33 RASHO: A Restricted Additive Schwarz Preconditioner with Harmonic Overlap

X.-C. Cai¹, M. Dryja², M. Sarkis³

Introduction

A restricted additive Schwarz (RAS) preconditioning technique was introduced recently for solving general nonsymmetric sparse linear systems [BGMS97, CFS98, CS99, FS01, GKK+00, LSHF01, SK00, QV99]. The RAS preconditioner improves the classical additive Schwarz preconditioner (AS), [SBG96], in the sense that it reduces the number of iterations of the iterative method, such as GMRES, and also reduces the communication cost per iteration when implemented on distributed memory computers. However, RAS in its original form is a nonsymmetric preconditioner and therefore cannot be used with the Conjugate Gradient method (CG). In this paper, we provide an extension of RAS for symmetric positive definite problems using the so-called harmonic overlaps (RASHO). Both RAS and RASHO outperform their counterparts of the classical additive Schwarz variants. Roughly speaking, the design of RASHO is based on a deeper understanding of the behavior of Schwarz type methods in the overlapping regions, and in the construction of the overlap. Under RASHO, the overlap is obtained by extending the nonoverlapping subdomains only in the directions that do not cut the boundaries of other subdomains, and all functions are made harmonic in the overlapping regions. As a result, the subdomain problems in RASHO are smaller than those of AS, and the communication cost is also smaller when implemented on distributed memory computers, since the right-hand sides of discrete harmonic systems are always zero which does not need to be communicated. We will show numerically that the RASHO preconditioned CG takes fewer iterations than the corresponding AS preconditioned CG. An almost optimal convergence theory will be presented for the RASHO for elliptic problems discretized with a finite element method.

Recall that the basic building blocks of classical Schwarz type algorithms are the operations of the form $(R_i^e)^T(A_i^e)^{-1}R_i^e$, where $A_i^e$ is the subdomain matrix and $R_i^e$ is the restriction operator for the extended subdomain (formal definitions will be given later in the paper). The multiplication of the such an operator with a vector, $v$, is realized by solving the linear system

$$A_i^e w = R_i^e v$$

(1)

on each extended subdomain. The key idea of RAS is that equation (1) is replaced by

$$A_i^f w = \begin{cases} v & \text{inside the unextended subdomain} \\ 0 & \text{in the overlapping part of the subdomain} \end{cases}$$

(2)

¹Department of Computer Science, University of Colorado, Boulder, CO 80309, (cai@cs.colorado.edu). The work was supported in part by the NSF grants ASC-9457534, ECS-9725504, and ACI-0072089

²Faculty of Mathematics, Informatics and Mechanics, Warsaw University, Warsaw, (dryja@mimuw.edu.pl). This work was supported in part by the NSF grant CCR-9732208 and in part by the Polish Science Foundation grant 2 P03A 021 16

³Mathematical Sciences Department, Worcester Polytechnic Institute, Worcester, MA 01609, (msarkis@wpi.edu). The work was supported in part by the NSF grant CCR-9984404
Note that the solution of (2) is discrete harmonic in the overlapping part of the subdomain, and therefore carries minimum energy in some sense. In this paper, we further explore the idea of “harmonic overlap” and at the same time keep the symmetry of the preconditioner.

The algorithm to be discussed below is applicable for symmetric positive definite problems. In order to provide a complete mathematical analysis, we restrict ourselves to the Poisson problem discretized with a finite element method. We consider a simple variational problem: Find \( u \in H_0^1(\Omega) \), such that

\[
 a(u, v) = f(v), \quad \forall \, v \in H_0^1(\Omega),
\]

where

\[
 a(u, v) = \int_{\Omega} \nabla u \cdot \nabla v \, dx \quad \text{and} \quad f(v) = \int_{\Omega} f(v) \, dx \quad \text{for} \quad f \in L^2(\Omega).
\]

For simplicity, let \( \Omega \) be a bounded polygonal region in \( \mathbb{R}^2 \) with a diameter of size \( O(1) \). The extension of the algorithm and results to \( \mathbb{R}^3 \) can be carried out easily. Let \( \mathcal{T}^h(\Omega) \) be a shape regular, quasi-uniform triangulation, of size \( O(h) \), of \( \Omega \) and \( \mathcal{V}(\Omega) \subset H_0^1(\Omega) \) the finite element space consisting of continuous piecewise linear functions associated with the triangulation. We are interested in solving the following discrete problem associated with (3): Find \( u^* \in \mathcal{V} \) such that

\[
 a(u^*, v) = f(v), \quad \forall \, v \in \mathcal{V}.
\]

Using the standard basis functions, (4) can be rewritten as a linear system of equations

\[
 Au^* = f.
\]

For simplicity, we understand \( u^* \) and \( f \) both as functions and vectors depending on the situation.

**Notations**

Let \( n \) be the total number of interior nodes of \( \mathcal{T}^h(\Omega) \) and \( W \) the set of nodes. We assume that a node-based partitioning has been applied and resulted in \( N \) nonoverlapping subsets \( W_i^0, i = 1, \ldots, N \), whose union is \( W \). For each \( W_i^0 \), we define a region \( \Omega_i^R \) as the union of all elements of \( \mathcal{T}^h(\Omega) \) that have all three vertices on \( W_i^0 \cup \partial \Omega \). We denote \( H \) as the representative size of the subregion \( \Omega_i^R \). We define the overlapping partition of \( W \) as follows. Let \( \{W_i^1\} \) be the one-overlap partition of \( W \), where \( W_i^1 \supseteq W_i^0 \) is obtained by including all the immediate neighboring vertices of the vertices in \( W_i^0 \). Using the idea recursively, we can define a \( \delta \)-overlap partition \( W = \bigcup_{i=1}^N W_i^\delta \). \( \delta h \) is approximately the extend of the extension.

We next define a subregion of \( \Omega \) induced by a set of nodes of \( \mathcal{T}^h(\Omega) \) as follows. Let \( Z \) be a subset of \( W \). The induced subregion, denoted as \( \Omega(Z) \), is defined as the union of: (1) the set \( Z \) itself; (2) the union all the open elements (triangles) of \( \mathcal{T}^h(\Omega) \) that have at least one vertex in \( Z \); and (3) the union of the open edges of these triangles that have at least one endpoint as a vertex of \( Z \). Note that \( \Omega(Z) \) is always an open region. The extended region \( \Omega_i^\delta \) is defined as \( \Omega(W_i^\delta) \). We introduce the subspace

\[
 \mathcal{V}_i^\delta \equiv \mathcal{V} \cap H_0^1(\Omega_i^\delta) \text{ extended by zero to } \Omega \setminus \Omega_i^\delta.
\]
It is easy to check that
\[ V = V_1^\delta + V_2^\delta + \cdots + V_N^\delta. \]

This decomposition is used in defining the classical one-level additive Schwarz algorithm without a coarse space [SBG96]. Let us define \( P_i^\delta : V \to V_i^\delta \) by
\[ a(P_i^\delta u, v) = a(u, v), \quad \forall u \in V, \quad \forall v \in V_i^\delta. \quad (6) \]

Then, the classical one-level additive Schwarz operator has the form
\[ P^\delta = P_1^\delta + \cdots + P_N^\delta. \]

Let \( \Gamma_i^\delta = \partial \Omega_i^\delta \setminus \partial \Omega \); i.e., the part of the boundary of \( \Omega_i^\delta \) that does belong to the Dirichlet part of the boundary. We define the interface overlapping boundary \( \Gamma^\delta \) as the union of all \( \Gamma_i^\delta \); i.e., \( \Gamma^\delta = \cup_{i=1}^N \Gamma_i^\delta \). We then define the following subsets of \( W \):

- \( W^I \equiv W \cap \Gamma^\delta \) (interface nodes)
- \( W_i^I \equiv W_i^I \cap W_i^\delta \) (local interface nodes)
- \( W_i^I_{,in} \equiv W_i^I \cap W_i^0 \) (local internal interface nodes)
- \( W_i^I_{,cut} \equiv W_i^I \setminus W_i^I_{,in} \) (local cut interface nodes)
- \( W_i^\delta_{,out} \equiv (W_i^\delta \setminus W_i^I) \cap (\bigcup_{j \neq i} W_j^\delta) \) (local overlapping nodes)
- \( W_i^\delta_{,non} \equiv W_i^\delta \setminus (W_i^I \cup W_i^\delta_{,out}) \) (local nonoverlapping nodes)
- \( W_i^\delta_{,in} \equiv W_i^\delta_{,non} \cup W_i^I_{,in} \) (internal nodes)

We note that the notions of subdomains, harmonic overlaps, the classification of nodal points can all be defined in terms of the graph of the sparse matrix.

We frequently use functions that are discrete harmonic at certain nodes. Let \( x_k \in W \) be a mesh point and \( \phi_{x_k}(x) \in V \) the finite element basis function associated with \( x_k \); i.e., \( \phi_{x_k}(x_k) = 1 \), and \( \phi_{x_k}(x_j) = 0, j \neq k \). We say \( u \in V \) is discrete harmonic at \( x_k \) if \( a(u, \phi_{x_k}) = 0 \). If \( u \) is discrete harmonic at a set of nodal points \( Z \), we say \( u \) is discrete harmonic in \( \Omega(Z) \).

Our new algorithm will be built on \( \bar{V}_i^\delta \) defined as a subspace of \( V_i^\delta \). \( \bar{V}_i^\delta \) consists of all functions in \( V_i^\delta \) that vanish on \( W_i^I_{,cut} \) and discrete harmonic at the nodes \( W_i^\delta_{,out} \). Note that the support of the subspace \( \bar{V}_i^\delta \) is
\[ \bar{W}_i^\delta \equiv W_i^\delta \setminus W_i^I_{,cut} \]
and, since the values at the harmonic nodes are not independent, they can not be counted toward the degree of freedoms. The dimension of \( \bar{V}_i^\delta \) is \( \text{dim} \left( \bar{V}_i^\delta \right) = |W_i^I_{,in}| \). Let \( \Omega_i^\delta \equiv \Omega(\bar{V}_i^\delta) \) be the induced domain. It is easy to see that \( \Omega_i^\delta \) is the same as \( \Omega_i^\delta \) but with cuts. We have then \( \bar{V}_i^\delta = V \cap H_0^1(\Omega_i^\delta) \) and discrete harmonic on \( \Omega_i^\delta \equiv \Omega(W_i^\delta_{,out}) \). We define \( \bar{V}_i^\delta \subset V_i^\delta \) as
\[ \bar{V}_i^\delta = \bar{V}_1^\delta + \cdots + \bar{V}_N^\delta, \]
which is a direct sum. We remark that functions in $\tilde{V}^\delta$ are, by definition, the sum of functions $u_i \in \tilde{V}^\delta_i$, $i = 1, \cdots, N$. Functions in $\tilde{V}^\delta$ can, in fact, be characterized easily as in the following lemma.

**Lemma 1** [CDS01] If $u \in \mathcal{V}$ and $u$ is discrete harmonic at all the overlapping nodes, i.e., on $\cup_{i=1}^N W^\delta_{i,\text{out}}$, then $u \in \tilde{V}^\delta$.

**RAS with Harmonic Overlap**

Let $\tilde{P}_i^\delta : \tilde{V}^\delta \to \tilde{V}^\delta_i$ be a projection operator satisfying

$$ a(\tilde{P}_i^\delta u, v) = a(u, v), \forall u \in \tilde{V}^\delta, \forall v \in \tilde{V}^\delta_i. \tag{7} $$

The RASHO operator can be defined as

$$ \tilde{P}^\delta = \tilde{P}_1^\delta + \cdots + \tilde{P}_N^\delta. \tag{8} $$

Note that the solution $u^*$ of (5) is not, generally speaking, in the subspace $\tilde{V}^\delta$, therefore, the operator $\tilde{P}^\delta$ can not be used to solve the linear system (5) directly. We will need to modify the right-hand side of the system; see Lemma 2. We will also show that the elimination of the variables associated with the overlapping nodes is not needed in order to apply $\tilde{P}^\delta$ to a vector $v \in \tilde{V}^\delta$.

We now introduce the matrix form of (8). We define the restriction operator, or a matrix, $\tilde{R}_i^\delta$ as follows. Let $v = (v_1, \ldots, v_n)^T$ be a vector corresponding to the nodal values of a function $u \in \mathcal{V}$; namely for any node $x_i \in W$, $v_i = u(x_i)$. For convenience, we say "$v$ is defined on $W$". Its restriction on $\tilde{W}^\delta_i$, $\tilde{R}_i^\delta v$, is defined as

$$ (\tilde{R}_i^\delta v)(x_i) = \begin{cases} v_i & \text{if } x_i \in \tilde{W}^\delta_i \\ 0 & \text{otherwise} \end{cases} \tag{9} $$

The matrix representation of $\tilde{R}_i^\delta$ is given by a diagonal matrix with 1 for nodal points in $\tilde{W}^\delta_i$ and zero for the remaining nodal points. We remark that, by way of definition, the operator $\tilde{R}_i^\delta$ is symmetric; i.e., $(\tilde{R}_i^\delta)^T = \tilde{R}_i^\delta$. Use this restriction operator, we define the subdomain stiffness matrix as

$$ \tilde{A}_i^\delta = \tilde{R}_i^\delta A (\tilde{R}_i^\delta)^T, $$

which can also be obtained by the discretization of the original problem on $\tilde{W}^\delta_i$ with zero Dirichlet data on nodes $W \setminus \tilde{W}^\delta_i$. The matrix $\tilde{A}_i^\delta$ is block diagonal with blocks corresponding to the structure of $\tilde{R}_i^\delta$ and its inverse is understood as an inverse of the nonzero block. A matrix representation of $\tilde{P}_i^\delta$ denoted also by $\tilde{P}_i^\delta$ is equal to

$$ \tilde{P}_i^\delta = \left( \tilde{A}_i^\delta \right)^{-1} A $$
and
\[ \tilde{P}^\delta = \left( (\tilde{A}_1^\delta)^{-1} + \cdots + (\tilde{A}_N^\delta)^{-1} \right) A. \] (10)

The next lemma tell us how to modify the system (5) so that its solution belongs to \( \tilde{\mathcal{V}}^\delta \).

**Lemma 2** [CDS01] Let \( u^* \) and \( f \) be the exact solution and the right-hand side of (5), and
\[ w = \sum_{i=1}^{N} (\tilde{A}_i^\delta)^{-1} \tilde{R}_i^0 f, \] (11)
where \( \tilde{R}_i^0 \) is defined by (9) with \( \delta = 0 \). Then, we have \( \tilde{u}^* = u^* - w \in \tilde{\mathcal{V}}^\delta \) is the solution of the modified linear system of equations
\[ A\tilde{u}^* = f - Aw = \tilde{f}. \]

We remark that RASHO has several advantages over the classical AS. Let us recall AS briefly. Let
\[ \left( R_i^\delta v \right) (x_i) = \begin{cases} v_i & \text{if } x_i \in W_i^\delta \\ 0 & \text{otherwise.} \end{cases} \] (12)
Then the AS operator takes the following matrix form
\[ P^\delta = \left( (A_1^\delta)^{-1} R_1^\delta + \cdots + (A_N^\delta)^{-1} R_N^\delta \right) A, \] (13)
where \( A_i^\delta = R_i^\delta A (R_i^\delta)^T \). We remark that the size of the matrix \( A_i^\delta \) is \( |W_i^\delta| \), which is bigger than the size of the matrix \( A_i^\delta \), which is \( |\tilde{W}_i^\delta| \). In a distributed memory implementation, the operation \( R_i^\delta v \) involves moving data from one processor to another, but the operation \( \tilde{R}_i^\delta v \) does not involve any communication. In RASHO, if \( u \in \tilde{\mathcal{V}}^\delta \), then it is easy to see that
\[ \tilde{R}_i^\delta Au = \tilde{R}_i^\delta Au, \] (14)
where \( \tilde{R}_i^\delta, Au \) is defined as
\[ \left( \tilde{R}_i^\delta Au \right) (x_i) = \begin{cases} v_i & \text{if } x_i \in W_i^\delta \\ 0 & \text{otherwise.} \end{cases} \] (15)
Therefore, for functions in \( \tilde{\mathcal{V}}^\delta \), we can rewrite \( \tilde{P}^\delta \), as in (10), in the following form
\[ \tilde{P}^\delta = \left( (\tilde{A}_1^\delta)^{-1} \tilde{R}_1^\delta + \cdots + (\tilde{A}_N^\delta)^{-1} \tilde{R}_N^\delta \right) A. \] (16)
Although the operator (16) does not look like a symmetric operator, it is indeed symmetric when applying to functions in the subspace \( \tilde{\mathcal{V}}^\delta \). The form (14) takes the advantage of the fact that the operator \( \tilde{R}_i^\delta, Au \) is communication-free in the sense that it needs only the residual associated with nodes in \( W_i^\delta \subset \Omega_i^0 \).
We make some further comments on how the residual $Au$ can be calculated in a distributed memory environment, for a given vector $u \in \mathcal{V}_h^\delta$. In a typical implementation, the matrix $A$ is constructed and stored in the form of $\{A_i^\delta\}$, each processor has one or several of the subdomain matrix $\tilde{A}_i^\delta$. Similarly $u$ is stored in the form of $\{u_i\}$, where $u_i \in \mathcal{V}_i^\delta$. We note, however, that to compute the residual at nodes $W_{i,m}^\delta$, some communications are required. The processor associated with subdomain $\Omega_i^\delta$ needs to obtain the local solution from the neighboring subdomains at nodes connected to $W_{i,m}^\delta$. It is important to note that the amount of communications does not depend on the size of the overlap since only one layer of nodes is required. This shows that, in terms of the communication cost, RASHO is superior to AS and RAS.

**Main Results**

The algorithm presented in the previous section is applicable for general sparse, symmetric positive definite linear systems. The notions of subdomains, harmonic overlaps, the classification of nodal points, etc, can all be defined in terms of the graph of the sparse matrix. The following theorem provides a nearly optimal estimate of the condition number of the RASHO operator $\tilde{P}_h^\delta$ in terms of the fine mesh size $h$, the subdomain size $H$, and the overlapping factor $\delta$ for a Poisson equation discretized with a piecewise linear finite element method. We note that because we do not include a coarse space, the constant will depend on the subdomain size $H$.

**Theorem 1** [CDS01] The RASHO operator $\tilde{P}_h^\delta$ is symmetric in the inner product $a(\cdot, \cdot)$, non-singular, and bounded in the following sense

$$C_0^{-2}a(u, u) \leq a(\tilde{P}_h^\delta u, u) \leq C_1 a(u, u) \quad \forall u \in \mathcal{V}_h^\delta.$$  

(17)

Here

$$C_0^2 = C \left( (1 + \log(\delta + 1)) \left( 1 + \log\left( \frac{H}{h} \right) \right) + \frac{1}{H^2} \left( 1 + \log(\delta + 1) + \frac{H}{(2\delta + 1)h} \right) \right).$$

The constants $C, C_1 > 0$ are independent of $h, H$, and $\delta$.

We remark that the corresponding convergence rate estimate for the regular one-level AS [DW94], in terms of the constant $C_0$, is

$$C_0^2 = C \left( 1 + \frac{1}{H(2\delta + 1)h} \right).$$

The lower bound $C_0^2$ of RASHO is theoretically slightly worse than the lower bound of AS in the case of large overlap, but roughly the same for the case of small overlap. On the other hand, the upper bound $C_1$ of RASHO is better, since the overlap between subspaces $\mathcal{V}_k^\delta$ is generally smaller than the overlap between subspaces $\mathcal{V}_k^\delta$. Because of the smaller upper bound, the numerical performance of RASHO presented in the next section is better than that of AS. It is interesting to point out that, for the case of generous overlap, our estimate is equivalent to the estimate for the iterative substructuring algorithms [DSW94] without a coarse space. We also remark that the results of the paper is for only one-level Schwarz algorithms. Because of the “harmonic overlap” requirement, the extension of the algorithm to multiply levels is not as trivial as the multilevel AS.
Table 1: RASHO and AS preconditioned CG for solving the Poisson equation on a $128 \times 128$ mesh decomposed into $2 \times 2 = 4$ subdomains with overlap = $overlap$. The AS/CG results are shown in ( ). The “+1” is for the preprocessing step needed for RASHO.

<table>
<thead>
<tr>
<th>overlap (overlap)</th>
<th>iter</th>
<th>cond</th>
<th>max</th>
<th>min</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>42 (42)</td>
<td>129 (129)</td>
<td>1.98 (1.98)</td>
<td>0.0154 (0.0154)</td>
</tr>
<tr>
<td>1</td>
<td>24+1 (28)</td>
<td>48.4 (86.3)</td>
<td>1.94 (4.00)</td>
<td>0.0402 (0.0464)</td>
</tr>
<tr>
<td>2</td>
<td>20+1 (23)</td>
<td>33.3 (51.8)</td>
<td>1.91 (4.00)</td>
<td>0.0574 (0.0773)</td>
</tr>
<tr>
<td>3</td>
<td>18+1 (20)</td>
<td>27.2 (37.0)</td>
<td>1.89 (4.00)</td>
<td>0.0694 (0.1081)</td>
</tr>
</tbody>
</table>

Table 2: RASHO and AS preconditioned CG for solving the Poisson equation on a $32 \times DOM \times 32 \times DOM$ mesh decomposed into $DOM \times DOM$ subdomains with overlap = 1.

<table>
<thead>
<tr>
<th>$DOM \times DOM$</th>
<th>iter</th>
<th>cond</th>
<th>max</th>
<th>min</th>
</tr>
</thead>
<tbody>
<tr>
<td>$2 \times 2$</td>
<td>19+1 (20)</td>
<td>26.8 (43.7)</td>
<td>1.89 (4.00)</td>
<td>0.0708 (0.0916)</td>
</tr>
<tr>
<td>$4 \times 4$</td>
<td>39+1 (42)</td>
<td>86.9 (145)</td>
<td>1.95 (4.00)</td>
<td>0.0225 (0.0276)</td>
</tr>
<tr>
<td>$8 \times 8$</td>
<td>75+1 (78)</td>
<td>328 (550)</td>
<td>1.97 (4.00)</td>
<td>0.0060 (0.0073)</td>
</tr>
<tr>
<td>$16 \times 16$</td>
<td>147+1 (156)</td>
<td>1295 (2168)</td>
<td>1.98 (4.00)</td>
<td>0.0015 (0.0018)</td>
</tr>
</tbody>
</table>

Numerical Experiments

We present some numerical results for solving the Poisson equation on the unit square with zero Dirichlet boundary conditions. We compare the performance of RASHO/CG and AS/CG in terms of the number of iterations and the condition numbers. We pay particular attention to the dependence on the number of subdomains and the size of the overlap.

In order to use RASHO/CG, we need to modify the linear system by forcing its modified solution to belong to $V^\delta$. To do so, we use (11). The stopping condition for CG is to reduce the energy norm of the initial residual by a factor of $10^{-6}$. The exact solution of the equation is taken to be $u(x,y) = e^{i(x+y)} \sin(\pi x) \sin(\pi y)$. All subdomain problems are solved exactly.

The iteration count (iter), the condition number (cond), the maximum (max) and minimum (min) eigenvalues of the preconditioned matrix are summarized in Table 1, and Table 2. It is clear that the newly introduced RASHO/CG is always better than the classical AS/CG in terms of the iteration counts and the condition numbers. Although we do not have any parallel results to report at this point, we are confident that RASHO/CG would be even better than AS/CG on a parallel computers with distributed memory since much less communication is required.

References


