On Hybrid Multigrid-Schwarz algorithms

Sébastien Loisel, Reinhard Nabben,
and Daniel B. Szyld

Research Report 07-8-28
August 2007

On Hybrid Multigrid-Schwarz algorithms*

Sébastien Loisel† Reinhard Nabben‡ Daniel B. Szyld†

Abstract

J. Lottes and P. Fischer in [*J. Sci. Comput.*, 24:45–78, 2005] studied many smoothers or preconditioners for hybrid Multigrid-Schwarz algorithms for the spectral element method. The behavior of several of these smoothers or preconditioners are analyzed in the present paper. Here it is shown that the Schwarz smoother that best performs in the above reference, is equivalent to a special case of the weighted restricted additive Schwarz, for which convergence analysis is presented. For other preconditioners which do not perform as well, examples and explanations are presented illustrating why this behavior may occur.

Keywords: Weighted restricted additive Schwarz methods; domain decomposition; convergence analysis.

1 Introduction

Fischer and Lottes studied in [9, 12] the effective solution of linear systems originating from spectral element discretizations of Poisson, Helmholtz, and Navier-Stokes equations. They compared several Schwarz methods used as smoothers or preconditioners. In this paper, we analyze two of those methods. We show that these methods can be interpreted within the class of *weighted restrictive additive Schwarz methods* (WRAS), and we use this interpretation to analyze their convergence properties. In other words, in this paper we provide an understanding of why some of these smoothers or preconditioners work well in practice, as well as an indication of why others may not always be advantageous. In section 2 we briefly describe some of the methods presented in [9, 12], but we refer the reader to these references for full details, as well as for further bibliography. We emphasize that we are not presenting new methods here, but only providing some analysis which we believe help in the understanding of existing methods.

*This version dated 28 August 2007

†Department of Mathematics, Temple University (038-16), 1805 N. Broad Street, Philadelphia, Pennsylvania 19122-6094, USA ({loisel,szyld}@temple.edu). Supported in part by the U.S. Department of Energy under grant DE-FG02-05ER25672.

‡Institute für Mathematik, Technische Universität Berlin, D-10623 Berlin, Germany (nabben@math.tu-berlin.de).

Since extensive numerical results are reported in [9, 12], here we present few examples and experiments illustrating our theoretical developments.

Restrictive additive Schwarz methods (RAS) were introduced as efficient alternatives to the classical additive Schwarz preconditioner in [5]. For description of general Schwarz methods; see, e.g., [17, 20, 21]. 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 [4, 5]. RAS preconditioners are widely used in practice and are the default preconditioner in the PETSc software package [1]. Convergence of RAS methods was first shown in [11]; see also [14]. Some of these methods, including selected weighted variants, are reviewed in section 3.

Using the algebraic structure of the smoothers and preconditioners given in sections 2 and 3, an analysis of the convergence properties of some of the methods of [12, 9] is presented in section 4. For one of the methods, examples are given illustrating the fact that while in some cases it may be effective, in others, it might not.

2 Hybrid Multigrid-Schwarz

The linear system obtained from the spectral element discretization of some elliptic operator with n degrees of freedom can be described as

$$Au = g. \tag{1}$$

One of the most effective methods used in [12, 9] for the solution of (1) is a version of multigrid for spectral elements (see, e.g., [13, 18]), and the smoother proposed is a version of the overlapping additive Schwarz method of [7]; see also [3, 8, 17, 20, 21]. The most effective smoother in [12] is now described, in part using the algebraic representation from [10, 11].

The domain Ω of the elliptic operator is decomposed into p disjoint subdomains Ω_i, so that $\bar{\Omega} = \cup\bar{\Omega}_i$. Given a mesh for Ω, we can enlarge each Ω_i by a thickness of one element by adjoining all elements adjacent to $\partial\Omega_i$, obtaining $\Omega_{i,1}$. If we repeat this process δ times to obtain $\Omega_{i,\delta}$, we have an overlapping domain decomposition whose overlap has thickness 2δ elements. Typical values for δ are 1 or 2. Larger values of δ typically result in faster convergence per iteration, although each iteration is more expensive.

Given a piecewise polynomial function in the finite element space, it can be identified by sampling it at n points, called the *nodes* of the finite element method. For example, piecewise linear elements on triangles are identified by their values at the vertices of the triangles, so the vertices are usually taken to be the nodes of the piecewise linear finite element method. For piecewise biquadratic functions on rectangles, the nodes are usually situated at the corners, edge midpoints and centroids of each rectangle. When solving a homogeneous

Dirichlet problem, those vertices and nodes on $\partial\Omega$ are known to be zero so they are eliminated from the matrix problem. If the boundary condition is Neumann or Robin, these nodes are not eliminated.

For each subdomain, and each amount of overlap given by δ, we define a restriction matrix $R_{i,\delta}$. For Dirichlet problems, $R_{i,\delta}$ restricts to the $n_i = n_{i,\delta}$ nodes which are in the interior[1] of $\Omega_{i,\delta}$, and for Neumann or Robin problems, $R_{i,\delta}$ restricts to those nodes in the interior of $\Omega_{i,\delta}$ as well as some nodes on $\partial\Omega$. One can write explicitly

$$R_{i,\delta} = [I_{i,\delta}|O]\,\pi_i, \tag{2}$$

where $I_{i,\delta}$ the identity on $\mathbb{R}^{n_i}$ and π_i is a permutation in $\mathbb{R}^n$ which depends on the numbering of the nodes. We also define a restriction of the operator A on Ω_i as

$$A_{i,\delta} = R_{i,\delta} A R_{i,\delta}^T. \tag{3}$$

The additive Schwarz preconditioner of [7] is

$$M_{AS,\delta}^{-1} = \sum_{i=1}^{p} R_{i,\delta}^T A_{i,\delta}^{-1} R_{i,\delta}. \tag{4}$$

We point out that the matrix defined in (4) is indeed nonsingular, and therefore the notation for it is consistent; see, e.g., [10].[2]

We introduce the weighting matrices

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

In order to define two other Schwarz operators, Lottes and Fischer [12] introduce the diagonal counting matrix

$$C = C_\delta = \sum_{i=1}^{p} E_{i,\delta} = \sum_{i=1}^{p} R_{i,\delta}^T R_{i,\delta}, \tag{5}$$

where each diagonal entry $C_{jj} \geq 1$ indicates the number of subdomains sharing the node j. The diagonal weight matrix $W = W_\delta$ is defined as

$$W_{jj} = 1/C_{jj}, \quad j = 1, \ldots, n. \tag{6}$$

For nonsymmetric problems, one smoother used for multigrid in [12] was

$$M_W^{-1} = W M_{AS,\delta}^{-1}, \tag{7}$$

which we can call a *weighted additive Schwarz* smoother.

[1] Nodes on the boundary of $\Omega_{i,\delta}$ play no role because the Schwarz iteration is in fact solving Dirichlet problems on each subdomain.

[2] Note that in [12] this preconditioner is denoted by M_{Schwarz} (without the inverse).

Another approach for the solution of (1) considered in [12] is the use of a Richardson iteration of the form

$$u^{k+1} = u^k + \theta M^{-1}(g - Au^k), \tag{8}$$

for some fixed damping parameter $\theta > 0$, where M is, for example, either the additive Schwarz operator (AS) defined in (4) or the weighted additive Schwarz operator (WAS) defined in (7). It follows directly from (8) that the relevant quantity in the analysis of the convergence rate of the Richardson iteration is the spectral radius $\rho(I - \theta M^{-1}A)$. Thus, for Richardson with the AS smoother, one needs to study $\rho(I - \theta M_{AS,\delta}^{-1}A)$ and for for Richardson with the WAS smoother, $\rho(I - \theta M_W^{-1}A)$. As we shall see, in the latter case, one can choose $\theta = 1$, and therefore we study $\rho(I - M_W^{-1}A) = \rho(I - WM_{AS,\delta}^{-1}A)$.

For symmetric (positive definite) problems, the method of choice is preconditioned conjugate gradient (CG). In this case, a symmetric preconditioner is called for, and the one considered in [12] is

$$\sqrt{W}M_{AS,\delta}^{-1}\sqrt{W}, \tag{9}$$

where the expression $\sqrt{W}$ refers to a diagonal matrix whose entries are $1/\sqrt{C_{jj}}$. The relevant quantity for the convergence of preconditioned CG here is then the condition number $\kappa(\sqrt{W}M_{AS,\delta}^{-1}\sqrt{W}A)$.

3 Restricted additive Schwarz

We present here an algebraic (matricial) representation of *restricted additive Schwarz methods*, following [11]. We introduce 'restricted' operators $\widetilde{R}_{i,\delta}$ as

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

The image of $\widetilde{R}_{i,\delta}^T = E_{i,0}R_{i,\delta}^T$ can be identified with $\Omega_{i,0} = \Omega_i$, and we can say that $\widetilde{R}_{i,\delta}^T$ 'restricts' $R_{i,\delta}^T$ in the sense that the image of the latter, corresponds to nodes in $\Omega_{i,\delta} \supset \Omega_i$. That is, $\widetilde{R}_{i,\delta}^T$ 'restricts' the nodes from the subdomain of the overlapping decomposition to those of the non-overlapping decomposition. The restricted additive Schwarz method from [4, 5] replaces the prolongation operator $R_{i,\delta}^T$ with $\widetilde{R}_{i,\delta}^T$ in (4), i.e., one uses[3]

$$M_{RAS,\delta}^{-1} = \sum_{i=1}^{p} \widetilde{R}_{i,\delta}^T A_{i,\delta}^{-1} R_{i,\delta}. \tag{10}$$

For practical parallel implementations, when each subdomain corresponds to a different processor, replacing $R_{i,\delta}^T$ by $\widetilde{R}_{i,\delta}^T$ means that the corresponding part of

[3] We note that the representations (4) and (10) using rectangular matrices $R_{i,\delta}$ and matrices $A_{i,\delta}$ of smaller size is consistent with the standard literature [6, 17, 20], and different than that of [5] where $n \times n$ matrices are used.

the computation does not require any communication, since the images of the $\widetilde{R}_{i,\delta}^T$ do not overlap. In addition, the numerical results in [5] 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 with RAS one looses symmetry, since if A is symmetric, $M_{AS,\delta}^{-1}$ is symmetric as well, whereas $M_{RAS,\delta}^{-1}$ is usually nonsymmetric. A symmetric version of RAS is possible, namely, the *restricted additive Schwarz with harmonic extension* (RASH) as follows

$$M_{RASH,\delta}^{-1} = \sum_{i=1}^{p} \widetilde{R}_{i,\delta}^T A_{i,\delta}^{-1} \widetilde{R}_{i,\delta}. \tag{11}$$

Note however that this version was shown to be inferior to RAS in several examples, and in fact it may not even converge [5, 10].

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

$$\sum_{i=1}^{p} E_{i,0} = I, \tag{12}$$

while for additive Schwarz we have

$$qI \geq C = \sum_{i=1}^{p} E_{i,\delta} \geq I, \tag{13}$$

where the inequalities are componentwise and

$$q = \max_j C_{jj} \tag{14}$$

is the maximum number of subdomains to which each node of the mesh belongs.

We briefly review now some convergence results for Richardson preconditioned by additive Schwarz (which can also be seen as additive Schwarz accelerated by Richardson), and for RAS; see [2, 10, 11] for proofs and further details. These results are given for nonsingular M-matrices, i.e., matrices with nonpositive off-diagonal elements and whose inverse are (componentwise) nonnegative. Thus, they apply in particular to Stieltjes matrices, i.e., symmetric positive definite M-matrices. (Results for symmetric positive semidefinite matrices can be found in [15]).

Theorem 3.1. *Let A be an M-matrix. If $\theta \leq 1/q$, then*

$$\rho(I - \theta M_{AS,\delta}^{-1} A) < 1,$$

where q is defined in (14). *Furthermore, one has that*

$$\rho(I - M_{RAS,\delta}^{-1} A) < 1.$$

The fact that no damping parameter θ is necessary for the convergence of RAS, can be traced directly to the observation (12) above, while the need for damping in Richardson preconditioned by additive Schwarz stems from (13). These observations lead to the following two particularly efficient variants of RAS presented in [5], and analyzed and extended in [11].

We introduce weighted restriction operators $R^{\omega}_{i,\delta}$ which result from $R_{i,\delta}$ of (2) by replacing the entry 1 in column $\pi_i(j)$ by "weights" $\omega_{i,j} > 0$, $j = 1, \ldots, n_i$ (i.e., whenever a column of $R_{i,\delta}$ is not a zero column), where the weights have to satisfy $\sum_i \omega_{i,j} = 1$. From this definition, it follows directly that

$$\sum R^T_{i,\delta} R^{\omega}_{i,\delta} = I. \tag{15}$$

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

$$M^{-1}_{WRAS,\delta} = \sum_{i=1}^{p} (R^{\omega}_{i,\delta})^T A^{-1}_{i,\delta} R_{i,\delta} \quad \text{and} \quad M^{-1}_{WASH,\delta} = \sum_{i=1}^{p} R^T_{i,\delta} A^{-1}_{i,\delta} R^{\omega}_{i,\delta}. \tag{16}$$

We postpone a result on the convergence of these preconditioners until the next section.

4 Analysis of weighted Schwarz smoothers

We begin with our analysis of the weighted additive Schwarz smoother (7) by relating it to the methods of the last section. We show first that this smoother can be interpreted as a WRAS smoother as in (16), with a particular choice of weighted restriction operator. This choice is $R^{\omega}_{i,\delta} = WR_{i,\delta}$, and with this choice we have

$$\begin{aligned} M^{-1}_{W} = WM^{-1}_{AS} = W \sum_{i=1}^{p} R^T_{i,\delta} A^{-1}_{i,\delta} R_{i,\delta} &= \sum_{i=1}^{p} (WR^T_{i,\delta}) A^{-1}_{i,\delta} R_{i,\delta} \\ &= \sum_{i=1}^{p} (R^{\omega}_{i,\delta})^T A^{-1}_{i,\delta} R_{i,\delta}. \end{aligned} \tag{17}$$

Furthermore, with the choice of $R^{\omega}_{i,\delta} = WR_{i,\delta}$, it follows from (6), (5), and (12), that (15) holds. Therefore, we can hope to apply the following convergence and comparison result from [11] to this smoother.

Theorem 4.1. *Let A be an M-matrix, and let*

$$R^{\omega}_{i,\delta} \geq \theta R_{i,\delta}, \text{ for } i = 1, \ldots, p. \tag{18}$$

Then,

$$\begin{aligned} \rho(I - M^{-1}_{WRAS,\delta}A) &\leq \rho(I - \theta M^{-1}_{AS,\delta}A) \\ \textit{and} \quad \rho(I - M^{-1}_{WASH,\delta}A) &\leq \rho(I - \theta M^{-1}_{AS,\delta}A). \end{aligned}$$

We show next that this theorem applies in particular to $R^{\omega}_{i,\delta} = WR_{i,\delta}$, and thus to the smoother (7). This follows directly from the definitions, since $W_{jj} = 1/C_{jj} \geq 1/q$, and thus for the range of values of θ that one can guarantee convergence (see Theorem 3.1) we have that $R^{\omega}_{i,\delta} = WR_{i,\delta} \geq (1/q)R_{i,\delta} \geq \theta R_{i,\delta}$, i.e., the hypothesis (18) holds.

Theorem 4.1 shows that the weighted additive Schwarz variants, as well the weighted additive Schwarz smoother, converge faster than Richardson preconditioned by the classical additive Schwarz method. In terms of performance, this is consistent with the experiments in [5] and in [12] (although there, the matrices are not always M-matrices or Stieltjes matrices). In our analysis this is to be attributed to the fact that for the weighted variants we have (15), while for Richardson preconditioned by additive Schwarz we have (13), which implies

$$\theta \sum_{i=1}^{p} E_{i,\delta} \leq I.$$

In other words, Theorem 4.1, together with our interpretation based on (17), explains why one would expect (at least for Stieltjes matrices) that the performance of the weighted additive Schwarz smoother (7) be superior to the performance of Richardson iteration with the additive Schwarz method (4).

Remark 4.2. *We have so far only mentioned one-level Schwarz methods. In practice, for preconditioners, it is customary to have a second level, usually a coarse-grid correction; see, e.g., [20, 21]. In terms of the description of the methods here, one has a 'global' restriction operator R_0 with nonzeros corresponding to nodes in each of the subdomains, and the corresponding $A_0 = R_0 A R_0^T$, as in* (3). *The multiplicative second level operator is $I - R_0^T A_0^{-1} R_0 A$. When it is combined, e.g., with WRAS, one has the two level operator*

$$(I - R_0^T A_0^{-1} R_0 A)(I - M^{-1}_{WRAS,\delta} A).$$

It was shown in [2, Section 7] that the use of the multiplicative second level can indeed improve the spectral radii of the operators discussed in Theorems 3.1 and 4.1, and under certain hypotheses on the matrix A, the spectral radii of these operators never deteriorates by the application of the multiplicative second level. Beyond this remark, we will not consider the second level methods in our study of the smoothers or preconditioners here.

We proceed now to analyze the symmetric preconditioner (9). We begin by noting that inspired by (11), one can introduce a *weighted restricted additive Schwarz preconditioner with harmonic extension* (WRASH) as follows

$$M^{-1}_{WRASH,\delta} = \sum_{i=1}^{p} (\widetilde{R}^{\omega}_{i,\delta})^T A^{-1}_{i,\delta} \widetilde{R}^{\omega}_{i,\delta}, \tag{19}$$

where $\widetilde{R}^{\omega}_{i,\delta}$ is such that

$$\sum_{i=1}^{p} (\widetilde{R}^{\omega}_{i,\delta})^T \widetilde{R}^{\omega}_{i,\delta} = I. \tag{20}$$

We can now interpret (9) as the WRASH operator (19) by setting $\widetilde{R}^{\omega}_{i,\delta} = \sqrt{W} R_{i,\delta}$. Indeed,

$$\sqrt{W} M^{-1}_{AS,\delta} \sqrt{W} = \sum_{i=1}^{p} \sqrt{W} R^{T}_{i,\delta} A^{-1}_{i,\delta} R_{i,\delta} \sqrt{W} = \sum_{i=1}^{p} (\widetilde{R}^{\omega}_{i,\delta})^T A^{-1}_{i,\delta} \widetilde{R}^{\omega}_{i,\delta}.$$

Note also that with this choice, (20) holds.

We already mentioned in the previous section that RASH is not usually recommended. Then, it would be natural to think that the symmetric weighted preconditioner WRASH may not always perform well either. We show below using a series of examples that while in some cases, WRASH performs better than additive Schwarz, in others, the opposite is true. Since these are considered preconditioners for CG, these comparisons are obtained for the appropriate condition numbers. We also show that if one thinks of WRASH as a smoother for multigrid, for these examples, a comparison of the appropriate spectral radii indicates that WRAS is a superior choice.

The examples we show are small, so that we can easily study the spectral properties of the operators. These spectral properties illustrate some of the convergence behavior observed in the extensive experiments reported in [9, 12].

In the examples below, $\sigma(A)$ denotes the spectrum of A, i.e., set of its eigenvalues, and S_i denotes the set of nodes in $\Omega_{i,\delta}$. We begin with an example of a symmetric positive definite M-matrix arising from a discretization of a simple one-dimensional partial differential equation (PDE).

Example 4.3. Consider the matrix

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

and consider $S_1 = \{1, 2, 3\}$ and $S_2 = \{3, 4, 5\}$, so that $C = diag(1, 1, 2, 1, 1)$. The spectrum $\sigma(A) = \{0.2679, 1.0, 2.0, 3.0, 3.7321\}$ and thus $\kappa(A) = 13.9309$. For additive Schwarz, we have $\sigma(M^{-1}_{AS}A) = \{1.0, 0.5, 1.5, 2.0, 1.0\}$ and thus $\kappa(M^{-1}_{AS}A) = 4.000$. For the symmetric weighted RAS, we have $\sigma(M^{-1}_{WRASH}A) = \{1.0, 0.4342, 1.5, 1.1516, 1.0\}$, and thus $\kappa(M^{-1}_{WRASH}A) = 2.6522$. We conclude that for this Stieltjes matrix, we have that $\kappa(M^{-1}_{WRASH}A) < \kappa(M^{-1}_{AS}A) < \kappa(A)$, i.e., WRASH is a better preconditioner than the classical additive Schwarz.

We further compute $\rho(I - \frac{1}{2}M^{-1}_{AS}A) = 0.75$, i.e., using $\theta = 1/q$, while $\rho(I - M^{-1}_{WRASH}A) = 0.5658$. This is then consistent with the previous observations on the condition numbers and with Theorem 4.1. On the other hand, observe that for the WRAS preconditioner, we obtain $\rho(I - \tilde{M}^{-1}_{WRAS}A) = 0.50$, illustrating that as a smoother WRAS is superior to both the classical additive Schwarz and WRASH.

Example 4.4. We now consider a more realistic version of Example 4.3, namely a two-dimensional PDE discretized using the spectral element method (SEM) as in [9, 12]. The SEM is a Finite Element Method (FEM) whose basis functions are piecewise polynomial of high degree; see, e.g., [16]. The elements are rectangular, and on each element, a basis function is given by $\phi = \ell_1(x)\ell_2(y)$, where ℓ_1 and ℓ_2 are Legendre polynomials with nodes at the Gauss-Lobatto points, which ensures that the basis functions are continuous at element boundaries[4]. To assemble the stiffness matrix, one must compute integrals of the form $\int_S \nabla\phi_j \cdot \nabla\phi_k \, dx\, dy$, where S is a rectangular element and ϕ_j, ϕ_k are two basis functions. The integral is replaced by a quadrature rule. Since the nodes are at the Gauss-Lobatto points, using the Gauss-Lobatto quadrature rule results in a formula which is easy to compute and which is also highly accurate.

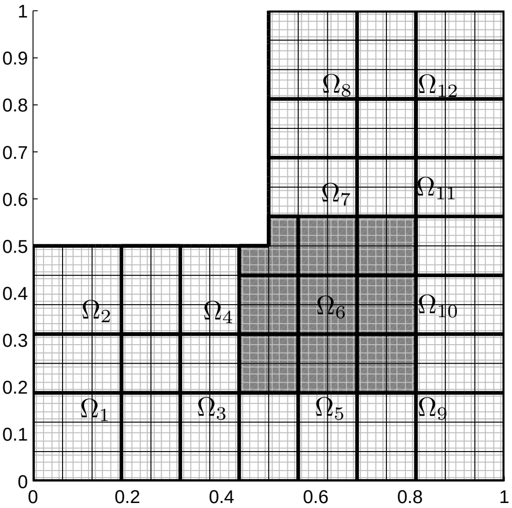

Figure 1: A domain decomposition of an L-shaped domain. Subdomain Ω_6 is highlighted in grey. It consists of 35 elements and its interior has 816 degrees of freedom.

We have meshed an L-shaped domain with 192 square elements of side 1/16; see Figure 1. Each element further comprises 6×6 nodes, so that the basis functions are piecewise biquintic. We are solving the problem

$$\begin{cases} -\Delta u &= f \text{ in } \Omega, \\ u &= 0 \text{ on } \partial\Omega. \end{cases}$$

We have decomposed this mesh of 192 elements into 12 subdomains as shown in

[4]Were we to use the Gauss nodes instead, adjacent elements would not share nodes and it would be hard to enforce continuity.

Figure 1. Each subdomain is created by starting with a 4×4 block of elements, and then adding all elements adjacent to the boundary.

Since we are solving a Dirichlet problem, the nodes on the boundary of Ω are not present in the discrete problem. Likewise, the restriction matrices R_i, $i = 1, ..., 12$, contain one row per node inside of Ω_i, the vertices on $\partial\Omega_i$ not being involved. The total number of degrees of freedom is 4641, so that A is a 4641×4641 matrix. This is a symmetric positive definite matrix, but not an M-matrix, i.e., it is not Stieltjes. This is the same situation as the matrices in [9, 12]. The 4641 nodes have coordinates $(x_1, y_1), ..., (x_{4641}, y_{4641}) \in \Omega$. The weights are chosen as points of a piecewise linear functions of the form

$$\omega_{i,j} = \max\{0, \min(1, 1 - 4/3\|(x_j, y_j) - (a_i, b_i)\|_\infty)\},$$

with $(a_i, b_i) \in \{0.5, 1.5, 2.5, 3.5\}^2$ corresponding to the center of the Ω_i subdomain, for $i = 1, 2, ..., 12$, which we then normalize for WRAS or WASH, so that $\sum_{i=1}^{12} \omega_{i,j} = 1$, or for WRASH so that $\sum_{i=1}^{12} \omega_{i,j}^2 = 1$; cf. (15) and (20).

The largest eigenvalue of A is 3.9967 and its smallest eigenvalue is 0.0016, so its condition number is 2655.85. We summarize our results in table form.

M^{-1}	$\kappa(M^{-1}A)$	θ	$\rho(I - \theta M^{-1}A)$
M_{AS}^{-1}	48.8733	1/4	0.9272
M_{WRAS}^{-1}	18.9619	1	0.7566
M_{WASH}^{-1}	59.1983	1	0.7566
M_{WRASH}^{-1}	15.1858	1	0.8121

As in Example 4.3, we have that M_{WRASH}^{-1} and M_{WRAS}^{-1} are better preconditioners than M_{AS}^{-1}, but M_{WASH}^{-1} is worse than M_{AS}^{-1}. The best smoothers are M_{WRAS}^{-1} and M_{WASH}^{-1} (cf. Theorem 4.1), followed by M_{WRASH}^{-1} and finally by M_{AS}^{-1}.

Example 4.5. This example is taken from [19]. Let

$$A = \begin{bmatrix} 2 & 13 & 18 & 1 & -1 \\ 13 & 102 & 139 & 11 & -6 \\ 18 & 139 & 191 & 15 & -9 \\ 1 & 11 & 15 & 2 & 0 \\ -1 & -6 & -9 & 0 & 2 \end{bmatrix},$$

which, as in the previous example, is a symmetric positive definite matrix, but not an M-matrix. Consider $S_1 = \{1, 2\}$, $S_2 = \{2, 3, 4\}$, and $S_3 = \{4, 5\}$, so that $C = diag(1, 2, 1, 2, 1)$, and $\sigma(A) = \{0.0057, 0.3876, 0.6998, 2.1881, 295.7187\}$, so that $\kappa(A) = 51180$. For additive Schwarz, we have

$$\sigma(M_{AS}^{-1}A) = \{2.9190, 2.4413, 1.0729, 0.5579, 0.0089\}$$

so that $\kappa(M_{AS}^{-1}A) = 327.9$. Turning to WRASH, we have $\sigma(M_{WRASH}^{-1}A) = \{13.3352, 1.6284, 0.6196, 0.0058, 0.1209\}$, so that $\kappa(M_{WRASH}^{-1}A) = 110.3$, indicating that WRASH a better symmetric preconditioner than additive Schwarz.

As a smoother we have $\rho(I - \frac{1}{2}M_{AS}^{-1}A) = 0.9956$. For WRAS, we have $\rho(I - M_{WRAS}^{-1}A) = 1.6308$, i.e., not convergent as an iteration. For curiosity we computed $\rho(I - \frac{1}{2}M_{WRAS}^{-1}A) = 0.9942$, i.e., if accelerated with Richardson with parameter $\theta = 1/2$. We also computed $\rho(I - M_{WRASH}^{-1}A) = 12.3352$ (i.e., $\theta = 1$). The following are values of the spectral radii for decreasing values of $\theta = 1/2, 1/3, 1/5, 1/6, 1/8$: $5.667, 3.445, 1.667, 1.2225, 0.9993$. In other words, for Richardson iterations, the damped classical additive Schwarz method is the best among the methods considered.

Example 4.6. Finally, consider the same matrix as in Example 4.5, but use now $S_1 = \{1, 2, 3, 4\}$ and $S_2 = \{2, 3, 4, 5\}$, so that $C = diag(1, 2, 2, 2, 1)$. In this case we have for additive Schwarz $\sigma(M_{AS}^{-1}A) = \{0.0619, 1.9381, 2.0, 2.0, 2.0\}$, so that $\kappa(M_{AS}^{-1}A) = 32.31$ (lower than the case with less overlap, as expected). For WRASH, we have $\sigma(M_{WRASH}^{-1}A) = \{3.5434, 0.0254, 1.4743, 0.9054, 1.0\}$, which gives $\kappa(M_{WRASH}^{-1}A) = 139.5$ implying that it is a worse preconditioner than additive Schwarz!

For completeness, we report that $\rho(I - \frac{1}{2}M_{AS}^{-1}A) = 0.9691$ and $\rho(I - M_{WRAS}^{-1}A) = 0.9381$, while $\rho(I - M_{WRASH}^{-1}A) = 2.5434$. Thus, WRAS would be a better smoother than damped additive Schwarz, and WRASH would not work at all as a smoother.

5 Conclusions

We have used an algebraic representation to analyze certain Schwarz methods presented in the literature. We were able to explain why for certain class of problems the weighted restricted additive Schwarz method works better than classical additive Schwarz, when used as a smoother accelerated by Richardson iterations.

We show examples where a symmetric weighted additive Schwarz preconditioner has worse performance than classical additive Schwarz, and this is consistent with other cases in the literature for which such preconditioners are not recommended.

Thus, we have illustrated how the algebraic analysis of Schwarz methods provides a useful tool in understanding their behavior. Similarly, this approach can also be used to inspire the design of new Schwarz smoothers or preconditioners.

Acknowledegement

We would like to thank Lothar Reichel and Olof B. Widlund for their comments on an earlier version of this paper, which have helped improve our presentation.

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, April 1998. Available at `http://www.mcs.anl.gov/petsc`.

[2] M. Benzi, A. Frommer, R. Nabben, and D. B. Szyld, Algebraic theory of multiplicative Schwarz methods. *Numerische Mathematik*, 89:605-639, 2001.

[3] P. E. Bjørstad and O. B. Widlund. To overlap or not to overlap: A note on a domain decomposition method for elliptic problems. *SIAM Journal on Scientific and Statistical Computing*, 10:1053–1061, 1989.

[4] X.-C. Cai, C. Farhat and M. Sarkis. A minimum overlap restricted additive Schwarz preconditioner and applications to 3D flow simulations. *Contemporary Mathematics*, 218:479–485, 1998.

[5] X.-C. Cai and M. Sarkis. A restricted additive Schwarz preconditioner for general sparse linear systems. *SIAM Journal on Scientific Computing*, 21:792–797, 1999.

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

[7] Maksymilian Dryja and Olof B. Widlund. An additive variant of the Schwarz alternating method for the case of many subregions. Technical Report 339, also Ultracomputer Note 131, Department of Computer Science, Courant Institute, New York University, 1987.

[8] M. Dryja and O. B. Widlund. Towards a Unified Theory of Domain Decomposition Algorithms for Elliptic Problems. In T. Chan, R. Glowinski, J. Périaux, and O. Widlund, editors, *Third International Symposium on Domain Decomposition Methods for Partial Differential Equations*, (held in Houston, Texas, March 20-22, 1989), pages 3–21, Philadelphia, 1989. SIAM.

[9] P. F. Fischer and J. .W. Lottes. Hybrid Schwarz-Multigrid methods for the spectral element method: Extensions to Navier-Stokes. In *Proceedings of the Fifteenth International Conference on Domain Decomposition Methods, Berlin, Germany, 2003*, R. Kornhuber, R. H. W. Hoppe, J. Périaux, O. Pironneau, O. B. Widlund, and J. Xu, eds. Lecture notes in Computational Science and Engineering, vol. 40, Springer, Berlin and Heidelberg, 2005, pp. 35–49.

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

[11] A. Frommer and D. B. Szyld. An algebraic convergence theory for restricted additive Schwarz methods using weighted max norms. *SIAM Journal on Numerical Analysis*, 39:463-479, 2001.

[12] J. .W. Lottes and P. F. Fischer. Hybrid Multigrid/Schwarz algorithms for the spectral element method. *Journal on Scientific Computing*, 24:45–78, 2005.

[13] Y. Maday, R. Muñoz, A.T. Patera, and E.M. Rønquist. Spectral element multigrid methods. In P. de Groen and R. Beauwens, eds., *Proc. of the IMACS Int. Symposium on Iterative Methods in Linear Algebra, Brussels, 1991.* Elsevier, Amsterdam, pp. 191–201.

[14] R. Nabben and D. B. Szyld. Convergence theory of restricted multiplicative Schwarz methods. *SIAM Journal on Numerical Analysis*, 40:2318–2336, 2003.

[15] R. Nabben and D. B. Szyld. Schwarz iterations for symmetric positive semidefinite problems. *SIAM Journal on Matrix Analysis and Applications*, 29:98–116, 2006.

[16] A. T. Patera. A spectral element method for fluid dynamics - Laminar flow in a channel expansion. *Journal of Computational Physics*, 54:468–488, 1984.

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

[18] E.M. Rønquist and A.T. Patera. Spectral element multigrid, I: Formulation and numerical results. *Journal of Scientific Computing*, 2:389–406, 1987.

[19] M. Schnitker. *Eine algebraische Konvergenztheorie der Schwarz-Verfahren für symmetrisch positiv definite Matrizen* (An algebraic convergence theory of the Schwarz method for symmetric positive definite matrices). Schriftliche Hausarbeit der 1. Staatspreufung (undergraduate thesis). Universität Bielefeld, Bielefeld, Germany, 2002. In German.

[20] 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.

[21] A. Toselli and O. B. Widlund. *Domain Decomposition: Algorithms and Theory.* Springer-Verlag, Heidelberg, 2004.