CUBO A Mathematical Journal
Vol. 6, № 4, (73 - 94). December 2004.

A Restricted Additive Schwarz Preconditioner with Harmonic Overlap for Symmetric Positive Definite Linear Systems

Xiao-Chuan Cai [1]
Department of Computer Science, University of Colorado
Boulder, CO 80309
cai@cs.colorado.edu

Maksymilian Dryja [2]
Faculty of Mathematics, Informatics and Mechanics
Warsaw University, Warsaw
dryja@mimuw.edu.pl

Marcus Sarkis [3]
Mathematical Sciences Department, Worcester Polytechnic Institute
Worcester, MA 01609
msarkis@wpi.edu

ABSTRACT

A restricted additive Schwarz (RAS) preconditioning technique was introduced recently for solving general nonsymmetric sparse linear systems. 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 (AS). The design of RASHO is based on a much deeper understanding of the behavior of Schwarz type methods in overlapping subregions, and in the construction of

[1]Supported in part by the NSF grants ASC-9457534, ECS-9725504, and ACI-0072089.
[2]Supported in part by the NSF grant CCR-9732208 and in part by the Polish Science Foundation grant 2 P03A 021 16.
[3]Supported in part by the NSF grant CCR-9984404.

the overlap. In 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 that 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 that do not need to be communicated. We also show numerically that RASHO preconditioned CG takes fewer number of iterations than the corresponding AS preconditioned CG. A nearly optimal theory is included for the convergence of RASHO/CG for solving elliptic problems discretized with a finite element method.

Key words and phrases: *Restricted additive Schwarz preconditioner, domain decomposition, harmonic overlap, elliptic equations, finite elements*

Math. Subj. Class.: *65N30, 65F10*

1 Introduction

A restricted additive Schwarz (RAS) preconditioning technique was introduced recently for solving general nonsymmetric sparse linear systems [1, 5, 7, 13, 15, 16, 17]. RAS outperforms the classical additive Schwarz preconditioner (AS) [8, 20] in the sense that it requires fewer number of iterations, as well as smaller communication and CPU time costs when implemented on distributed memory computers, [1]. Unfortunately, RAS in its original form is nonsymmetric and therefore the conjugate gradient method (CG) cannot be used [14]. Although a symmetrized version was constructed in [7], our numerical experiments show that it often takes more iterations than the corresponding AS/CG. In this paper we propose another modification of RAS and show in both theory and numerical experiments that this new variant works well for symmetric positive definite sparse linear systems and is superior to AS. Recall that the basic building blocks of classical Schwarz type algorithms are realized by solving the linear systems of the form

$$A_i^\delta w = R_i^\delta v \tag{1}$$

on each extended subdomain, where A_i^δ is the extended subdomain stiffness matrix and R_i^δ is the restriction operator for the extended subdomain (formal definitions will be given later in the paper). The key idea of RAS is that equation (1) is replaced by

$$A_i^\delta w = \begin{cases} v & \text{inside the un-extended subdomain} \\ 0 & \text{in the overlapping part of the subdomain.} \end{cases} \tag{2}$$

Note that the solution of (2) is discrete harmonic in the overlapping part of the subdomain, and therefore carries minimum energy in some sense. Setting part of the right-hand side vector to zero reduces the energy of the solution, and also destroys the symmetry of the additive Schwarz operator. In this paper, we further explore the idea of "harmonic overlap" and at the same time keep the symmetry of the Schwarz preconditioner. We mention that other approaches can also be taken to modify the Schwarz algorithm in the overlapping regions, such as allowing the functions to be discontinuous [4].

The algorithm to be discussed below is applicable for general symmetric positive definite problems. However, in order to provide a complete mathematical analysis, we restrict our discussion to a finite element problem, [3]. We consider a simple variational problem: Find $u \in H_0^1(\Omega)$, such that

$$a(u,v) = f(v), \quad \forall\ v \in H_0^1(\Omega), \tag{3}$$

where

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

For simplicity, let Ω be a bounded polygonal region in $\Re^2$ with a diameter of size $O(1)$. The extension of the results to $\Re^3$ can be carried out easily by using the theory developed here in this paper and the well-known three-dimensional additive Schwarz techniques; [9, 10, 12]. Let $\mathcal{T}^h(\Omega)$ be a shape regular, quasi-uniform triangulation, of size $O(h)$, of Ω and $\mathcal{V} \subset H_0^1(\Omega)$ 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

$$a(u^*,v) = f(v), \quad \forall\ v \in \mathcal{V}. \tag{4}$$

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

$$Au^* = f. \tag{5}$$

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

The paper is organized as follows. In section 2, we introduce notations. The new algorithm is described in section 3. Section 4 is devoted to the mathematical analysis of the new algorithm. We conclude the paper in section 5 by providing some numerical results and final remarks. Through out this paper, C and C_0, are positive generic constants that are independent of any of the mesh parameters and the number of subdomains. All the domains and subdomains are assumed to be open; i.e., boundaries are not included in their definitions.

2 Notations

Let n be the total number of interior nodes of $\mathcal{T}^h(\Omega)$ and W the set containing all the interior 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 subregion Ω_i^R as the union of all elements of $\mathcal{T}^h(\Omega)$ that have all three vertices in $W_i^0 \cup \partial\Omega$. Note that $\cup\bar{\Omega}_i^R$ is not equal to $\bar{\Omega}$; see Fig. 1(b). We denote by H as the representative size (diameter) of the subregion Ω_i^R.

We define the overlapping partition of W as follows. Let $\{W_i^1\}$ be the one-overlap partition of W, where $W_i^1 \supset W_i^0$ is obtained by including all the immediate neighboring vertices of all vertices in W_i^0; see Fig. 1(c). Using the idea recursively, we can define a δ-overlap partition of W,

$$W = \bigcup_{i=1}^{N} W_i^\delta.$$

Here the integer δ indicates the level of overlap with its neighboring subdomains and δh is approximately the extend of the extension. The definition of W_i^δ, as well as many other subsets, can be found in an illustrative picture, Fig. 1.

We next define a subregion of Ω induced by a subset 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 subregion Ω_i^δ is defined as $\Omega(W_i^\delta)$, and the corresponding subspace as

$$\mathcal{V}_i^\delta \equiv \mathcal{V} \cap H_0^1(\Omega_i^\delta) \text{ extended by zero to } \Omega\backslash\Omega_i^\delta.$$

It is easy to verify that

$$\mathcal{V} = \mathcal{V}_1^\delta + \mathcal{V}_2^\delta + \cdots + \mathcal{V}_N^\delta.$$

This decomposition is used in defining the classical one-level additive Schwarz algorithm [8]. Note that for $\delta = 0$ this decomposition is a direct sum. Let us define $P_i^\delta : \mathcal{V} \to \mathcal{V}_i^\delta$ by: for any $u \in \mathcal{V}$,

$$a(P_i^\delta u, v) = a(u, v), \quad \forall v \in \mathcal{V}_i^\delta. \tag{6}$$

Then, the classical one-level additive Schwarz operator has the form

$$P^\delta = P_1^\delta + \cdots + P_N^\delta.$$

In the classical AS as defined above, all the nodes of W_i^δ are treated equally even through some subsets of the nodes play different roles in determining the convergence rate of the AS preconditioned CG. To further understand the issue, we classify the nodes as follows. Let $\Gamma_i^\delta = \partial\Omega_i^\delta\backslash\partial\Omega$; i.e., the part of the boundary of Ω_i^δ that does not belong to the Dirichlet part of the physical boundary $\partial\Omega$. We define the interface overlapping boundary Γ^δ as the union of all Γ_i^δ; i.e., $\Gamma^\delta = \cup_{i=1}^N \Gamma_i^\delta$. We also need to define the following subsets of W, see, for examples, Fig. 1, where $\delta = 1$

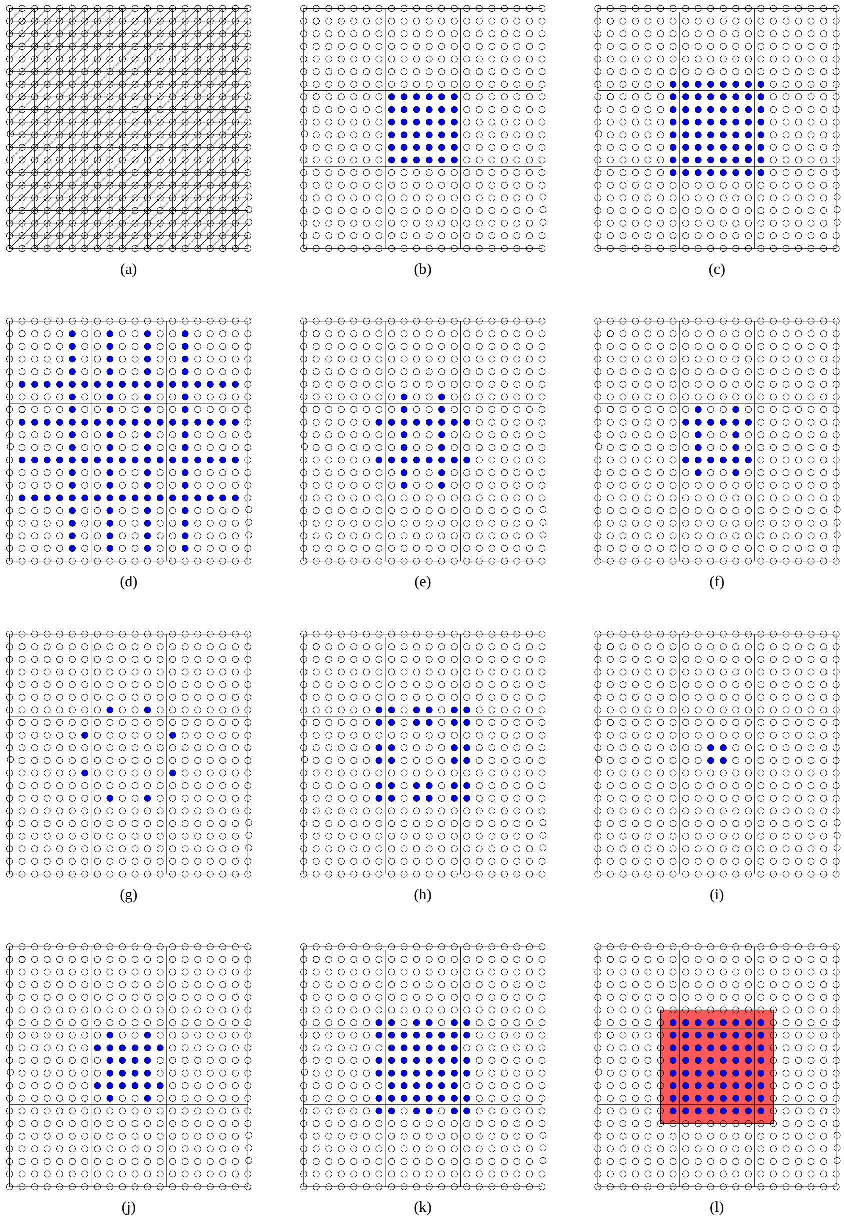

Figure 1: The partition of a finite element mesh into 9 subdomains with the overlapping factor $\delta = 1$. (a) the finite element mesh and nodal points; (b) a node-based partition of the mesh into 9 nonoverlapping subsets, and the collection of "•" forms the set W_i^0; (c) W_i^δ; (d) W^{Γ^δ}; (e) $W_i^{\Gamma^\delta}$; (f) $W_{i,in}^{\Gamma^\delta}$; (g) $W_{i,cut}^{\Gamma^\delta}$; (h) $W_{i,ovl}^\delta$; (i) $W_{i,non}^\delta$; (j) $W_{i,in}^\delta$; (k) $\widetilde{W}_i^\delta$; (l) the shadowed area is Ω_i^δ.

- $W^{\Gamma^\delta} \equiv W \bigcap \Gamma^\delta$ (interface nodes)
- $W_i^{\Gamma^\delta} \equiv W^{\Gamma^\delta} \bigcap W_i^\delta$ (local interface nodes)
- $W_{i,in}^{\Gamma^\delta} \equiv W^{\Gamma^\delta} \bigcap W_i^0$ (local internal interface nodes)
- $W_{i,cut}^{\Gamma^\delta} \equiv W_i^{\Gamma^\delta} \backslash W_{i,in}^{\Gamma^\delta}$ (local cut interface nodes)
- $W_{i,ovl}^{\delta} \equiv (W_i^\delta \backslash W_i^{\Gamma^\delta}) \bigcap (\bigcup_{j \neq i} W_j^\delta)$ (local overlapping nodes)
- $W_{i,non}^{\delta} \equiv W_i^\delta \backslash (W_i^{\Gamma^\delta} \bigcup W_{i,ovl}^\delta)$ (local nonoverlapping nodes)
- $W_{i,in}^{\delta} \equiv W_{i,non}^\delta \bigcup W_{i,in}^{\Gamma^\delta}$ (internal nodes)

We note that the most northwest and the southeast nodes in (c) were added to Γ_i^δ in order to make Ω_i^δ a rectangle. This just to simplify the presentation and it is not required in the implementation of the algorithms.

We frequently use functions that are discrete harmonic at certain nodes. Let $x_k \in W$ be a mesh point and $\phi_{x_k}(x) \in \mathcal{V}$ 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 $u \in \mathcal{V}$ 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 the subspace $\widetilde{\mathcal{V}}_i^\delta$ defined as a subspace of $\mathcal{V}_i^\delta$. $\widetilde{\mathcal{V}}_i^\delta$ consists of all functions that vanish on the cuting nodes $W_{i,cut}^{\Gamma^\delta}$ and discrete harmonic at the nodes of $W_{i,ovl}^\delta$. Note that the support of the subspace $\widetilde{\mathcal{V}}_i^\delta$ is

$$\widetilde{W}_i^\delta \equiv W_i^\delta \backslash W_{i,cut}^{\Gamma^\delta}$$

and, since the values at the harmonic nodes are not independent, they can not be counted toward the degree of freedoms. The dimension of $\widetilde{\mathcal{V}}_i^\delta$ is

$$dim\left(\widetilde{\mathcal{V}}_i^\delta\right) = |W_{i,in}^\delta|.$$

Let $\Omega(\widetilde{W}_i^\delta)$ be the induced domain. It is easy to see that $\Omega(\widetilde{W}_i^\delta)$ is the same as Ω_i^δ but with cuts. We denote $\Omega(\widetilde{W}_i^\delta)$ by $\widetilde{\Omega}_i^\delta$. We have then $\widetilde{\mathcal{V}}_i^\delta = \mathcal{V} \cap H_0^1(\widetilde{\Omega}_i^\delta)$ and discrete harmonic on $\Omega(W_{i,ovl}^\delta)$. We denote $\Omega(W_{i,ovl}^\delta)$ by $\Omega_{i,ovl}^\delta$.

We define $\widetilde{\mathcal{V}}^\delta \subset \mathcal{V}^\delta$ as

$$\widetilde{\mathcal{V}}^\delta = \widetilde{\mathcal{V}}_1^\delta \oplus \cdots \oplus \widetilde{\mathcal{V}}_N^\delta,$$

which is a direct sum. We remark that functions in $\widetilde{\mathcal{V}}^\delta$ are, by definition, the sum of functions $u_i \in \widetilde{\mathcal{V}}_i^\delta$, $i = 1, \cdots, N$. Functions in $\widetilde{\mathcal{V}}^\delta$ can, in fact, be characterized easily as in the following lemma.

Lemma 2.1 *If $u \in \mathcal{V}$ and u is discrete harmonic at all the overlapping nodes, i.e., on $\cup_{i=1}^N W_{i,ovl}^\delta$, then $u \in \widetilde{\mathcal{V}}^\delta$.*

Proof. To prove that $u \in \widetilde{\mathcal{V}}^\delta$, all we need is to find a decomposition

$$u = \sum_{i=1}^N u_i, \quad \text{with } u_i \in \widetilde{\mathcal{V}}_i^\delta, \ \ i = 1, \cdots, N.$$

For the given u, we define u_i piece by piece as follows. On the nodes in $W_{i,in}^\delta$ we let $u_i = u$. On the nodes in $W_{i,cut}^\delta$ we let u_i be zero. On the nodes outside W_i^δ we set u_i to zero. We now need only to define u_i on the nodes belong to $W_{i,ovl}^\delta$. There, we extend u_i as a discrete harmonic function with boundary data given by u_i just defined. ∎

3 Restricted additive Schwarz with harmonic overlap (RASHO)

Using notations introduced in the previous section, we now describe a new method, namely a restricted additive Schwarz with harmonic overlap.

We first define $\widetilde{P}_i^\delta : \widetilde{\mathcal{V}}^\delta \to \widetilde{\mathcal{V}}_i^\delta$ as a projection operator, such that, for any $u \in \widetilde{\mathcal{V}}^\delta$

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

The RASHO operator can then be defined as

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

Note, however, that the solution u^* of (4), see also (5), is not, generally speaking, in the subspace $\widetilde{\mathcal{V}}^\delta$, therefore, the operator $\widetilde{P}^\delta$ cannot be used to solve the linear system (5) directly. We will need to modify the right-hand side of the system (5). A reformulated (5) will be presented in Lemma 3.1 below. We will show that the elimination of the variables associated with the overlapping nodes is not needed in order to apply $\widetilde{P}^\delta$ to any given vector $v \in \widetilde{P}^\delta$.

We now introduce a matrix form of (8). We define the restriction operator, or a matrix, $\widetilde{R}_i^\delta$ 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 $\widetilde{W}_i^\delta$, $\widetilde{R}_i^\delta v$, is defined as

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

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

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

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

$$\widetilde{P}_i^\delta = \left(\widetilde{A}_i^\delta\right)^{-1} \; A$$

and

$$\widetilde{P}^\delta = \left((\widetilde{A}_1^\delta)^{-1} + \cdots + (\widetilde{A}_N^\delta)^{-1}\right) \; A. \tag{10}$$

Using the matrix notations, the next lemma shows how to modify the system (5) so that its solution belongs to $\widetilde{\mathcal{V}}^\delta$.

Lemma 3.1 *Let u^* and f be the exact solution and the right-hand side of (5), and*

$$w = \sum_{i=1}^{N} (\widetilde{A}_i^\delta)^{-1} \widetilde{R}_i^0 f, \tag{11}$$

then, we have $\widetilde{u}^* = u^* - w \in \widetilde{\mathcal{V}}^\delta$, which is the solution of the modified linear system of equations

$$A\widetilde{u}^* = f - Aw = \widetilde{f}.$$

Proof. If we can show that

$$a(w, \phi_k) = f(\phi_k),$$

for a regular basis function associated with an arbitrary overlapping node $x_k \in W_{i,ovl}^\delta$, for some i, then we will have

$$a(u^* - w, \phi_k) = f(\phi_k) - f(\phi_k) = 0, \tag{12}$$

which says that $\widetilde{u}^* = u^* - w$ is discrete harmonic at the overlapping node x_k. We can then use Lemma 2.1 to conclude the proof. Let us now consider

$$w_i = (\widetilde{A}_i^\delta)^{-1} \widetilde{R}_i^0 f,$$

which, by definition, is the same as

$$a(w_i, \phi_j) = (\phi_j, \widetilde{R}_i^0 f), \;\; \forall x_j \in \widetilde{W}_i^\delta.$$

Here and in the rest of the proof, ϕ_j is the basis function associated with the node $x_j \in \widetilde{W}_i^\delta$. Using that $\widetilde{R}_i^0$ is symmetric and

$$(\phi_j, \widetilde{R}_i^0 f) = (f, \widetilde{R}_i^0 \phi_j) = a(u^*, \widetilde{R}_i^0 \phi_j),$$

we get

$$a(w_i, \phi_j) = a(u^*, \widetilde{R}_i^0 \phi_j). \tag{13}$$

Let us compute $a(w_i, \phi_k)$. Since x_k is an overlapping node, it cannot be on the boundary of $\widetilde{\Omega}_i^\delta$. This leaves us with the following two cases.

Case (1): The support of $\phi_k(x)$ belongs to the exterior of $\widetilde{\Omega}_i^\delta$. Since the supports of w_i and ϕ_k do not overlap, we have

$$a(w_i, \phi_k) = 0.$$

Case (2): The support of $\phi_k(x)$ belongs to the interior of $\widetilde{\Omega}_i^\delta$. In this case, we have

$$a(w_i, \phi_k) = a(u^*, \widetilde{R}_i^0 \phi_k).$$

Taking the sum of the above equality for $i = 1, \cdots, N$,

$$a(w, \phi_k) = a\left(\sum_{i=1}^N w_i, \phi_k\right) = a\left(u^*, \sum_{i=1}^N \widetilde{R}_i^0 \phi_k\right) = a(u^*, \phi_k),$$

which proves (12). Here the fact $\sum_{i=1}^N \widetilde{R}_i^0 = I$ is used. ∎

There are basically two ways to compute w in practice. Suppose that subdomain problems are solved using some LU factorization based method. One can use the same factorization of $\widetilde{A}_i^\delta$ to modify the right-hand side of the system and to solve subdomain problems in the preconditioning steps, as what was suggested in Lemma 3.1. Or, one can obtain w by solving several small Poisson problems on each subdomain with zero Dirichlet boundary conditions in the overlapping regions $\Omega_{i,ovl}^\delta$. In both strategies, the computation can be done in parallel and no communication is needed in a distributed memory implementation.

Let $\widetilde{f} = f - Aw$, then $\widetilde{u}^*$ is the solution of the following linear system of equations

$$A\widetilde{u}^* = \widetilde{f}. \tag{14}$$

Since $\widetilde{u}^* \in \widetilde{\mathcal{V}}^\delta$,

$$g \equiv \widetilde{P}^\delta \widetilde{u}^*$$

is well defined, and can be computed without knowing $\widetilde{u}^*$ by using the following relations:

$$a(\widetilde{P}_i^\delta \widetilde{u}^*, v) = a(\widetilde{u}^*, v) = (\widetilde{f}, v), \ \ \forall v \in \widetilde{\mathcal{V}}_i^\delta \text{ and } i = 1, \cdots, N.$$

More precisely speaking, we can obtain g by solving the subdomain problems

$$a(g_i, v) = (\widetilde{f}, v), \quad \forall v \in \widetilde{\mathcal{V}}_i^\delta,$$

for $i = 1, \cdots, N$, and taking $g = g_1 + \cdots + g_N$. With such a right-hand side, we introduce a new linear system

$$\widetilde{P}^\delta \widetilde{u}^* = g, \tag{15}$$

which is equivalent to the linear system (14). The system (15) is a symmetric positive definite system under the usual energy inner product, and therefore, can be solved using the conjugate gradient method. RASHO has a few advantages over the classical AS preconditioner. Let us recall AS briefly. Let

$$\left(R_i^\delta v\right)(x_i) = \begin{cases} v_i & \text{if } x_i \in W_i^\delta \\ \\ 0 & \text{otherwise.} \end{cases} \tag{16}$$

Then the AS operator takes the following matrix form

$$P^\delta = \left((A_1^\delta)^{-1} + \cdots + (A_N^\delta)^{-1}\right) \ A \tag{17}$$

where $A_i^\delta = R_i^\delta A (R_i^\delta)^T$. Because of the inclusion of the cut interface nodes, the size of the matrix A_i^δ is $|W_i^\delta|$, which is slightly larger than the size of the matrix $\widetilde{A}_i^\delta$, which is $|\widetilde{W}_i^\delta|$. In a distributed memory implementation, the operation $R_i^\delta v$ involves moving data from one processor to another, but the operation $\widetilde{R}_i^\delta v$ does not involve any communication. More precisely speaking, in RASHO, if $u \in \widetilde{\mathcal{V}}^\delta$, then it is easy to see that

$$\widetilde{R}_i^\delta A u = \widetilde{R}_{i,in}^\delta A u, \tag{18}$$

where $\widetilde{R}_{i,in}^\delta$ is defined as

$$\left(\widetilde{R}_{i,in}^\delta v\right)(x_i) = \begin{cases} v_i & \text{if } x_i \in W_{i,in}^\delta \\ \\ 0 & \text{otherwise.} \end{cases} \tag{19}$$

Therefore, for functions in $\widetilde{\mathcal{V}}^\delta$, we can rewrite $\widetilde{P}^\delta$, as in (10), in the following form

$$\widetilde{P}^\delta = \left((\widetilde{A}_1^\delta)^{-1}\widetilde{R}_{1,in}^\delta + \cdots + (\widetilde{A}_N^\delta)^{-1}\widetilde{R}_{N,in}^\delta\right) \ A. \tag{20}$$

Although the operator (20) does not look like a symmetric operator, but it is indeed symmetric when applying to functions in the subspace $\widetilde{\mathcal{V}}^\delta$. The form (18) takes the advantage of the fact that the operator $\widetilde{R}_{i,in}^\delta$ is communication-free in the sense that it needs only the residual associated with nodes in $W_{i,in}^{\Gamma^\delta} \subset \Omega_i^0$.

We make some further comments on how the residual Au can be calculated in a distributed memory environment, for a given vector $u \in \widetilde{\mathcal{V}}^\delta$. In a typical implementation, the matrix A is constructed and stored in the form of $\{\widetilde{A}_i^\delta\}$, each processor has one or several of the subdomain matrix $\widetilde{A}_i^\delta$. Similarly u is stored in the form of $\{u_i\}$, where $u_i \in \widetilde{\mathcal{V}}_i^\delta$. We note, however, that to compute the residual at nodes $W_{i,in}^{\Gamma^\delta}$ some communications are required. The processor associated with subdomain Ω_i^δ needs to obtain the local solution from the neighboring subdomains at nodes connected to $W_{i,in}^{\Gamma^\delta}$. 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 communications, the RASHO is superior to AS and RAS.

4 Theoretical analysis

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 the nodal points, etc, can all be defined in terms of the graph of the sparse matrix. In this section we provide a nearly optimal estimate for a Poisson equation discretized with a piecewise linear finite element method. We estimate the condition number of the RASHO operator $\widetilde{P}^\delta$ in terms of the fine mesh size h, the subdomain size H, and the overlapping factor δ. We note that because we do not include a coarse space, the constant will depend on the subdomain size H. We shall follow the abstract additive Schwarz theory [20]:

Lemma 4.1 *Suppose the following assumptions hold:*

i) There exists a constant C_0 such that for any $u \in \widetilde{\mathcal{V}}^\delta$ there exists a decomposition

$$u = \sum_{i=1}^{N} u_i,$$

where $u_i \in \widetilde{\mathcal{V}}_i^\delta$, and

$$\sum_{i=1}^{N} |u_i|^2_{H^1(\Omega_i^\delta)} \le C_0^2 |u|^2_{H^1(\Omega)}.$$

ii) There exist constants $\epsilon_{ij}, i, j = 1, \ldots, N$ such that

$$a(u_i, u_j) \le \epsilon_{ij}\, a(u_i, u_i)^{1/2} a(u_j, u_j)^{1/2}, \quad \forall u_i \in \widetilde{\mathcal{V}}_i^\delta, \quad \forall u_j \in \widetilde{\mathcal{V}}_j^\delta.$$

Then, $\widetilde{P}^\delta$ is invertible, symmetric; i.e., $a(\widetilde{P}^\delta u, v) = a(u, \widetilde{P}^\delta v)$, $\forall u, v \in \widetilde{\mathcal{V}}^\delta$, and

$$C_0^{-2} a(u,u) \le a(\widetilde{P}^\delta u, u) \le \rho(\mathcal{E}) a(u,u), \quad \forall u \in \widetilde{\mathcal{V}}^\delta. \tag{21}$$

Here $\rho(\mathcal{E})$ is the spectral radius of $\mathcal{E}$, which is a $(N) \times (N)$ matrix made of $\{\epsilon_{ij}\}$.

It is trivial to see that $\rho(\mathcal{E}) \le C$. So our focus in the rest of the paper is in bounding C_0.

4.1 A partition of unity and a comparison function

The construction of a partition of unity is one of the key steps in an additive Schwarz analysis. We construct $\phi^i(x)$ as follows:

$$\phi^i(x_k) = \begin{cases} 1 & \text{if } x_k \in W_{i,in}^{\Gamma^\delta} \\ \text{discrete harmonic} & \text{if } x_k \in W_{i,ovl}^{\delta} \cup W_{i,non}^{\delta} \\ 0 & \text{if } x_k \in W \backslash \widetilde{W}_i^\delta \end{cases}$$

Note that $\phi^i(x) = 0$ if $x \notin \widetilde{\Omega}_i^\delta$. Let us denote $\Omega(W_{i,non}^\delta)$ by $\Omega_{i,non}^\delta$, then $\phi^i(x_k) = 1$ at $x_k \in W_{i,non}^\delta$ for the case $\Omega_{i,non}^\delta \cap \partial\Omega = \emptyset$ since all the boundary nodes of $\Omega_{i,non}^\delta$ belong to $W_{i,in}^{\Gamma^\delta}$. Also, it is easy to see that $\{\phi^i(x), i = 1, \ldots, N\}$ restricted to W^{Γ^δ} form a partition of unit.

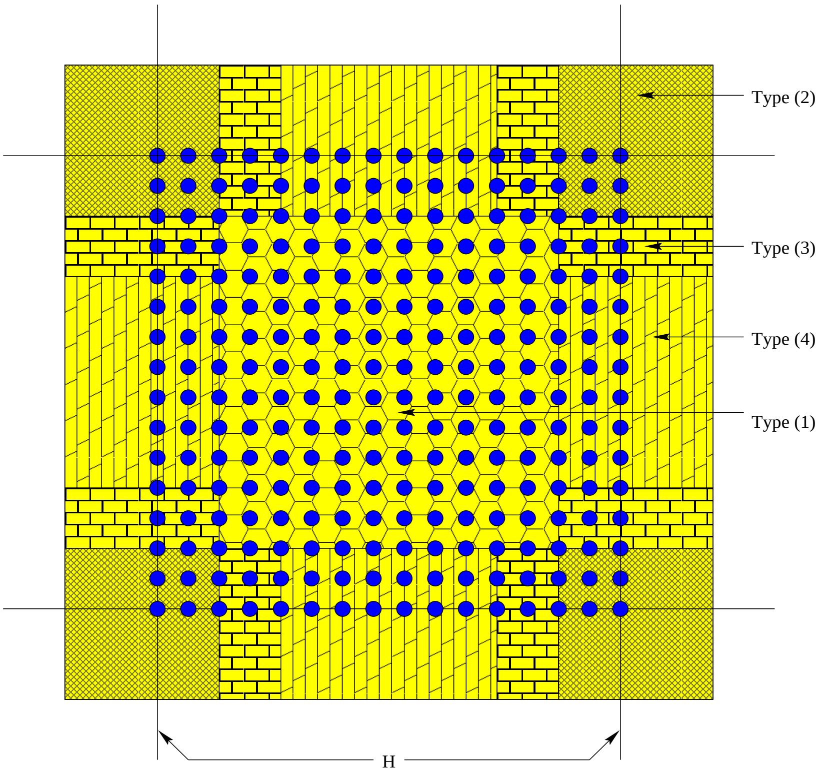

Figure 2: The partition of Ω_i^δ into the union of four types of subregions. This is a 'floating' subdomain with $\delta = 2$. The collection of "•" forms the set W_i^0.

In addition to $\phi^i(x)$, we also need to construct a *comparison function* $\theta_i(x)$ for each subdomain Ω_i^δ. Comparison functions, or barrier functions, are very useful for many Schwarz algorithms, such as these on non-matching grids [6]. We will show that, even though $\theta_i(x) \in \mathcal{V}_i^\delta$, not in $\widetilde{\mathcal{V}}_i^\delta$ as we wished, it can still be used to bound functions in $\widetilde{\mathcal{V}}_i^\delta$. Both $\theta_i(x)$ and $\phi^i(x)$ depend on the overlapping factor δ. Because $\phi^i(x)$ is discrete harmonic at $W_{i,ovl}^\delta \cup W_{i,non}^\delta$, we have

$$a(\phi^i, \phi^i) \leq a(\theta_i, \theta_i).$$

To construct the function $\theta_i(x)$, we first consider the case when Ω_i^0 is a floating square subdomain. "Floating" refers to the fact that the subdomain doesn't touch the boundary $\partial\Omega$. The extension to cases when Ω_i^δ touches the boundary is simple and we will comment on it later. To further simplify our arguments, we assume that Ω_i^δ and its neighboring extended subdomains Ω_j^δ are squares of the same size; i.e. sides length equals to $H + 2(\delta + 1)h$. This assumption is equivalent to that Ω^R has size H and δ levels of overlap is applied; see Fig. 2. And we also assume the overlap is not too large; for the analysis given below δh no larger than $H/4$ is enough. Our techniques can be modified to consider larger overlaps and more complex subdomains, although too large of an overlap has little practical value.

Roughly speaking, $\theta_i(x)$ equals to $\phi^i(x)$ on $W \backslash W_{i,ovl}^\delta$. On the overlapping region $W_{i,ovl}^\delta$ we need to define $\theta_i(x)$ carefully so that we can control its energy in the semi H^1 norm. For this purpose, we decompose Ω_i^δ into subregions of four types: $\Omega_{i,non}^\delta$, $\Omega_i^{\delta\delta}$, $\Omega_i^{\delta H}$, and $\Omega_i^{\delta\tilde{\delta}}$ and define $\theta_i(x)$ on each piece of the subregion separately.

Type (1): The first subregion is $\Omega_{i,non}^\delta$, which a square with sides of size $H - 2\delta h$.

Type (2): The second subregion $\Omega_i^{\delta\delta}$ is the area where Ω_i^δ overlaps simulatneously with three neighbors Ω_j^δ. $\Omega_i^{\delta\delta}$ therefore represents the union of the four corner pieces of Ω_i^δ; i.e. four squares with sides of size $(2\delta + 1)h$.

Type (3) and (4): The area where Ω_i^δ overlaps only one neighbor are four rectangles of size $H - 2\delta h$ by $(2\delta + 1)h$. We further partition each of the four rectangles into three smaller rectangles; i.e. two of them are of $\Omega_i^{\delta\tilde{\delta}}$ type and one of them of $\Omega_i^{\delta H}$ type. For instance, without lost of generality, Let us consider the intersection of Ω_i^δ with its right neighbor Ω_j^δ, excluding the corner parts. In this case, the subregion to be partitioned is a rectangle of size $(2\delta + 1)h$ in the x direction and $H - 2\delta h$ in the y direction. The partition of this rectangles gives two smaller rectangles of $\Omega_i^{\delta\tilde{\delta}}$ type with dimensions $2(\delta + 1)h \times \delta h$ and each one has an edge in common with a square of $\Omega_i^{\delta\delta}$ type. We denote them as transition subregions because they are placed between a corner type subregion $\Omega_i^{\delta\delta}$ and a face type subregion $\Omega_i^{\delta H}$. The $\Omega_i^{\delta H}$ face type subregions are the smaller rectangles that are placed between the two smaller rectangles of $\Omega_i^{\delta\tilde{\delta}}$ type. $\Omega_i^{\delta H}$ face type regions are of size $(2\delta + 1)h$ by $H - 4\delta h$.

For any node x belonging to a **Type** (1) region $\Omega_{i,non}^\delta$, we define $\theta_i(x)$ to be equal to one; i.e., equals to $\phi^i(x)$. Therefore

$$|\phi^i(x)|^2_{H^1(\Omega_{i,non}^\delta)} = |\theta_i(x)|^2_{H^1(\Omega_{i,non}^\delta)} = 0.$$

We next define $\theta_i(x)$, node by node, in $\Omega_{i,ovl}^\delta$, which is the union of corner, transition and face type regions defined above.

For a **Type** (2) region $\Omega_i^{\delta\delta}$. Let Q be such a square with vertices $V_1 = (a, b)$, $V_2 = (a + (2\delta + 1)h, b)$, $V_3 = (a, b + (2\delta + 1)h)$, and $V_4 = (a + (2\delta + 1)h, b + (2\delta + 1)h)$. We assume that V_1, V_2, and V_4 belong to $\partial\Omega_i^{\delta}$. In other words, Q is located on the southeast corner of Ω_i^{δ}. Let use also introduce another square region $\widetilde{Q}$, with vertices $V_3 = (a, b + (2\delta + 1)h)$, $\widetilde{V}_1 = (a, b + \delta h)$, $\widetilde{V}_2 = (a + (\delta + 1)h, b + \delta h)$, and $\widetilde{V}_4 = (a + (\delta + 1)h, b + (2\delta + 1)h)$. Note that $\widetilde{Q}$ is contained in Q, with V_3 as the common vertex. To define $\theta_i(x)$ on Q, we set $\theta_i(V_3) = 1$, $\theta_i(\widetilde{V}_1) = 0$, $\theta_i(\widetilde{V}_2) = 0$, $\theta_i(\widetilde{V}_4) = 0$. At the remaining nodes x on the edges $\widetilde{V}_1\widetilde{V}_2$ and $\widetilde{V}_2\widetilde{V}_4$ we set $\theta_i(x) = 0$, and on the edges $V_3\widetilde{V}_1$ and $V_3\widetilde{V}_4$ we set $\theta_i(x) = 1$. For nodes on $Q\backslash\widetilde{Q}$ we set $\theta_i(x) = 0$. It remains only to define $\theta_i(x)$ for nodes x in the interior of $\widetilde{Q}$. To define $\theta_i(x)$ there we use a well-known cutoff function technique, such as the one introduced in Lemma 4.4 of [10] but for two-dimensional square regions. An illustrative picture of $\theta_i(x)$ in a typical region $\Omega_i^{\delta\delta}$ is shown in Fig 3.

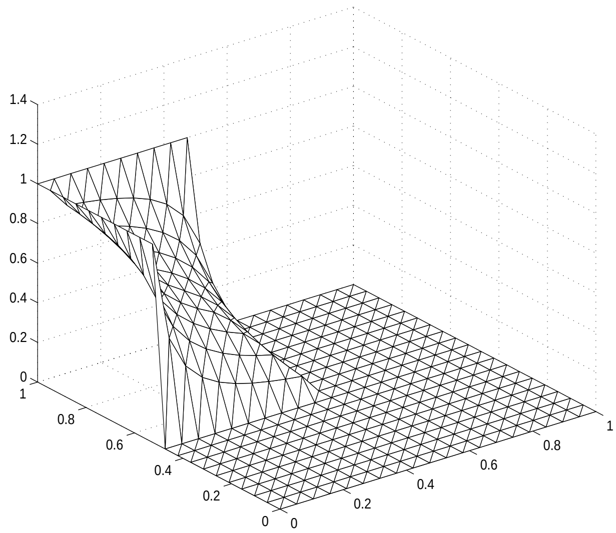

Figure 3: An illustrative picture of $\theta_i(x)$ in a typical region $\Omega_i^{\delta\delta}$.

For the completeness of this paper, we include the construction below. Let C be the center of the square $\widetilde{Q}$. The construction of $\theta_i(x)$ is defined by the following steps:

(1) Define $\theta_i(V_3) = 1$, $\theta_i(\widetilde{V}_2) = 0$, $\theta_i(\widetilde{V}_1) = 0$ and $\theta_i(\widetilde{V}_4) = 0$.

(2) For a point P that belongs to the segments $V_3\widetilde{V}_4$ or $V_3\widetilde{V}_1$, define $\theta_i(P) = 1$. For a point P that belongs to the segments $\widetilde{V}_4\widetilde{V}_2$ or $\widetilde{V}_1\widetilde{V}_2$, define $\theta_i(P) = 0$.

(3) For a point Y that belongs to the line segment connecting C to V_3, define $\theta_i(Y)$ by linear interpolation between values $\theta_i(C) = 1/2$ and $\theta_i(V_3) = 1$. For a point

Y that belongs to the line segment connecting C to $\widetilde{V}_2$, define $\theta_i(Y)$ by linear interpolation between values $\theta_i(C) = 1/2$ and $\theta_i(\widetilde{V}_2) = 0$.

(4) For a point S that belongs to a line segment connecting a point Y to a vertex $\widetilde{V}_1$ or $\widetilde{V}_4$, define $\theta_i(S) = \theta_i(Y)$.

(5) Note that the θ_i is defined everywhere on $\widetilde{Q} \cup \partial\widetilde{Q}$. θ_i is continuous everywhere except at the points $\widetilde{V}_1$ and $\widetilde{V}_4$. We redefine θ_i as the continuous piecewise linear finite element function given by the standard pointwise interpolation.

The most important observation of the construction of $\theta_i(x)$ inside $\widetilde{Q}$ is that $|\nabla\theta_i(x)| \le C/r$ near $\widetilde{V}_1$ or $\widetilde{V}_4$. Here r is the distance of x to $\widetilde{V}_1$ or $\widetilde{V}_4$. Therefore, we obtain (see [10] and [19])

$$|\theta_i(x)|^2_{H^1(Q)} = |\theta_i(x)|^2_{H^1(\widetilde{Q})} \le C\left(1 + \log\left(\frac{(\delta+1)h}{h}\right)\right) = C(1 + \log(\delta+1)).$$

Since inside of Ω_i^δ there are four of those squares we obtain

$$|\theta_i(x)|^2_{H^1(\Omega_i^{\delta\delta})} \le C\left(1 + \log(\delta+1)\right).$$

Type (3) regions consist of transition type rectangles. Let us consider one of them and denote it by T, which we assume has vertices at $V_3 = (a, b + (2\delta+1)h)$, $V_4 = (a+(2\delta+1)h, b+(2\delta+1)h)$, $V_5 = (a, b+(3\delta+1)h)$, and $V_6 = (a+(2\delta+1)h, b+(3\delta+1)h)$. Note that T stands on the top of the square Q introduced above and has the common edge V_3V_4. We define $\theta_i(x)$ over the edge V_3V_4 to be equal to $\phi^i(x)$. Over the edge V_3V_5, we set $\theta_i(x) = 1$. Over the edge V_4V_6, we set $\theta_i(x) = 0$. And over the edge V_5V_6 we let $\theta_i(x)$ decrease linearly from the value 1 to 0. What remains is to define $\theta_i(x)$ inside T. Let us define the nodes $V_l = (a + \delta h, b + (2\delta+1)h)$ and $V_r = (a + (\delta+1)h, b + (2\delta+1)h)$, which is the same as the node $\widetilde{V}_4$ used in the description of **Type** (2) regions. The nodes V_l and V_r are exactly the places on the edge V_3V_4 where $\phi^i(x)$ jumps from 1 to 0. On the triangle $V_3V_lV_5$ we set $\theta_i(x) = 1$. On the triangle $V_rV_4V_6$ we set $\theta_i(x) = 0$. On the region $V_lV_rV_6V_5$, we let $\theta_i(x)$ decrease linearly in the x direction from the value 1 to 0. We note that next to the nodes V_lV_r, $\theta_i(x)$ has a singular behavior similar to $|\nabla\theta_i(x)| \le C/r$ where r is the distance from x to the line $V_l\ V_r$. Similarly, we have

$$|\theta_i(x)|^2_{H^1(T)} \le C\left(1 + \log(\delta+1)\right)).$$

Since there are eight rectangles of **Type** (3) inside $\Omega_i^{\delta\tilde{\delta}}$, we obtain

$$|\theta_i(x)|^2_{H^1(\Omega_i^{\delta\delta})} \le C\left(1 + \log(\delta+1)\right).$$

Type (4) regions are rectangles of face type. Let R be one of them, and we assume that the vertices are given by $V_5 = (a, b+(3\delta+1)h)$, $V_6 = (a+(2\delta+1)h, b+(3\delta+1)h)$, $V_7 = (a, b + H - (\delta-1)h)$, and $V_8 = (a + (2\delta+1)h, b + H - (\delta-1)h)$. Note that R is on the top of the rectangle T defined above and its height is $H - 4\delta h$. The vertices V_6 and V_8 are the vertices that belong to $\partial\Omega_i^\delta$. We define $\theta_i(x) = 1$ if x is on the edge V_5V_7, and equals zero if x is on the edge V_6V_8, and linear in the x direction for the remaining points. We obtain then

$$|\theta_i(x)|^2_{H^1(R)} \leq \frac{H - 4\delta h}{(2\delta + 1)h}.$$

Since there are four of those rectangles inside $\Omega_i^{\delta H}$, we obtain

$$|\theta_i(x)|_{H^1(\Omega_i^{\delta H})} \leq C\frac{H - 4\delta h}{(2\delta + 1)h} \leq C\frac{H}{(2\delta + 1)h}.$$

For the cases in which Ω_i^0 touches the boundary $\partial\Omega$, the analysis needs to be modified slightly. The first modification is because the shape of the overlapping region changes slightly, i.e. the longer side is shorter. It is easy to see that we get similar bounds as before. The other modification is because ϕ^i on $\Omega_{i,non}^\delta$ is not identically equal to one and therefore the corresponding energy is not necessarily zero. For this case we can design θ_i similarly and obtain

$$|\theta_i(x)|^2_{H^1(\Omega^\delta_{i,non})} \leq C\left(1 + \log\left(\frac{H}{h}\right)\right).$$

Putting all pieces of $\theta_i(x)$ together, we see that $\theta_i(x) \in \mathcal{V}_i^\delta$ and it equals to $\phi^i(x)$ on W^{Γ^δ}. Adding all the estimates on subregions of four types, we arrive at the following lemma.

Lemma 4.2 *For $i = 1, \cdots, N$. $\theta_i(x) \in \mathcal{V}_i^\delta$, and $\phi^i(x) \in \tilde{\mathcal{V}}_i^\delta$, and also*

(1) $|\phi^i|^2_{H^1(\Omega_i^\delta)} \leq |\theta_i|^2_{H^1(\Omega_i^\delta)}$.

(2) $|\theta_i|^2_{H^1(\Omega_i^\delta \backslash \Omega^\delta_{i,non})} \leq C\left(1 + \log(\delta + 1) + \frac{H}{(2\delta + 1)h}\right)$.

(3) if $\Omega^\delta_{i,non} \cap \partial\Omega = \emptyset$ *then* $|\theta_i|^2_{H^1(\Omega^\delta_{i,non})} = 0$.

(4) if $\Omega^\delta_{i,non} \cap \partial\Omega \neq \emptyset$ *then* $|\theta_i|^2_{H^1(\Omega^\delta_{i,non})} \leq C\left(1 + \log\left(\frac{H}{h}\right)\right)$.

Here $C > 0$ is independent of the parameters h, H and δ.

4.2 A bounded partition lemma

To obtain the parameter C_0 of Assumption i) of the abstract additive Schwarz theory, see Lemma 4.1, we construct a decomposition of $\widetilde{\mathcal{V}}^\delta$ and prove its boundedness below.

Lemma 4.3 *There exists a constant $C > 0$, independent of h, H, and δ, such that for any $u \in \widetilde{\mathcal{V}}^\delta$, there exist $u_i \in \widetilde{\mathcal{V}}_i^\delta$, such that*

$$u = \sum_{i=1}^{N} u_i, \tag{22}$$

and

$$\begin{aligned}\sum_{i=1}^{N} |u_i|^2_{H^1(\Omega)} \quad &\leq \quad C\left(1+\log(\delta+1)\right)\left(1+\log\left(\tfrac{H}{h}\right)\right)|u|^2_{H^1(\Omega)} + \\ & C\frac{1}{H^2}\left(1+\log(\delta+1)+\frac{H}{(2\delta+1)h}\right)|u|^2_{H^1(\Omega)}.\end{aligned} \tag{23}$$

Proof. We first construct the decomposition (22). For any given $u \in \widetilde{\mathcal{V}}^\delta$, we define $u_i \in \widetilde{\mathcal{V}}_i^\delta$ as

$$u_i(x_k) = \begin{cases} u(x_k) & \text{if } x_k \in W^\delta_{i,in} \\ \text{discrete harmonic} & \text{if } x_k \in W^\delta_{i,ovl} \\ 0 & \text{if } x_k \in W \backslash \widetilde{W}_i^\delta. \end{cases}$$

It is easy to see (22) holds. For $i = 1, \ldots, N$, let us define $v_i \in \widetilde{\mathcal{V}}_i^\delta$ by

$$v_i = u_i - \bar{u}_i \phi^i \in \widetilde{\mathcal{V}}_i^\delta,$$

where

$$\bar{u}_i = \frac{1}{|\Omega_i^\delta|}\int_{\Omega_i^\delta} u dx$$

is the average of u on the extended region Ω_i^δ. Here $|\Omega_i^\delta|$ is the area of the region Ω_i^δ. The next step is to bound the sums $\sum_{i=1}^N |v_i|^2_{H^1(\Omega)}$ and $\sum_{i=1}^N |\bar{u}_i\phi^i|^2_{H^1(\Omega)}$. For the second sum, we use Lemma 4.2 to obtain

$$\begin{aligned}\sum_{i=1}^{N} |\bar{u}_i\phi^i|^2_{H^1(\Omega)} \leq C\left(1+\log\left(\frac{H}{h}\right)\right)\sum_{i\in\partial\Omega}|\bar{u}_i|^2 + \\ C\left(1+\log(\delta+1)+\frac{H}{(2\delta+1)h}\right)\sum_i |\bar{u}_i|^2.\end{aligned}$$

Here we use $i \in \partial\Omega$ to denote the subdomains Ω_i^0 that touch the boundary $\partial\Omega$ with a face.

By Cauchy-Schwarz and Friedrichs inequalities we have

$$\begin{aligned}\sum_{i=1}^{N}|\bar{u}_i|^2 = \sum_{i=1}^{N}\left(\frac{1}{|\Omega_i^\delta|}\int_{\Omega_i^\delta} u dx\right)^2 \leq C\sum_{i=1}^{N}\frac{1}{H^2}\|u\|^2_{L^2(\Omega_i^\delta)} \\ \leq C\frac{1}{H^2}\|u\|^2_{L^2(\Omega)} \leq C\frac{1}{H^2}|u|^2_{H^1(\Omega)}.\end{aligned}$$

And for the cases $i \in \partial\Omega$, we can use a Poincaré inequality to obtain

$$\sum_{i\in\partial\Omega} |\bar{u}_i|^2 \leq C \sum_{i\in\partial\Omega} \frac{1}{H^2}\|u\|^2_{L^2(\Omega_i^\delta)} \leq C \sum_{i\in\partial\Omega} |u|^2_{H^1(\Omega_i^\delta)} \leq C|u|^2_{H^1(\Omega)}.$$

To bound the other terms $|v_i|^2_{H^1(\Omega)}, i = 1,\ldots,N$, we use $\theta_i(x), i = 1,\ldots,N$, introduced before. Consider $\tilde{v}_i \in \mathcal{V}_i^\delta$ defined as follows

$$\tilde{v}_i(x) = I_h(\theta_i(x)(u(x) - \bar{u}_i)).$$

Note that $\tilde{v}_i(x)$ is equal to $v_i(x)$ on $W_i^{\Gamma^\delta}$ and on $\partial\Omega_i^\delta$. On $\Omega_{i,ovl}^\delta$ and $\Omega_{i,non}^\delta$, v_i is discrete harmonic. Therefore, we have

$$|v_i|^2_{H^1(\Omega_i^\delta)} \leq |\tilde{v}_i|^2_{H^1(\Omega_i^\delta)}.$$

The rest of the proof will be devoted to the estimate of $|\tilde{v}_i|^2_{H^1(\Omega_i^\delta)}$ in terms of $|u|^2_{H^1(\Omega_i^\delta)}$.

Let K be an element of Ω_i^δ and let us denote $w_i = u - \bar{u}_i$ then

$$|\tilde{v}_i|^2_{H^1(K)} = |I_h(\theta_i w_i)|^2_{H^1(K)} \leq 2|\bar{\theta}_i w_i|^2_{H^1(K)} + 2|I_h((\bar{\theta}_i - \theta_i)w_i)|^2_{H^1(K)}. \tag{24}$$

Here, $\bar{\theta}_i$ is the average of θ_i on K, and I_h is the standard pointwise interpolation. To estimate the first part of (24) we use the fact that $|\bar{\theta}_i| \leq 1$, to obtain

$$|\bar{\theta}_i w_i|^2_{H^1(K)} = |\bar{\theta}_i(u - \bar{u}_i)|^2_{H^1(K)} \leq |u - \bar{u}_i|^2_{H^1(K)} = |u|^2_{H^1(K)}.$$

The last equality comes from the fact that $\bar{u}_i$ is a constant. For the second part of (24), according to an inverse inequality we have

$$|I_h((\bar{\theta}_i - \theta_i)w_i)|^2_{H^1(K)} \leq C\frac{1}{h^2}\|I_h((\bar{\theta}_i - \theta_i)w_i)\|^2_{L^2(K)}. \tag{25}$$

To obtain the bound for the right-hand side of (25), we consider the element K in four different situations corresponding to the four types of subregions into which the the subregion Ω_i^δ is split i.e., $\Omega_{i,non}^\delta$, $\Omega_i^{\delta H}$, $\Omega_i^{\delta\tilde{\delta}}$ and $\Omega_i^{\delta\delta}$.

The proof for the cases $K \subset \Omega_i^{\delta H}$ and $K \subset \Omega_i^{\delta\tilde{\delta}}$ are nearly the same, so we only consider one of them here. For $K \subset \Omega_i^{\delta H}$, since

$$\|\bar{\theta}_i - \theta_i\|^2_{L^\infty(K)} \leq C\left(\frac{h}{(2\delta+1)h}\right)^2,$$

we obtain

$$\frac{1}{h^2}\|I_h((\bar{\theta}_i - \theta_i)w_i)\|^2_{L^2(K)} \leq C\frac{1}{((2\delta+1)h)^2}\|w_i\|^2_{L^2(K)}.$$

Applying a technique developed in Dryja and Widlund [11], we obtain

$$\frac{1}{((2\delta+1)h)^2}\|w_i\|^2_{L^2(\Omega_i^{\delta H})} \leq C\left(\frac{H}{(2\delta+1)h}|w_i|^2_{H^1(\Omega_i^\delta)} + \frac{1}{H((2\delta+1)h)}\|w_i\|^2_{L^2(\Omega_i^\delta)}\right). \tag{26}$$

Using the fact $|w_i|^2_{H^1(\Omega_i^\delta)} = |u|^2_{H^1(\Omega_i^\delta)}$ and a Friedrichs inequality

$$\|w_i\|^2_{L^2(\Omega_i^\delta)} \le CH^2|u|^2_{H^1(\Omega_i^\delta)}. \tag{27}$$

Combining the estimates (26) and (27), we obtain

$$\frac{1}{((2\delta+1)h)^2}\|w_i\|^2_{L^2(\Omega_i^{\delta H})} \le C\frac{H}{(2\delta+1)h}|u|^2_{H^1(\Omega_i^\delta)}.$$

For the case when $K \subset \Omega_i^{\delta\delta}$, we use similar arguments as in Dryja, Smith and Widlund [10] to obtain

$$\sum_{K\in\Omega_i^{\delta\delta}} \frac{1}{h^2}\|I_h((\bar{\theta}_i-\theta_i)w_i\|^2_{L^2(K)} \le \sum_{K\in\Omega_i^{\delta\delta}} C\frac{1}{r^2}\|w_i\|^2_{L^2(K)}, \tag{28}$$

where $ch \le r \le C((\delta+1)h)$ is the distance to those "cut pieces". We have used here that $\theta_i(x)$ has the singular behavior C/r on $\Omega_i^{\delta\delta}$. We have then

$$\sum_{K\in\Omega_i^{\delta\delta}} \frac{1}{r^2}\|w_i\|^2_{L^2(K)} \le C\int_{ch}^{C(\delta+1)h}\int_\alpha r^{-2}r\|w_i\|^2_{L^\infty(\Omega_i^{\delta\delta})}d\alpha dr \tag{29}$$

and

$$\|w_i\|^2_{L^\infty(\Omega_i^{\delta\delta})} \le C\left(1+\log\left(\frac{H}{h}\right)\right)|u|^2_{H^1(\Omega_i^\delta)}. \tag{30}$$

For the inequality (30), we have used a well-known result (see Bramble [2])

$$\|u-\bar{u}_i\|^2_{L^\infty(\Omega_i^{\delta\delta})} \le \|u-\bar{u}_i\|_{L^\infty(\Omega_i^\delta)} \le C\left(1+\log\left(\frac{H}{h}\right)\right)\|u-\bar{u}_i\|^2_{H^1(\Omega_i^\delta)}$$

and that $\bar{u}_i$ is the average of u on Ω_i^δ; i.e., a Friedrichs inequality

$$\|u-\bar{u}_i\|^2_{H^1(\Omega_i^\delta)} \le C|u|^2_{H^1(\Omega_i^\delta)}.$$

Putting (29) and (30) together, we obtain

$$\sum_{K\in\Omega_i^{\delta\delta}} \frac{1}{r^2}\|w_i\|^2_{L^2(K)} \le C\left((1+\log(\delta+1))\left(1+\log\left(\frac{H}{h}\right)\right)\right)|u|^2_{H^1\Omega_i^\delta}. \tag{31}$$

For the case $K \subset \Omega_{i,non}^\delta$. If Ω_i^0 is a floating subdomain, which is to say that $\Omega_{i,non}^\delta$ does not touch $\partial\Omega$, then $\bar{\theta}_i-\theta_i$ is zero. If $\Omega_{i,non}^\delta$ touches the boundary $\partial\Omega$, then the estimate becomes

$$\begin{aligned} |v_i|^2_{H^1(\Omega_{i,non}^\delta)} &\le C\left(|u|^2_{H^1(\Omega_{i,non}^\delta)} + |\bar{u}_i|^2|\phi^i|^2_{H^1(\Omega_{i,non}^\delta)}\right) \\ &\le C\left(1+\log\left(\frac{H}{h}\right)\right)|u|^2_{H^1(\Omega_i^\delta)}. \end{aligned} \tag{32}$$

∎

4.3 The main theorem

We state the main theorem of this paper here, the proof follows directly from the abstract AS theory and the lemmas just proved.

Theorem 4.1 *The RASHO operator $\widetilde{P}^\delta$ is symmetric in the inner product $a(\cdot,\cdot)$, nonsingular, and bounded in the following sense*

$$C_0^{-2} a(u,u) \le a(\widetilde{P}^\delta u, u) \le C_1 a(u,u) \quad \forall u \in \widetilde{\mathcal{V}}^\delta. \tag{33}$$

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 δ.

We remark that the corresponding convergence rate estimate for the regular one-level AS [11], 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 case of large overlap, but roughly the same for small overlap. On the other hand, the upper bound C_1 of RASHO is smaller than the upper bound of AS. We can see this since $\widetilde{\mathcal{V}}_k^\delta \subset \mathcal{V}_k^\delta, \forall k$, implies that the positive numbers ϵ_{ij} defined in Lemma 4.1 are smaller for RASHO than the correspondent ϵ_{ij} for AS. Consequently, the spectral radius of $\mathcal{E}$ in RASHO is smaller. Because C_1 of RASHO is smaller, the numerical performance of RASHO presented in the next section is better than that of AS. We also remark that the results of the paper is for 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.

5 Numerical experiments

In this section, we present some numerical results for solving the Poisson's equation on the unit square with zero Dirichlet boundary conditions. We compare the performance of RASHO and AS preconditioned Conjugate Gradient methods 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 overlap.

We first discuss a few implementation issues related to the new preconditioner. In order to apply the RASHO/CG method, it is necessary to force the solution to belong to $\widetilde{\mathcal{V}}^\delta$. To do so, a pre-CG-computation is needed, and it is done through the formula (11). We note that $u = u^* - w \in \widetilde{\mathcal{V}}^\delta$, see Lemma 3.1, and therefore, we can apply the regular PCG to the RASHO preconditioned system (15). The AS/CG is

Table 1: RASHO and AS preconditioned CG for solving the Poisson's equation on a 128×128 mesh decomposed into $2 \times 2 = 4$ subdomains with overlap $= ovlp$. The AS/CG results are shown in ( ). The "+1" is for the preprocessing step needed for RASHO.

ovlp	iter	cond	max	min
h	42 (42)	129.(129.)	1.98 (1.98)	0.0154 (0.0154)
3h	24+1 (28)	48.4 (86.3)	1.94 (4.00)	0.0402 (0.0464)
5h	20+1 (23)	33.3 (51.8)	1.91 (4.00)	0.0574 (0.0773)
7h	18+1 (20)	27.2 (37.0)	1.89 (4.00)	0.0694 (0.1081)

Table 2: RASHO and AS preconditioned CG for solving the Poisson's equation on a $32 * DOM \times 32 * DOM$ mesh decomposed into $DOM \times DOM$ subdomains with overlap $= 3h$, i.e. $\delta = 1$.

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

the classical additive Schwarz preconditioned CG as described in [8]. We note that in the case $\delta = 0$, i.e. $ovlp = h$, RASHO and AS are the same.

The stopping condition for CG is to reduce the initial residual by a factor of 10^{-6}. The exact solution of the equation is $u(x, y) = e^{5(x+y)} \sin(\pi x) \sin(\pi y)$. All subdomain problems are solved exactly. The iteration counts (iter), condition numbers (cond), maximum (max) and minimum (min) eigenvalues of the preconditioned matrix are summerized in Table 1, and Table 2.

From Table 1 and Table 2, it is clear that RASHO/CG is always better than the classical AS/CG in terms of the iteration counts and condition numbers. Note that there is a practical suggestion for AS that the overlap should be $3h - 5h$ width. In this case the condition number of RASHO is almost twice smaller than AS. This is an important result since it is easy to modify a (parallel) AS/CG code to obtain a RASHO/CG implementation. Although we do not have any parallel results to report here, we are confident to predict that RASHO/CG would be even better than AS/CG on a parallel computer with distributed memory since much less communications are required. Also the local solvers in RASHO are slightly cheaper since the local solvers have slightly smaller numbers of unknowns than for the regular AS.

References

[1] S. Balay, W. Gropp, L. McInnes, and B. Smith, *The Portable Extensible Toolkit for Scientific Computing (PETSc)*, www.mcs.anl.gov/petsc, 2001.

[2] J. Bramble, *A second order finite difference analogue of the first biharmonic boundary value problem*, Numer. Math., 9(1966), pp. 236-249.

[3] S. Brenner and R. Scott, *The Mathematical Theory of Finite Element Methods*, Springer, 1994.

[4] X.-C. Cai, M. Casarin, F. Elliot, and O. Widlund, *Overlapping Schwarz algorithms for solving Helmholtz's equation,* Contemporary Mathematics, 218 (1998), pp. 391-399.

[5] X.-C. Cai, C. Farhat, and M. Sarkis, *A minimum overlap restricted additive Schwarz preconditioner and applications in 3D flow simulations*, Contemporary Mathematics, 218 (1998), pp. 479-485.

[6] X.-C. Cai, T. Mathew, and M. Sarkis, *Maximum norm analysis of overlapping nonmatching grid discretizations of elliptic equations*, SIAM J. Numer. Anal., 37 (2000), pp. 1709-1728.

[7] 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.

[8] M. Dryja and O. 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, 1987.

[9] M. Dryja, M. Sarkis, and O. Widlund, *Multilevel Schwarz methods for elliptic problems with discontinuous coefficients in three dimensions*, Numer. Math., 72(1996), pp. 313–348.

[10] M. Dryja, B. Smith, and O. Widlund, *Schwarz analysis of iterative substructuring algorithms for elliptic problems in three dimensions*, SIAM J. Numer. Anal., 31(1994), pp. 1662–1694.

[11] M. Dryja and O. Widlund, *Domain decomposition algorithms with small overlap*, SIAM J. Sci. Comp., 15(1994), pp. 604–620.

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

[13] A. Frommer and D. Szyld, *An algebraic convergence theory for restricted additive Schwarz methods using weighted max norms*, Research Report 00-3-23, Department of Mathematics, Temple University, 2000. (SIAM J. Numer. Anal., to appear)

[14] G. Golub and C. van Loan, *Matrix Computations* The Johns Hopkins University Press, 1983.

[15] W. Gropp, D. Kaushik, D. Keyes, and B. Smith, *Performance modeling and tuning of an unstructured mesh CFD application*, Proceedings of SC2000, IEEE Computer Society, 2000.

[16] M. Lesoinne, M. Sarkis, U. Hetmaniuk, and C. Farhat, *A linearized method for the frequency analysis of three-dimensional fluid/structure interaction problems in all flow regimes*, Comp. Meth. Appl. Mech. Eng., 2001. (to appear).

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

[18] Y. Saad and M. Schultz, *GMRES: A generalized minimum residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 7(1986), pp. 856–869.

[19] M. Sarkis, *Nonstandard coarse spaces and Schwarz methods for elliptic problems with discontinuous coefficients using non-conforming elements*, Numer. Math., 77(1997), pp. 383-406.

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