}$$

and $D_{\neg i,\delta}$ is the diagonal part of the principal submatrix of A "complementary" to $A_{i,\delta}$, i.e.,

$$D_{\neg i,\delta} = \operatorname{diag}\left([O|I_{\neg i,\delta}] \cdot \pi_i \cdot A \cdot \pi_i^T \cdot [O|I_{\neg i,\delta}]^T\right)$$

with $I_{\neg i,\delta}$ the identity on $\mathbb{R}^{n-n_{i,\delta}}$. Here, we assume that $A_{i,\delta}$ and $D_{\neg i,\delta}$ are nonsingular.

Proof. Let us show that

$$\widetilde{R}_{i,\delta}^T A_{i,\delta}^{-1} R_{i,\delta} = E_{i,0} M_{i,\delta}^{-1}, \qquad i = 1, \ldots, p,$$

which implies (2.8). The identity (2.7) follows in a similar manner; see also [2, 23]. Clearly, by (2.5) and (2.3), we have

$$\widetilde{R}_{i,\delta}^T A_{i,\delta}^{-1} R_{i,\delta} = E_{i,0} R_{i,\delta}^T A_{i,\delta}^{-1} R_{i,\delta} = E_{i,0} \pi_i^T \begin{bmatrix} A_{i,\delta}^{-1} & O \\ O & O \end{bmatrix} \pi_i. \tag{2.10}$$

For any $(n - n_{i,\delta}) \times (n - n_{i,\delta})$ matrix X

$$\begin{aligned} E_{i,0}\pi_i^T \begin{bmatrix} O & O \\ O & X \end{bmatrix} \pi &= \pi_i^T \begin{bmatrix} I_{i,0} & O \\ O & O \end{bmatrix} \pi_i \pi_i^T \begin{bmatrix} O & O \\ O & X \end{bmatrix} \pi_i \\ &= \pi_i^T \begin{bmatrix} I_{i,0} & O \\ O & O \end{bmatrix} \begin{bmatrix} O & O \\ O & X \end{bmatrix} \pi_i = O. \end{aligned}$$

Choosing $X = D_{\neg i,\delta}$, we see that (2.10) yields

$$\widetilde{R}_{i,\delta}^T A_{i,\delta}^{-1} R_{i,\delta} = E_{i,0}\pi_i^T \begin{bmatrix} A_{i,\delta}^{-1} & O \\ O & O \end{bmatrix} \pi_i = E_{i,0}\pi_i^T \begin{bmatrix} A_{i,\delta}^{-1} & O \\ O & D_{\neg i,\delta}^{-1} \end{bmatrix} \pi_i = E_{i,0} M_{i,\delta}^{-1},$$

since

$$M_{i,\delta}^{-1} = \pi_i^T \left[\begin{array}{cc} A_{i,\delta}^{-1} & O \\ O & D_{\neg i,\delta}^{-1} \end{array} \right] \pi_i. \qquad \square$$

We note that with the RAS preconditioning the corresponding weighting matrices satisfy

$$\sum_{i=1}^{p} E_{i,0} = I, \tag{2.11}$$

consistent with the traditional multisplitting theory [8, 29], while for additive Schwarz we have

$$qI \geq \sum_{i=1}^{p} E_{i,\delta} \geq I, \tag{2.12}$$

where the inequalities are componentwise and

$$q = \max_{j=1,\ldots,n} |\{i : j \in S_{i,\delta}\}|. \tag{2.13}$$

In the partial differential equation setting, q is the maximum number of subdomains to which each node of the mesh belongs.

3. RAS viewed as sums of projections. The theory of orthogonal projections plays an important role in the analysis of classical Schwarz methods; see, e.g., [25, Ch. 11], [36], and especially [5]. Let

$$P_{i,\delta} = R_{i,\delta}^T A_{i,\delta}^{-1} R_{i,\delta} A = E_{i,\delta} M_{i,\delta}^{-1} A.$$

It is not hard to see that this is a projection onto the subspace $W_{i,\delta}$ and that when A is symmetric positive definite (s.p.d.) this projection is orthogonal with respect to the A-inner product $\langle x, y \rangle_A = x^T A y$, i.e., $P_{i,\delta}^2 = P_{i,\delta}$ and $AP_{i,\delta} = P_{i,\delta}^T A$; see, e.g., [25, Ch. 11], [36]. Thus, the preconditioned matrix with the additive Schwarz preconditioning (Algorithm 2.1) can be viewed as

$$M_{AS,\delta}^{-1} A = \sum_{i=1}^{p} P_{i,\delta},$$

i.e., as a sum of A-orthogonal projections.

In the restrictive additive Schwarz method, the operator corresponding to $P_{i,\delta}$ is

$$Q_{i,\delta} = \tilde{R}_{i,\delta}^T A_{i,\delta}^{-1} R_{i,\delta} A = E_{i,0} M_{i,\delta}^{-1} A. \tag{3.1}$$

As we show in the following lemma, this operator is a projection onto the subspace $W_{i,0}$, but this projection is not orthogonal. Therefore, the preconditioned matrix with RAS is

$$M_{RAS,\delta}^{-1} A = \sum_{i=1}^{p} Q_{i,\delta},$$

i.e., a sum of oblique projections.

LEMMA 3.1. *Let $Q_{i,\delta}$ be defined as in* (3.1)*; then the following identity holds:*

$$Q_{i,\delta} = E_{i,0}P_{i,\delta}. \tag{3.2}$$

Furthermore, $Q_{i,\delta}$ is a projection.

Proof. First we note that $E_{i,0}E_{i,\delta} = E_{i,0}$, which implies (3.2). To prove that $Q_{i,\delta}$ is a projection, we first observe that

$$A_{i,\delta}^{-1}R_{i,\delta}AR_{i,0}^T = \begin{bmatrix} I_{i,0} & O \\ O & O \end{bmatrix} \in \mathbb{R}^{n_{i,\delta}\times n_{i,\delta}}.$$

By multiplying this equation by $R_{i,0}^T$ we get $R_{i,0}^T A_{i,\delta}^{-1}R_{i,\delta}AR_{i,0}^T = R_{i,0}^T$. Putting all this together and using the identity $E_{i,0} = R_{i,0}^T R_{i,0}$ we have

$$\begin{aligned} Q_{i,\delta}^2 &= R_{i,0}^T R_{i,0} R_{i,\delta}^T A_{i,\delta}^{-1} R_{i,\delta} A R_{i,0}^T R_{i,0} R_{i,\delta}^T A_{i,\delta}^{-1} R_{i,\delta} A \\ &= E_{i,0} R_{i,\delta}^T A_{i,\delta}^{-1} R_{i,\delta} A = Q_{i,\delta}. \qquad \square \end{aligned}$$

The matrix $E_{i,0}$ is also a projection which is orthogonal with respect to the Euclidean inner product. Lemma 3.1 indicates that the projection $Q_{i,\delta}$ is the product of two projections. These projections are orthogonal with respect to *different* inner products, and thus the theory from [5] does not apply. In the case of A s.p.d., it turns out that $Q_{i,\delta}$ is not in general an orthogonal projection (with respect to the A-inner product). Therefore, the convergence theory developed using A-norms (or energy norms) cannot be developed using the same analytical or algebraic tools as for the classical additive Schwarz method (Algorithm 2.1); see, e.g., [2, 23, 25, 31, 36] and the references therein.

This is why in this paper, we concentrate our study on the case where A in (2.1) is a nonsingular M-matrix, and in particular an s.p.d. M-matrix (Stieltjes matrix). Many of our results hold for more general H-matrices, and the proofs can be obtained in a similar manner using the techniques found, e.g., in [22]. We will not dwell on these here.

M-matrices arise naturally in discretizations of (non-self-adjoint) convection-diffusion equations when the boundary conditions are such that the problem is of "monotone kind" in the sense of Collatz [14]. See [14] and [15] for a detailed discussion and many physical examples. Discretizations which reflect monotone kind result in M-matrices. In particular, this is the case for standard finite difference discretizations if the mesh size is small enough or if upwind differences are used for the convection term [30]. It is also the case for exponentially adapted finite difference or finite element discretizations, independently of the mesh size. This was recently shown by Schrader [35].

4. Convergence of RAS.

The purpose of this section is to show that for M-matrices the spectral radius $\rho(I - M_{RAS,\delta}^{-1}A)$ of the RAS iteration matrix is less than 1 for all values of $\delta \geq 0$. This implies in particular that the spectrum of the preconditioned system $\sigma(M_{RAS,\delta}^{-1}A)$ is located in the right half plane and contained in a disk of radius less than 1 around the point 1.

We start by recalling some basic terminology. The natural partial ordering $\leq$ between matrices $A = (a_{ij})$, $B = (b_{ij})$ of the same size is defined componentwise, i.e., $A \leq B$ iff $a_{ij} \leq b_{ij}$ for all i, j. If $A \geq O$ we call A nonnegative. If all entries of A are positive, we say that A is positive and write $A > O$. This notation and terminology carries over to vectors as well.

A matrix $A \in \mathbb{R}^{n\times n}$ is called a (nonsingular) M-matrix if it has nonpositive off-diagonal elements and $A^{-1} \geq O$. The following lemma states some useful properties of M-matrices; see, e.g., [4, 40].

LEMMA 4.1. *Let $A, B \in \mathbb{R}^{n\times n}$ be two nonsingular M-matrices with $A \leq B$. Then we have the following:*

(i) *Every principal submatrix of A or B is again an M-matrix.*
(ii) *Every matrix C such that $A \leq C \leq B$ is an M-matrix—in particular, if $A \leq C \leq \operatorname{diag}(A)$, then C is an M-matrix.*
(iii) $B^{-1} \leq A^{-1}$.

Our convergence results will be formulated in terms of nonnegative splittings according to the following definition and theorem.

DEFINITION 4.2. *Consider the splitting $A = M - N \in \mathbb{R}^{n\times n}$ with M nonsingular. This splitting is said to be*

(i) regular *if $M^{-1} \geq O$ and $N \geq O$,*
(ii) weak nonnegative of the first type *(also called* weak regular*) if $M^{-1} \geq O$ and $M^{-1}N \geq O$,*
(iii) weak nonnegative of the second type *if $M^{-1} \geq O$ and $NM^{-1} \geq O$, and*
(iv) nonnegative *if $M^{-1} \geq O$, $M^{-1}N \geq O$, and $NM^{-1} \geq O$.*

THEOREM 4.3. *Let $A = M - N$ be a splitting satisfying one of the following conditions:*

(i) *The splitting is regular.*
(ii) *The splitting is weak nonnegative of the first type.*
(iii) *The splitting is weak nonnegative of the second type.*
(iv) *The splitting is nonnegative.*

Then $\rho(I - M^{-1}A) < 1$ iff A is nonsingular and $A^{-1} \geq 0$.

Proof. The case (i) goes back to Varga [40], case (ii) can be found in [30, 2.4.17], and case (iii) is from [42]. Case (iv) follows from (ii) or (iii). ∎

Thus, in particular, if A is a nonsingular M-matrix, weak nonnegative splittings of the first or the second type are convergent.

We are now able to formulate the central result of this section, which is very similar to [19, Theorem 2.1].

THEOREM 4.4. *Let A be a nonsingular M-matrix. Then for each value of $\delta \geq 0$, the splitting $A = M_{RAS,\delta} - N_{RAS,\delta}$, corresponding to the RAS method, is weak nonnegative of the first type. In particular, the iteration matrix $M^{-1}_{RAS,\delta}N_{RAS,\delta} = I - M^{-1}_{RAS,\delta}A$ satisfies*

$$\rho(I - M^{-1}_{RAS,\delta}A) < 1. \tag{4.1}$$

Proof. By Lemma 2.2, we have that the identity (2.8) holds with $M_{i,\delta}$ of the form (2.9). Clearly, $A \leq M_{i,\delta} \leq \operatorname{diag}(A)$, since $M_{i,\delta}$ arises from A by setting those off-diagonal elements to zero which do not belong to the block defined by the indices in $S_{i,\delta}$. Therefore, by Lemma 4.1(ii) each $M_{i,\delta}$ is an M-matrix with $M^{-1}_{i,\delta} \geq O$, and therefore

$$M^{-1}_{RAS,\delta} = \sum_{i=1}^{p} E_{i,0} M^{-1}_{i,\delta} \geq O.$$

In addition, since $A \leq M_{i,\delta}$, we obtain, after multiplication by $M^{-1}_{i,\delta} \geq O$, that

$M_{i,\delta}^{-1}A \le I$ for $i = 1, \ldots, p$. Using (2.11) we therefore get

$$I - M_{RAS,\delta}^{-1}A = I - \sum_{i=1}^{p} E_{i,0} M_{i,\delta}^{-1} A \ge I - \sum_{i=1}^{p} E_{i,0} = O.$$

This shows that the RAS-splitting is weak nonnegative of the first type. Theorem 4.3 therefore yields (4.1). ∎

We point out that in general a convergence result such as (4.1) does not hold for the classical additive Schwarz preconditioner (2.4). To guarantee convergence, a damping (or relaxation) parameter $\theta > 0$ is introduced. It can be shown that if $\theta \le 1/q$, then $\rho(I - \theta M_{AS,\delta}^{-1}A) < 1$, where q is defined in (2.13); see [2, 23] and also [25]. Thus, one of the attractive features of the RAS preconditioner is that no damping parameter is needed for convergence.

5. The effect of overlap on RAS. We study in this section the effect of varying the overlap. More precisely, we prove comparison results on the spectral radii and/or on certain weighted max norms for the corresponding iteration matrices

$$T_{RAS,\delta} = I - M_{RAS,\delta}^{-1}A$$

for different values of $\delta \ge 0$.

We start with a result which compares one RAS splitting, defined through the sets $S_{i,\delta'}$ with another one with more overlap defined through sets $S_{i,\delta}$, where $S_{i,\delta'} \subseteq S_{i,\delta}, i = 1, \ldots, p$. We show that the larger the overlap ($\delta \ge \delta'$), the faster RAS method converges as measured in certain weighted max norms. This is consistent with the experiments in tables 1 and 2 of [11], where an increase of the overlap is associated with fewer iterations.

For a positive vector w we denote $\|x\|_w$ the weighted max norm in $\mathbb{R}^n$ given by

$$\|x\|_w = \max_{i=1,\ldots,n} |x_i|/w_i.$$

The resulting operator norm in $\mathbb{R}^{n\times n}$ is denoted similarly and for $B = (b_{ij}) \in \mathbb{R}^{n\times n}$ we have (see, e.g., [32])

$$\|B\|_w = \max_{i=1,\ldots,n} \left(\sum_{j=1}^{n} |b_{ij}| w_j \right) / w_i. \tag{5.1}$$

The following lemma follows directly from (5.1).

LEMMA 5.1. *Let T, $\tilde{T}$ be nonnegative matrices. Assume that $Tw \le \tilde{T}w$ for some vector $w > 0$. Then $||T||_w \le ||\tilde{T}||_w$.*

THEOREM 5.2. *Let A be a nonsingular M-matrix and let $w > 0$ be any positive vector such that $Aw > 0$, e.g., $w = A^{-1}v$ with $v > 0$. Then, if $\delta \ge \delta'$,*

$$\|T_{RAS,\delta}\|_w \le \|T_{RAS,\delta'}\|_w. \tag{5.2}$$

Moreover, if the Perron vector $w_{\delta'}$ of $T_{RAS,\delta'}$ satisfies $w_{\delta'} > 0$ and $Aw_{\delta'} \ge 0$, then we also have

$$\rho(T_{RAS,\delta}) \le \rho(T_{RAS,\delta'}). \tag{5.3}$$

Proof. Since $S_{i,\delta'} \subseteq S_{i,\delta}, i = 1, \ldots, p$, we have $A \le M_{i,\delta} \le M_{i,\delta'} \le \operatorname{diag}(A)$. Since A is an M-matrix, this yields

$$M_{i,\delta}^{-1} \ge M_{i,\delta'}^{-1} \quad \text{for} \quad i = 1, \ldots, p, \tag{5.4}$$

so that we obtain for all $w > 0$ such that $v = Aw > 0$,

$$\begin{aligned} 0 \le T_{RAS,\delta} w &= \left(I - \sum_{i=1}^{p} E_{i,0} M_{i,\delta}^{-1} A \right) w = w - \sum_{i=1}^{p} E_{i,0} M_{i,\delta}^{-1} v \\ &\le w - \sum_{i=1}^{p} E_{i,0} M_{i,\delta'}^{-1} v = T_{RAS,\delta'} w, \end{aligned}$$

and using Lemma 5.1, we get (5.2). Now, if the Perron vector $w_{\delta'}$ can be chosen as w, we have $\|T_{RAS,\delta'}\|_{w_{\delta'}} = \rho(T_{RAS,\delta'})$, so that (5.2) yields $\|T_{RAS,\delta}\|_{w_{\delta'}} \le \rho(T_{RAS,\delta'})$, and since the spectral radius is never larger than any operator norm, thus we have (5.3). ∎

In the case that (5.3) holds, Theorem 5.2 indicates that the spectrum of the preconditioned matrix is included in a possibly smaller disk when the overlap is increased.

We remark here that (5.2) (as well as most results using the weighted max norms in the paper) holds for *any* positive vector w such that Aw is positive, so that one has a lot of freedom in choosing the norm. For example, if all row-sums of A are positive we can choose as w the vector of all ones, and thus the weighted max norm is simply the max norm. A commonly chosen vector w is the row-sums of A^{-1}, which is always positive.

For $\delta' = 0$, i.e., for the block Jacobi preconditioner we can always provide the comparison of the spectral radii (5.3) in addition to the comparison (5.2). To this purpose, we will use the following comparison theorem due to Woźnicki [42]; see also [13].

THEOREM 5.3. *Let $A^{-1} \ge O$ and two splittings $A = M - N = \tilde{M} - \tilde{M}$, where one of them is weak nonnegative of the first type and the other is weak regular of the second type. If*

$$M^{-1} \ge \tilde{M}^{-1}, \tag{5.5}$$

then

$$\rho(M^{-1}N) \le \rho(\tilde{M}^{-1}\tilde{N}). \tag{5.6}$$

Furthermore, if the inequality (5.5) *is strict and $A^{-1} > O$, then the inequality* (5.6) *is also strict.*

Note that this result holds particularly if one of the splittings is regular, in which case it goes back to Elsner [17].

The following theorem is in fact [19, Theorem 2.2].

THEOREM 5.4. *Let A be a nonsingular M-matrix. Then, for any value of $\delta \ge 0$, $\rho(T_{RAS,\delta}) \le \rho(T_{RAS,0})$.*

Proof. We verify the hypotheses of Theorem 5.3. By Theorem 4.3, the splitting induced by $T_{RAS,\delta}$ is nonnegative of the first kind. The splitting induced by $T_{RAS,0}$ is a regular splitting. This follows by observing that the inverse of a block diagonal matrix is block diagonal with the appropriate inverses in the diagonal blocks, i.e., we have

$M_{RAS,0} = \sum_{i=1}^{p} E_{i,0} M_{i,0}$ and consequently $M_{RAS,0} - A = \sum_{i=1}^{p} E_{i,0}(M_{i,0} - A) \geq O$ since $M_{i,0} - A \geq O$ for all i. Furthermore, since (5.4) holds for $\delta' = 0$, we have

$$M_{RAS,\delta}^{-1} = \sum_{i=1}^{p} E_{i,0} M_{i,\delta}^{-1} \geq \sum_{i=1}^{p} E_{i,0} M_{i,0}^{-1} = M_{RAS,0}^{-1},$$

from which the theorem follows. $\square$

We note that results similar to those of Theorems 5.2 and 5.4 were obtained in [2] for classical, but damped, additive Schwarz methods.

6. RAS variants: ASH, RASH, WRAS, and WASH. As a variant of RAS, the authors of [11] introduced the *additive Schwarz preconditioner with harmonic extension (ASH)*, which in our notation reads

$$M_{ASH,\delta}^{-1} = \sum_{i=1}^{p} R_{i,\delta}^T A_{i,\delta}^{-1} \widetilde{R}_{i,\delta} = \sum_{i=1}^{p} M_{i,\delta}^{-1} E_{i,0}. \tag{6.1}$$

This variant is similar to multisplittings with *postweighting*; see [18, 41].

As was observed in [11, Remark 2.4], the ASH preconditioner exhibits a similar convergence behavior as the RAS preconditioner. Indeed, if A is symmetric, one can show that the spectrum of the preconditioned matrix is the same for both ASH and RAS.

THEOREM 6.1. *Let $M_{RAS,\delta}^{-1}$ and $M_{ASH,\delta}^{-1}$ be the RAS and ASH preconditioners* (2.6) *and* (6.1) *corresponding to the same subspaces $W_{i,\delta}$ and assume that A is symmetric. Then*

$$\sigma(M_{ASH,\delta}^{-1} A) = \sigma(M_{RAS,\delta}^{-1} A).$$

Proof. Note that all matrices $A_{i,\delta}$ are symmetric. Since $\sigma(BC) = \sigma(CB)$ for any two square matrices B and C, and since the spectrum remains invariant under transposition, we have

$$\sigma(M_{ASH,\delta}^{-1} A) = \sigma(A M_{ASH,\delta}^{-1}) = \sigma\left(\left(M_{ASH,\delta}^{-1}\right)^T A\right).$$

However,

$$\left(M_{ASH,\delta}^{-1}\right)^T = \sum_{i=1}^{p} \left(R_{i,\delta}^T A_{i,\delta}^{-1} \widetilde{R}_{i,\delta}\right)^T = \sum_{i=1}^{p} \widetilde{R}_{i,\delta}^T A_{i,\delta}^{-1} R_{i,\delta} = M_{RAS,\delta}^{-1}. \qquad \square$$

In general, we obviously have

$$\left(M_{ASH,\delta}^{-1}\right)^T = \sum_{i=1}^{p} \widetilde{R}_{i,\delta}^T \left(A_{i,\delta}^{-1}\right)^T R_{i,\delta}$$

so that every ASH-splitting of A gives rise to a corresponding RAS-splitting of A^T. If A is an M-matrix, then A^T is an M-matrix too. Therefore, applying Theorems 4.4, 5.2, and 5.4 to the transposed matrices while restating the hypotheses and the conclusions in terms of the original matrices, wherever this is appropriate, we obtain the following result.

THEOREM 6.2. *Let A be a nonsingular M-matrix. Then we have the following:*

(i) *For any value of $\delta \geq 0$, the splitting $A = M_{ASH,\delta} - N_{ASH,\delta}$, corresponding to the ASH method, is weak nonnegative of the second type.*

(ii) *The iteration matrix $T_{ASH,\delta} = M_{ASH,\delta}^{-1} N_{ASH,\delta} = I - M_{RAS,\delta}^{-1} A$ satisfies*

$$\rho(T_{ASH,\delta}) < 1.$$

(iii) *For any positive vector w such that $w^T A > 0$ and for $\delta \geq \delta'$ we have*

$$\|T_{ASH,\delta}\|_{1,w} \leq \|T_{ASH,\delta'}\|_{1,w},$$

where $\|\cdot\|_{1,w}$ is the weighted column sum norm defined for $B = (b_{ij}) \in \mathbb{R}^{n \times n}$ as

$$\|B\|_{1,w} = \max_{j=1,\ldots,n} \left(\sum_{i=1}^{n} |b_{ij}| w_i \right) / w_j.$$

Moreover, if the Perron vector $w_{\delta'}$ of $T_{ASH,\delta'}^T$ satisfies $w_{\delta'} > 0$ and $A^T w_{\delta'} \geq 0$, then we also have

$$\rho(T_{ASH,\delta}) \leq \rho(T_{ASH,\delta'}).$$

(iv) *For any value of $\delta \geq 0$, $\rho(T_{ASH,\delta}) \leq \rho(T_{ASH,0})$.*

As another variant, let us consider the symmetrized version of RAS and ASH, which is the RASH preconditioner of [11], where

$$M_{RASH,\delta}^{-1} = \sum_{i=1}^{p} \widetilde{R}_{i,\delta}^T A_{i,\delta}^{-1} \widetilde{R}_{i,\delta}.$$

It was observed in [11] that this preconditioner is less efficient than ASH or RAS in practice. From the theoretical point of view, RASH also seems to be less attractive. Even for a symmetric M-matrix, convergence cannot be guaranteed, as the following example shows.

Example 6.3. Consider the symmetric M-matrix

$$A = \frac{1}{2} \cdot \begin{bmatrix} 7 & -2 & -2 & -2 \\ -2 & 7 & -2 & -2 \\ -2 & -2 & 7 & -2 \\ -2 & -2 & -2 & 7 \end{bmatrix} \quad \text{with} \quad A^{-1} = \frac{1}{9} \cdot \begin{bmatrix} 6 & 4 & 4 & 4 \\ 4 & 6 & 4 & 4 \\ 4 & 4 & 6 & 4 \\ 4 & 4 & 4 & 6 \end{bmatrix}$$

and take two nonoverlapping subspaces defined through the sets $S_{1,0} = \{1, 2\}$, $S_{2,0} = \{3, 4\}$ with restriction operators

$$R_{1,0} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad R_{2,0} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

and for the larger, overlapping subspaces let us take $S_{1,1} = \{1, 2, 3\}$, $S_{2,1} = \{2, 3, 4\}$ with restrictions

$$R_{1,1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad R_{2,1} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Then we have

$$A_{1,1} = A_{2,1} = \frac{1}{2} \cdot \begin{bmatrix} 7 & -2 & -2 \\ -2 & 7 & -2 \\ -2 & -2 & 7 \end{bmatrix} \text{ and } A_{1,1}^{-1} = A_{2,1}^{-1} = \frac{1}{27} \begin{bmatrix} 10 & 4 & 4 \\ 4 & 10 & 4 \\ 4 & 4 & 10 \end{bmatrix},$$

resulting in

$$M_{RASH}^{-1} = \sum_{i=1}^{2} E_{i,0} R_{i,1}^T A_{i,1}^{-1} R_{i,1} E_{i,0} = \frac{1}{27} \cdot \begin{bmatrix} 10 & 4 & 0 & 0 \\ 4 & 10 & 0 & 0 \\ 0 & 0 & 10 & 4 \\ 0 & 0 & 4 & 10 \end{bmatrix}.$$

It is then easy to compute

$$I - M_{RASH}^{-1} A = I - A M_{RASH}^{-1} = \frac{1}{27} \cdot \begin{bmatrix} -4 & -4 & 14 & 14 \\ -4 & -4 & 14 & 14 \\ 14 & 14 & -4 & -4 \\ 14 & 14 & -4 & -4 \end{bmatrix},$$

which shows that the RASH-splitting is neither nonnegative of the first type nor of the second type. Moreover, $4/3$ is an eigenvalue of $I - M_{RASH}^{-1} A$ with eigenvector $(1, 1, -1, -1)^T$, and therefore RASH for this matrix is not convergent.

It has been observed heuristically in [11] that yet other modifications of the restricted additive Schwarz method and the classical additive Schwarz methods are particularly efficient. For these new modifications, one introduces weighted restriction operators $R_{i,\delta}^{\omega}$ which result from $R_{i,\delta}$ by replacing the entry 1 in column j by $1/k$, where k is the number of sets $S_{i,\delta}$ the component j belongs to (note that $k \leq q$). With this notation we have

$$\sum R_{i,\delta}^T R_{i,\delta}^{\omega} = I, \tag{6.2}$$

and the weighted restricted additive Schwarz (WRAS) preconditioner and the weighted restricted additive Schwarz preconditioner with harmonic extension (WASH) are then defined as

$$M_{WRAS,\delta}^{-1} = \sum_{i=1}^{p} (R_{i,\delta}^{\omega})^T A_{i,\delta}^{-1} R_{i,\delta} \quad \text{and} \quad M_{WASH,\delta}^{-1} = \sum_{i=1}^{p} R_{i,\delta}^T A_{i,\delta}^{-1} R_{i,\delta}^{\omega}. \tag{6.3}$$

We point out that these preconditioners can be defined in a more general way as that used in [11], namely, by assigning different weights to each component, not just the same value $1/k$, as long as these weights add up to 1, and thus (6.2) holds, just as in the classical multisplitting theory.

Both preconditioners (6.3) yield convergent iterations as a consequence of the following theorem, together with Theorem 4.3.

THEOREM 6.4. *Let A be a nonsingular M-matrix. Then for any value of $\delta \geq 0$ both splittings*

$$A = M_{WRAS,\delta} - (M_{WRAS,\delta} - A) \quad \textit{and} \quad A = M_{WASH,\delta} - (M_{WASH,\delta} - A)$$

are weak nonnegative of the first type and the second type, respectively.

Proof. The proof can be done in a similar manner as for the corresponding result for RAS in Theorem 4.4. □

As a final result in this section, we now compare the resulting iteration matrices $T_{WRAS,\delta}$ and $T_{WASH,\delta}$ with that of the damped classical Schwarz iteration with damping factor $\theta \leq 1/q$ (q as in (2.13)), i.e., with $T_\theta = I - \theta \sum_{i=1}^p R_{i,\delta}^T A_{i,\delta}^{-1} R_{i,\delta} A = I - \theta \sum_{i=1}^p E_{i,\delta} M_{i,\delta}^{-1}$. In fact, all we need is that $R_{i,\delta}^\omega \geq \theta R_{i,\delta}$, i.e., that each nonzero weight in $R_{i,\delta}^\omega$ be larger than the damping factor.

THEOREM 6.5. *Let $R_{i,\delta}^\omega \geq \theta R_{i,\delta}$ for $i = 1, \ldots, p$. Then*

$$\rho(T_{WRAS,\delta}) \leq \rho(T_\theta) \text{ and } \rho(T_{WASH,\delta}) \leq \rho(T_\theta). \tag{6.4}$$

Proof. As was shown in [2], the unique splitting $A = M_\theta - N_\theta$ corresponding to T_θ is a nonnegative splitting. From (6.3), and using that $R_{i,\delta}^\omega \geq \theta R_{i,\delta}$, we have

$$M_{WRAS}^{-1} \geq M_\theta^{-1} \text{ and } M_{WASH}^{-1} \geq M_\theta^{-1}.$$

Therefore, Theorem 6.4 together with Theorem 5.3 proves (6.4). ∎

This theorem shows that the weighted additive Schwarz variants converge faster than the damped classical additive Schwarz method. In terms of the quality of the preconditioner this was observed in [11]. In our analysis this is to be attributed to the fact that for the weighted variants we have (6.2), while for damped additive Schwarz we have (2.12) which implies $\theta \sum_{i=1}^p E_{i,\delta} \leq I$.

An important consequence of Theorem 6.5 is the fact that it indirectly makes it possible to get quantitative upper bounds on $\rho(T_{WRAS,\delta})$ and $\rho(T_{WASH}, \delta)$. Indeed, if such bound is obtained for T_θ using the well-established theory for additive Schwarz methods for symmetric and positive definite operators (see [6, 9, 10, 11, 16, 25, 36], for example), it is also an upper bound on $\rho(T_{WRAS,\delta})$ and $\rho(T_{WASH,\delta})$.

7. Inexact local solves. Very often in practice, instead of solving a local problem $A_{i,\delta} x_i = z_i$ exactly, its solution is approximated by $\tilde{x}_i = \tilde{A}_{i,\delta}^{-1} z_i$ where $\tilde{A}_{i,\delta}$ is an approximation of $A_{i,\delta}$. This is the case, for example, when we perform some "inner" iteration, based on a splitting of $A_{i,\delta}$, to approximate x_i; see, e.g., [2, 7, 23, 31, 36]. In this section, we briefly study the effect of such inexact local solves.

The methods with inexact local solves can be described by replacing $A_{i,\delta}$ with $\tilde{A}_{i,\delta}$ in the expression for the preconditioners and the iteration matrices.

As in [23], suppose that the inexact solves are such that the splittings

$$A_{i,\delta} = \tilde{A}_{i,\delta} - (\tilde{A}_{i,\delta} - A_{i,\delta}),\ i = 1, \ldots, p, \tag{7.1}$$

are weak nonnegative splittings of the first type. This condition is quite natural, since if $\tilde{A}_{i,\delta}$ corresponds to some ν, say, steps of an inner iteration belonging to a weak nonnegative splitting of the first type, $A_{i,\delta} = B_{i,\delta} - C_{i,\delta}$, it was shown in [3] that the resulting splitting $A_i = \tilde{A}_i - (\tilde{A}_i - A_i)$ is again weak regular of the first type. Moreover, we have

$$O \ \leq \ \tilde{A}_{i\delta} = \left(\sum_{\mu=0}^{\nu-1} (B_{i,\delta}^{-1} C_{i,\delta})^\mu \right) B_{i,\delta}^{-1} \ \leq \ \left(\sum_{\mu=0}^{\infty} (B_{i,\delta}^{-1} C_{i,\delta})^\mu \right) B_{i,\delta}^{-1} \ = \ A_{i,\delta}^{-1}.$$

Since $A_{i,\delta}$ is an M-matrix, many of the standard iterative methods indeed represent regular splittings (and thus weak nonnegative splittings of either type): Jacobi, Gauss–Seidel and their block variants, and also several ILU splittings; see [28, 39].

With these observations it should be obvious that we can now establish a theory for inexact restricted additive Schwarz methods by following the lines of the previous sections. In particular, analogous to Theorem 4.4 we get that if A is an M-matrix and (7.1) is satisfied, inexact RAS is convergent.

We do not want to go into details here, but we would like to end this section with an interesting result which can be used to compare exact and inexact local solves.

THEOREM 7.1. *Let A be an M-matrix and consider two inexact RAS methods where the matrices $\tilde{A}_{i,\delta}$ and $\hat{A}_{i,\delta}$ corresponding to inexact solves satisfying* (7.1) *and*

$$O \le \hat{A}_{i,\delta}^{-1} \le \tilde{A}_{i,\delta}^{-1} \le A_{i,\delta}^{-1},\ i = 1, \dots, p.$$

Let the corresponding iteration matrices be

$$\tilde{T} = I - \sum_{i=1}^{p} \widetilde{R}_{i,\delta}^T \tilde{A}_{i,\delta}^{-1} R_{i,\delta} A \quad \text{and} \quad \hat{T} = I - \sum_{i=1}^{p} \widetilde{R}_{i,\delta}^T \hat{A}_{i,\delta}^{-1} R_{i,\delta} A.$$

Then for any positive vector w such that $Aw > 0$ we have

$$\|T_{RAS,\delta}\|_w \le \|\tilde{T}\|_w \le \|\hat{T}\|_w < 1. \tag{7.2}$$

Proof. From the hypothesis it follows that both matrices $\tilde{T}$ and $\hat{T}$ are nonnegative and that $T_{RAS,\delta} w \le \tilde{T} w \le \hat{T} w$. By Lemma 5.1, this establishes (7.2). □

If the inexact solves are due to several steps of an "inner" iteration based on a weak regular splitting, this theorem shows that taking more steps in the inner iteration results in an improved convergence (measured in the $\|\cdot\|_w$ norm) of the inexact RAS iteration.

As a final observation on inexact local solves we note that Theorems 3.2 and 3.3 in [27] can be interpreted as a comparison between nonoverlapping inexact RAS with overlapping inexact RAS.

8. Coarse grid corrections. It has been shown theoretically, and confirmed in practice, that a coarse preconditioner (or coarse grid correction) improves the performance of the classical additive Schwarz preconditioner (2.4). This coarse correction can be applied either additively or multiplicatively; see, e.g., [2, 12, 23, 31, 36]. This corresponds to a two-level scheme, the coarse correction being the second level.

We discuss in this section the use of a coarse preconditioner in conjunction with the RAS method and its variants and its effect on the convergence behavior. We confine ourselves to coarse corrections of the form $R_0^T A_0^{-1} R_0$, where $A_0 = R_0 A R_0^T$, and R_0 is a prolongation of the form (2.3); see [2, 23]. The idea is that at least one nonzero row of $R_0 \in \mathbb{R}^{n_0 \times n}$ corresponds to an index in each set S_i, $i = 1, \dots, p$, and thus $n_0 \ge p$. In particular, the coarse grid correction proposed in [38] and used, e.g., in [26], is of this form.

Consider a general iteration matrix $T = I - M^{-1}A$, where M^{-1} is any of the RAS preconditioners discussed throughout the paper (including those with inexact local solves). The iteration matrices with the correction are

$$T_{ac} = I - \theta(R_0^T A_0^{-1} R_0 + M^{-1})A$$

and

$$T_{mc} = (I - R_0^T A_0^{-1} R_0 A)(I - M^{-1}A).$$

It can be shown, following the techniques used in [2, 23] and in sections 5 and 6, that for any $w > 0$ with $Aw > 0$,

$$\|T_{mc}\|_w \le \|T\|_w < 1 \quad \text{and} \quad \|T_{ac}\|_w \le \|I - \theta M^{-1} A\|_w < 1$$

for $\theta \leq 1/2$. Furthermore, if $A = M - (A - M)$ is nonnegative of the first or second type, then the splitting induced by the iteration matrix of the preconditioner with the correction is nonnegative of the same type. Moreover, in the case of the multiplicative correction, augmenting the coarse subspace, i.e., adding rows to R_0, yields faster convergence as measured in the corresponding weighted max norm. For the sake of brevity, these proofs are omitted.

9. Nonstationary and asynchronous versions. We end the paper by pointing out that the theory developed here is pretty general and applies to other important situations. In particular we can talk about *nonstationary* WRAS iterations, in which the weights in $R^{\omega}_{i,\delta}$ change from one iteration to the next, for example, giving more weight to a set of variables (subdomain) where the approximation is closer to the solution, as suggested, e.g., in [20, 21]. Since we can chose a *common* weighted max norm, for which $\|T_{WRAS,\delta}\|_w < 1$ for all values of the weights of $R^{\omega}_{i,\delta}$ satisfying (6.2), e.g., $w = A^{-1}(1, \ldots, 1)^T$, this nonstationary WRAS iteration converges to the solution of (2.1).

In terms of preconditioners, the nonstationary WRAS implies that the preconditioner is changed from one step of the Krylov subspace method to the next. In this situation one needs to use a Krylov method with variable (or flexible) preconditioning such as FGMRES [33] or FQMR [37]; see also [24].

Similarly, one can consider situations in which the overlap used may change from step to step. Again, we have a nonstationary iteration, this time of the RAS method, for example, and the same considerations as in the previous two paragraphs apply.

Furthermore, one can consider parallel asynchronous versions of RAS methods, namely those in which the correction to the residual contributed by each summand $\widetilde{R}^T_{i,\delta} A^{-1}_{i,\delta} R_{i,\delta} = E_{i,0} M^{-1}_{i,\delta}$ takes place when it is completed independently of the others. This situation is essentially treated in [21], and we will not repeat the details here. All we will say is that we can use the theory of [21] precisely because we have (2.11) or (6.2).

Acknowledgments. We thank Olof B. Widlund, who suggested we study the convergence of the RAS preconditioner. We are also grateful to an anonymous referee for helpful comments. The remark after Theorem 6.5 concerning qualitative results is due to that referee.

Most of this paper was prepared while the second author visited Bergische Universität GH Wuppertal. Their support and warm hospitality is very much appreciated.

REFERENCES

[1] S. Balay, W. D. Gropp, L. C. McInnes, and B. F. Smith, *PETSc* 2.0 *User's Manual*, Technical report ANL-95/11-Revision 2.0.22, Argonne National Laboratory, Argonne, IL, 1998; available online from http://www.mcs.anl.gov/petsc.

[2] M. Benzi, A. Frommer, R. Nabben, and D. B. Szyld, *Algebraic theory of multiplicative Schwarz methods*, Numer. Math., to appear.

[3] 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.

[4] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, 3rd ed., Academic Press, New York, 1979, SIAM, Philadelphia, 1994.

[5] P. E. Bjørstad and J. Mandel, *On the spectra of sums of orthogonal projections with applications to parallel computing*, BIT, 31 (1991), pp. 76–88.

[6] 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.

[7] 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.

[8] R. Bru, V. Migallón, J. Penadés, and D. B. Szyld, *Parallel, synchronous and asynchronous two-stage multisplitting methods*, Electron. Trans. Numer. Anal., 3 (1995), pp. 24–38.

[9] X.-C. Cai, C. Farhat, and M. Sarkis, *A minimum overlap restricted additive Schwarz preconditioner and applications to 3D flow simulations*, Contemp. Math., 218 (1998), pp. 479–485.

[10] X.-C. Cai and Y. Saad, *Overlapping domain decomposition algorithms for general sparse matrices*, Numer. Linear Algebra Appl., 3 (1996), pp. 221–237.

[11] X.-C. Cai and M. Sarkis, *A restricted additive Schwarz preconditioner for general sparse linear systems*, SIAM J. Sci. Comput., 21 (1999), pp. 792–797.

[12] T. F. Chan and T. P. Mathew, *Domain decomposition methods*, Acta Numerica (1994), pp. 61–143.

[13] 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.

[14] L. Collatz, *Aufgaben monotoner Art*, Arch. Math., 3 (1952), pp. 366–376.

[15] L. Collatz, *Functional Analysis and Numerical Mathematics*, Academic Press, New York, 1966.

[16] 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. Periaux, and O. B. Widlund, eds., SIAM, Philadelphia, 1990, pp. 3–21.

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

[18] A. Frommer and G. Mayer, *On the theory and practice of multisplitting methods in parallel computation*, Computing, 49 (1992), pp. 63–74.

[19] A. Frommer and B. Pohl, *A comparison result for multisplittings and waveform relaxation methods*, Numer. Linear Algebra Appl., 2 (1995), pp. 335–346.

[20] A. Frommer and H. Schwandt, *A unified representation and theory of algebraic additive Schwarz and multisplitting methods*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 893–912.

[21] A. Frommer, H. Schwandt, and D. B. Szyld, *Asynchronous weighted additive Schwarz methods*, Electron. Trans. Numer. Anal., 5 (1997), pp. 48–61.

[22] A. Frommer and D. B. Szyld, *H-splittings and two-stage iterative methods*, Numer. Math., 63 (1992), pp. 345–356.

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

[24] G. H. Golub and Q. Ye, *Inexact preconditioned conjugate gradient method with inner-outer iteration*, SIAM J. Sci. Comput., 21 (1999), pp. 1305–1320.

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

[26] E. W. Jenkins, R. C. Berger, J. P. Hallberg, S. E. Howington, C. T. Kelley, J. H. Schmidt, A. K. Stagg, and M. D. Tocci, *A two-level aggregation-based Newton-Krylov-Schwarz method for hydrology*, in Parallel Computational Fluid Dynamics, D. E. Keyes, A. Ecer, J. Periaux, and N. Satofuka, eds., North-Holland, Amsterdam, pp. 257–264.

[27] M. T. Jones and D. B. Szyld, *Two-stage multisplitting methods with overlapping blocks*, Numer. Linear Algebra Appl., 3 (1996), pp. 113–124.

[28] 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*, Math. Comp., 31 (1977), pp. 148–162.

[29] D. P. O'Leary and R. E. White, *Multi-splittings of matrices and parallel solution of linear systems*, SIAM J. Alg. Disc. Meth., 6 (1985), pp. 630–640.

[30] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

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

[32] 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*, Math. Comp., 27 (1973), pp. 139–145.

[33] Y. Saad, *A flexible inner-outer preconditioned GMRES algorithm*, SIAM J. Sci. Comput., 14 (1993), pp. 461–469.

[34] Y. Saad, *Iterative Methods for Sparse Linear Systems*, PWS, Boston, 1996.

[35] U. Schrader, *Invers-isotone Diskretisierungsmethoden für invers-isotone lineare und quasilin-*

eare Zwei-Punkt-Randwertaufgaben, Ph.D. thesis, Department of Mathematics, Technical University of Chemnitz, Chemnitz, Germany, 2000.

[36] 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.

[37] D. B. Szyld and J. A. Vogel, *A flexible quasi-minimal residual method with inexact preconditioning*, SIAM J. Sci. Comput., to appear.

[38] P. Vaněk, J. Mandel, and M. Brezina, *Algebraic multigrid by smoothed aggregation for second and fourth order elliptic problems*, Computing, 56 (1996), pp. 179–196.

[39] R. S. Varga, *Factorization and normalized iterative methods*, in Boundary Problems in Differential Equations, R. E. Langer, ed., The University of Wisconsin Press, Madison, WI, 1960, pp. 121–142.

[40] R. S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1962, 2nd ed., Springer, Berlin, 2000.

[41] R. E. White, *Multisplitting with different weighting schemes*, SIAM J. Matrix Anal. Appl., 10 (1989), pp. 481–493.

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