# A FAMILY OF OVERLAPPING SCHWARZ ALGORITHMS FOR NONSYMMETRIC AND INDEFINITE ELLIPTIC PROBLEMS

XIAO-CHUAN CAI*

**Abstract.** The classical Schwarz alternating method has recently been generalized in several directions. This effort has resulted in a number of new powerful domain decomposition methods for solving general elliptic problems, including the nonsymmetric and indefinite cases. In this paper, we present several overlapping Schwarz preconditioned Krylov space iterative methods for solving elliptic boundary value problems with operators that are dominated by the self-adjoint, second-order terms, but need not be either self-adjoint or definite. All algorithms discussed in this paper involve two levels of preconditioning, and one of the critical components is a global coarse grid problem. We show that, under certain assumptions, the algorithms are optimal in the sense that the convergence rates of the preconditioned Krylov iterative methods are independent of the number of unknowns of the linear system and also the number of subdomains. The optimal convergence theory holds for problems in both two- and three-dimensional spaces, and for both structured and unstructured grids. Some numerical results are presented also.

**1. Introduction.** In this paper, we present a family of overlapping domain decomposition methods for the solution of large, sparse, nonsymmetric and/or indefinite linear systems of equations obtained by discretizing elliptic partial differential equations. This family of methods originates from the classical Schwarz alternating algorithm, introduced in 1870 by H. A. Schwarz [37] in an existence proof for elliptic boundary value problems defined in certain irregular regions. This method has attracted much attention as a convenient computational method for the solution of a large class of elliptic or parabolic equations, see e.g., [14, 38], especially on parallel machines, [22]. There are essentially two ways to use the algorithm as a computational tool. The first approach is to use it directly on the continuous partial differential equation defined on a physical domain. The mesh partitioning and the PDE discretization are then carried out subdomain by subdomain, which may sometimes result in nonmatching grids between overlapping subdomains. The second approach is to use it on the already discretized PDE, i.e., a linear system of algebraic equations. In this approach, a global grid is assumed to have been introduced before the domain, or mesh, is partitioned into subdomains. We shall consider only the second approach. Some of the material presented in this paper can also be found the references [8, 10, 11, 12].

This family of overlapping Schwarz algorithms has been shown to be efficient and robust for solving differential equations of many different types under a wide range of circumstances. In this paper, we shall focus only on the class of nonsymmetric and/or indefinite second order elliptic finite element, or finite difference, equations. The solution of such problems is an important computational kernel in implicit methods, for example, the Jacobian problems that need to be solved in any Newton-like method used in the solution of nonlinear partial differential equations such as in computational fluid dynamics [9].

An efficient iterative algorithm for solving general elliptic equations requires three basic steps, namely (a) a discretization scheme, (b) a basic iterative method, and (c) a preconditioning strategy. There is a significant difference between symmetric

* Department of Computer Science, University of Colorado at Boulder, Boulder, CO 80309. cai@cs.colorado.edu . The work is supported in part by the NSF grant ASC-9457534, and in part by the NSF Grand Challenges Applications Group grant ASC-9217394 and by the NASA HPCC Group grant NAG5-2218.

and nonsymmetric problems, the latter being considerably harder to deal with both theoretically and algorithmically. The main reasons are the lack of a generally applicable discretization technique for the general nonsymmetric elliptic operator, the lack of "good" algebraic iterative methods (such as CG for symmetric, positive definite problems), and the incompleteness of the mathematical theory for the performance of the algebraic iterative methods that do exist, such as GMRES [35, 36]. By a "good" method, we mean a method that is provably convergent within memory requirements proportional to a small multiple of the number of degrees of freedom in the system, independent of the operator. One must assume that the symmetric part is positive definite and be able to afford amounts of memory roughly in proportion to the number of iterations, in order to obtain rapid convergence with GMRES. The task of finding a good preconditioner for nonsymmetric or indefinite problems is more important than for symmetric, positive definite problems, since, first, the preconditioner can force the symmetric part of the preconditioned system to be positive definite, and second, a better-conditioned system implies both more rapid convergence and smaller memory requirements. The focus of this paper is on the construction of efficient, parallel and scalable preconditioners by using domain decomposition methods.

Domain decomposition methods are commonly classified according to a few criteria. "Overlapping" and "nonoverlapping" methods are differentiated by the decomposition into territories on which the elemental subproblems are defined. We shall not discuss any nonoverlapping algorithms in this paper, interested readers should consult the paper [13] for recent progress. For a comparison of some of the overlapping and nonoverlapping algorithms, we refer to the paper [8]. Overlapping methods generally permit simple (Dirichlet) updating of the boundary data of the subregions at the expense of having to solve some larger linear systems, defined on subregions, per iteration from the redundantly degrees of freedom. An advantage of the overlapping methods, over non-overlapping substructuring type methods, is that the solution of the so-called interface problems (see [8, 13]) can always be avoided. We remark here that a general purpose, robust interface solver that guarantees the optimal convergence for the class of general variable coefficients, nonsymmetric and indefinite elliptic problems is yet to be introduced.

We shall restrict our attention to the so-called optimal algorithms, i.e., algorithms whose convergence rates are independent of the number of unknowns as well as the number of subregions. All the algorithms under consideration can be used in either two- or three-dimensional spaces, with either structured or unstructured meshes. A coarse space, which is used in all the algorithms, plays an extremely important role in obtaining the optimality. It essentially reduces the original nonsymmetric and/or indefinite elliptic problem to a positive definite problem [11, 12], which may not be symmetric. Most of the theory concerning the convergence rate of domain decomposition methods is in the framework of the Galerkin finite element method. In some cases the Galerkin results transfer immediately to finite difference discretizations, though this is less true for nonsymmetric problems than for symmetric. We shall describe the algorithms by using a matrix language which is independent of the underlying discretization schemes, however, we shall switch to the finite element language when discussing the convergence theory.

We remark that algorithms based on preconditioned iterative solution of the normal equations can also be used to solve nonsymmetric and/or indefinite linear systems, but are beyond the scope of this paper. Interested readers should consult, for exam-

ples, [3, 28, 32].

The paper is organized as follows. In the rest of this section, we shall define our model elliptic problem and its discretization. Section 2 is devoted to the description of an overlapping partitioning of the mesh, as well as algorithms for subdomain coloring. Both nested and nonnested coarse meshes are discussed in Section 2. The main algorithms of this paper are introduced in Section 3. This section includes the discussion of a number of optimal overlapping Schwarz algorithms including the additive Schwarz algorithm, the multiplicative Schwarz algorithm and some polynomial Schwarz algorithms. Several inexact subdomain problems solving techniques, and an algebraic extension of the Schwarz algorithms for general sparse linear systems are also discussed in Section 3. A brief overview of the available theory for the optimality of the Schwarz algorithms is given in Section 4. The paper ends with Section 5, which contains some numerical results.

We confine ourselves to the following model problem. Let $\Omega$ be a polygonal region, in $R^d$ $(d = 2, 3)$, with boundary $\partial\Omega$, and let

$$\left\{ \begin{array}{rcll} Lu & = & f & \text{in } \Omega \\ u & = & 0 & \text{on } \partial\Omega \end{array} \right. \tag{1}$$

be a second-order linear elliptic operator with a homogeneous Dirichlet boundary condition. Here

$$Lu(x) = -\sum_{i,j=1}^{d} \frac{\partial}{\partial x_i}\left(a_{ij}(x)\frac{\partial u(x)}{\partial x_j}\right) + 2\sum_{i=1}^{d} b_i(x)\frac{\partial u(x)}{\partial x_i} + c(x)u(x). \tag{2}$$

We assume that the matrix $\{a_{ij}(x)\}$ is symmetric and uniformly positive definite for any $x \in \Omega$ and the right-hand side $f \in L^2(\Omega)$. Only Dirichlet boundary conditions are considered here; however, the algorithms can be used to solve problems with other boundary conditions as well, such as Neumann or mixed boundary conditions. We also assume that a finite element mesh, structured or unstructured, has been introduced on $\Omega$. A finite element, or finite difference, discretization of the elliptic problem (1) on the given mesh in $\Omega$ gives us a linear system of algebraic equations

$$Bu^* = f, \tag{3}$$

where $B$ is an $n \times n$ sparse matrix and $n$ is the total number of interior nodes in $\Omega$. Here and in the rest of the paper $u^*$ denotes the exact solution of the linear system (3). We shall use $h$, even in the unstructured case, to characterize the mesh interval of the grid, which will be referred to as the $h$-level or fine grid. The nodal points in the fine grid will be referred to as the $h$-level nodes. We shall use the $n \times n$ matrix $A$ to denote the discretization of the symmetric, positive definite part of the operator $L$. Let $(\cdot, \cdot)$ denote the Euclidean inner product with the corresponding norm $\|\cdot\|$. We denote the energy norm associated with the matrix $A$ as

$$\|\cdot\|_A = (A\cdot, \cdot)^{1/2}.$$

In practice, there are many discretization schemes can be used to obtain the linear system (3), such as the artificial diffusion and streamline diffusion methods [23] and the methods in [1]. Multiple discretizations can also be combined in the same iterative process; see, e.g., [24]. The preconditioning techniques to be discussed in the next few sections can easily be used together with these discretization schemes.

**2. Overlapping partitioning, subdomain coloring and coarse spaces.** In this section, we discuss a number of issues, mostly non-numerical, related to the partitioning the finite element mesh, and the definition of a coarse mesh space, which is an important component of the algorithms of this paper.

We begin with the overlapping partitioning of the mesh. Let $\{\Omega_i, i = 1, \cdots, N\}$, be nonoverlapping subregions of $\Omega$, such that

$$\bigcup_{i=1}^{N} \bar{\Omega}_i = \bar{\Omega},$$

where $\bar{\Omega}$ means the closure of $\Omega$. Some earlier theory on Schwarz algorithms, [11, 12, 17, 18], required that the partitioning $\{\Omega_i\}$ forms a regular finite element subdivision of $\Omega$, but recent development shows that this requirement is not necessary [7]. These subdomains can be of any shapes. In the case of unstructured meshes, this partitioning is often achieved by using certain graph partitioning techniques; namely, we first define an adjacency graph for the finite element mesh, then partition the graph into a number of disjoint subgraphs. We refer to [10, 21, 26, 33] for issues of graph partitioning. We assume that the vertices of any $\Omega_i$, not on $\partial\Omega$, coincide with the $h$-level nodes. Following [11, 18], we can obtain an overlapping decomposition of $\Omega$, denoted by

$$\{\Omega_i^{'}, i = 1, \cdots, N\}.$$

Here $\Omega_i^{'}$ is obtained by extending each $\Omega_i$ to a larger region which is cut off at the physical boundary of $\Omega$. We assume that

$$\text{Distance}(\partial\Omega_i^{'} \cap \Omega, \partial\Omega_i \cap \Omega) \geq \ \delta, \ \ \forall i,$$

for a constant $\delta > 0$. Here 'Distance' is in the usual Euclidean sense. In the uniform mesh case, $\delta$ is usually equal to an integer multiple of the mesh size $h$. $\delta$ is an important parameter in these overlapping algorithms. Usually, using a larger overlapping can result in a reduced total number of iterations, however, per-iteration arithmetic operations and local memory requirement may increase. Let $n_i$ be the total number of $h$-level interior nodes in $\Omega_i^{'}$, and $B_i$ the $n_i \times n_i$ stiffness matrix corresponding to the discretization of $L$ on the fine grid in $\Omega_i^{'}$, with a zero Dirichlet boundary condition on $\partial\Omega_i^{'}$. Since the matrices $B_i$ are used only in the preconditioner part of the algorithms, they need not be calculated exactly. A detailed discussion on the use of inexact subdomain solvers can be found in Section 3.4. The size of the matrix $B_i$ depends not only on the size of the substructure $\Omega_i$ but also on the degree of overlap. The cost for solving the linear systems corresponding to the matrix $B_i$ is determined not only by the size of the matrix but also by the type of solver. We note that a less accurate solver, such as an ILU [30], or ILUT [34], with a small number of fill-ins and a relatively large drop tolerance, can keep the overall cost down, even if the overlap is not too small.

When using some of the multiplicative algorithms (Section 3.1), the subdomains are usually colored with the purpose of reducing the number of the sequential steps and speed up the overall convergence. The coloring is realized as follows. Associated with the decomposition $\{\Omega_j^{'}\}$, we define an undirected graph in which nodes represent the extended subregions and the edges intersections of the extended subregions. This graph can be colored by using colors $1, \cdots, J$, such that no connected nodes have the

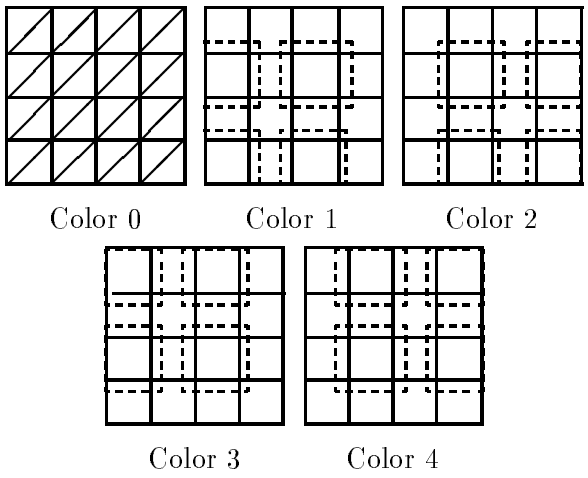


FIG. 1. *The coloring pattern of 16 fine grid overlapped subregions and a coarse grid region. Color "0" is for the global coarse grid. The extended subregions of the other colors are indicated by the dotted boundaries.*

same color. Obviously, colorings are not unique. Simple greedy heuristic subgraph coloring algorithms have been discussed in the literature; see for examples, [10]. Numerical experiments support the expectation that minimizing the number of colors enhances convergence. An optimal five-color strategy ($J = 4$) is shown for the decomposition in Figure 1, in which the total number of subregions (including the coarse grid on the global region) is $N + 1 = 17$.

Let $R_i$ be an $n_i \times n$ matrix representing the algebraic restriction of an $n$-vector on $\Omega$ to the $n_i$-vector on $\Omega_i^{'}$. Thus, if $v$ is a vector corresponding to all the $h$-level interior nodes in $\Omega$, then $R_i v$ is a vector corresponding to the $h$-level interior nodes in $\Omega_i^{'}$. The transpose $(R_i)^t$ is an extension-by-zero matrix, which extends a length $n_i$ vector to a length $n$ vector by padding with zero.

All the algorithms discussed in the next section involve a coarse level discretization. Let us define it here. Suppose that there is another mesh defined on $\Omega$, which contains $n_0$ nodes, and is coarser than the fine mesh. Let $B_0$ be the discretization of $L$ on this coarse mesh. Let $R_0^T$ be an extension operator, which maps any coarse mesh vector to the corresponding fine mesh vector. There is a variety of ways that one can define such an operator. Here we discuss only one example in the finite element context. Let $\phi_j(x)$ be the basis function defined at the $j$th coarse node. Let $\{x_i \in \Omega, i = 1, \cdots, n\}$ be the fine mesh nodes. Then the $n \times n_0$ matrix $R_0^T = \{r_{ij}\}$ can be defined by $r_{ij} = \phi_j(x_i)$. $R_0$ is the transpose of $R_0^T$, and is used as a restriction operator that maps a fine mesh vector to a coarse mesh vector. In practice, the coarse mesh space needs not to be a subspace of the fine mesh space; we refer to [7] for a detailed discussion. We shall use $\Omega_0^{'}$ to denote the coarse grid, which is always assumed to have color 0.

We conclude this section by introducing several frequently used notations. For each subdomain $\Omega_i^{'}$, $i = 0, \cdots, N$, we define two $n \times n$ matrices

$$M_i^{-1} = R_i^T B_i^{-1} R_i \quad \text{and } P_i = M_i^{-1} B.$$

For $j = 0, 1, \cdots, J$, we denote by $Q_j$ the sum of all $P_i^{'}$ and by $N_j^{-1}$ the sum of all $M_i^{-1^{'}}$ that correspond to subregions of the $j^{th}$ color. These matrices will serve as the basic building blocks of the overlapping Schwarz algorithms to be discussed.

**3. Some Schwarz algorithms.** In this section, we describe several overlapping Schwarz type algorithms constructed by using $P_i$ as the basic building blocks. We shall begin with the so-called multiplicative Schwarz algorithm, which is a direct extension of the classical Schwarz alternating algorithm. Then, we discuss a much simpler, additive version of the multiplicative Schwarz algorithm, which will be referred to as the additive Schwarz algorithm. In the third subsection, we introduce a family of Schwarz algorithms constructed by using a multivariable matrix-valued polynomial with $P_i$ as the variables. In fact, both additive and multiplicative Schwarz algorithms are special cases of this family of *polynomial Schwarz* algorithms. Finally, in the last subsection, we briefly discuss an algebraic extension of the overlapping Schwarz algorithms, introduced recently in [10], for solving general sparse linear systems.

**3.1. Multiplicative Schwarz method (MSM).** Unlike other preconditioners, such as the additive Schwarz, MSM algorithm can be employed either as an iterative algorithm by itself or used as a preconditioner. As an iterative algorithm, or equivalently as a preconditioner accelerated by a simple Richardson method, MSM is rather sensitive to some of the problem parameters, such as the size of the first order terms in the partial differential operator, and sometimes loses its convergence, see for example Table 1. However, it is an excellent and robust preconditioner, especially when accelerated by a Krylov space iterative method, such as the GMRES. Along with the other algorithms to be described below, we shall normally employ it as a preconditioner for GMRES, but because of its historical importance, and to illustrate certain robustness advantages of acceleration, we also include the Richardson version in our discussion. In this paper, we shall use the abbreviation MSM for the multiplicative Schwarz-preconditioned GMRES method, and MSR for the simple Richardson process that corresponds to the classical Schwarz alternating algorithm with an extra coarse grid solver.

To obtain parallelism, one needs a good subdomain coloring strategy so that a set of independent subproblems can be introduced within each sequential step and the total number of sequential steps can be minimized. A detailed description of the coloring algorithm and its theoretical aspects can be found in [4, 12, 25]. We now describe the MSR algorithm in terms of a subspace correction process.

Let $u^k$ be the current approximate solution. Then $u^{k+1}$ is computed as follows. For $j = 0, 1, \cdots, J$:

(i) Compute the residual in subregions with the $j^{th}$ color:

$$r^{k+\frac{j}{J+1}} = f - Bu^{k+\frac{j}{J+1}}.$$

(ii) Solve for the subspace correction in all $\Omega_i^{'}$s that share the $j^{th}$ color:

$$B_i e^{k+\frac{j}{J+1}} = R_i r^{k+\frac{j}{J+1}}.$$

(iii) Update the approximate solution in all $\Omega_i^{'}$s that share the $j^{th}$ color:

$$u^{k+\frac{j+1}{J+1}} = u^{k+\frac{j}{J+1}} + R_i^T e^{k+\frac{j}{J+1}}.$$

At each iteration, every subproblem is solved once. For $j \neq 0$, applications of operators $R_j$ and $R_j^T$ do not involve any arithmetic operations. For $j \neq 0$, within

each series of steps (i)–(iii), the operations in subregions sharing the same color can be done in parallel. MSR can also be written in the following more compact form: For a given initial approximate solution $u^0$, and $k = 0, 1, \cdots$,

$$u^{k+1} = E_{J+1} u^k + g,$$

where the error propagation operator $E_{J+1}$ is defined as $E_{J+1} = (I - Q_J) \cdots (I - Q_0)$ and $g \equiv g_J$ is computed at a pre-iteration step by the following $J + 1$ sequential steps:

$$\begin{aligned} g_0 &= N_0^{-1} f \\ g_1 &= g_0 + N_1^{-1}(f - B g_0) \\ &\vdots \\ g_J &= g_{J-1} + N_J^{-1}(f - B g_{J-1}). \end{aligned}$$

Next, we shall discuss an accelerated version of MSR. We begin with the observation that if the matrix $I - E_{J+1}$ is invertible, then the exact solution of equation (3) also satisfies

$$(I - E_{J+1}) u^* = g, \tag{4}$$

which is sometimes referred to as the *transformed*, or preconditioned, system corresponding to (3). We next observe that for a given vector $v \in R^n$, the matrix-vector product $(I - E_{J+1})v$, denoted as $v_J$, can be computed in a manner similar to that of $g$, namely,

$$\begin{aligned} v_0 &= Q_0 v \\ v_1 &= v_0 + Q_1(v - v_0) \\ &\vdots \\ v_J &= v_{J-1} + Q_J(v - v_{J-1}). \end{aligned} \tag{5}$$

Now, the multiplicative Schwarz preconditioned GMRES method (MSM) can be described as follows: Find the solution of equation (3) by solving the equation (4) with the GMRES method for a given initial guess and inner product.

Even in the case that the matrix $B$ is symmetric positive definite, the iteration matrix $I - E_{J+1}$ is not symmetric. An obvious symmetrization exists, upon which a conjugate gradient method can be used as the acceleration method; however, we shall not emphasize the case of a symmetric $B$ in this paper.

**3.2. Additive Schwarz algorithm (ASM).** An additive variant of the Schwarz alternating method was originally proposed in [15, 17, 31] for selfadjoint elliptic problems and extended to nonselfadjoint elliptic cases in [5, 11]. The idea is simply to give up the data dependency between the subproblems defined on subregions with different colors, as in going from Gauss-Seidel to Jacobi. Instead of iterating with (5), one uses

$$\begin{aligned} v_0 &= Q_0 v \\ v_1 &= v_0 + Q_1 v \\ &\vdots \\ v_J &= v_{J-1} + Q_J v. \end{aligned} \tag{6}$$

Of course, similar changes have to be made to the right-hand side vector $g$. Coloring does not play a role at all in (6). Because of the lack of data dependency, the method is

usually not to be recommended as a simple Richardson process (it may not converge), but as a preconditioner for some algebraic iterative methods of CG type. We denote by $M_{\rm ASM}^{-1}$ the preconditioning part of (6). Following [11] and using the notation of the previous subsection, we can define the inverse of the matrix $M_{\rm ASM}$, referred to as the additive Schwarz preconditioner, as

$$M_{\rm ASM}^{-1} = R_0^T B_0^{-1} R_0 + R_1^T B_1^{-1} R_1 + \cdots + R_N^T B_N^{-1} R_N. \tag{7}$$

The key ingredients for the success of the ASM are the use of overlapping subregions and the incorporation of a coarse grid solver. At each iteration, all subproblems are solved once. It is obvious that all subproblems are independent of each other and can therefore be solved in parallel. The ASM discussed in this subsection can be used recursively for the solving the subdomain problems. The result is the multilevel ASM, as developed in [2, 12, 19, 40, 41].

**3.3. Some polynomial Schwarz algorithms.** In this subsection, we discuss a family of Schwarz algorithms constructed by using some matrix-valued, multivariable polynomials. The previously discussed ASM and MSM algorithms can be viewed as two extreme cases of polynomial Schwarz algorithms, namely polynomials with the lowest and highest possible degrees. We remark here that the degree of the polynomial usually equals the number of sequential steps of the algorithm. Related subjects can be found in the papers [3, 6, 16, 27, 39], and references therein.

Let us define

$$T = poly(P_0, P_1, \cdots, P_N)$$

as a matrix-valued polynomial in the variables $P_i^{'}$, and we assume the polynomial satisfies $poly(0, \cdots, 0) = 0$, which simply means that the constant term in the polynomial is zero. It is not difficult to see that $Tu^* \in V^h$ can be computed without knowing $u^*$ itself. This is because that $P_i u^*$, $i = 0, \cdots, N$, can be computed directly from the right-hand side function $f$. By denoting $g = Tu^*$, we can define a new linear system

$$Tu^* = g \tag{8}$$

which will be referred to as the *transformed system* of (3). It can be proved that if the matrix $T$ is invertible, then the equation (8) has the same solution as the equation (3). To obtain the matrix $T$ explicitly is usually not possible; however, for any $v \in R^n$, the matrix-vector multiply $Tv$ can be computed easily. This makes the linear system (8) a good candidate for using Krylov space iterative methods.

We next look at some special examples. The first and simplest, in which the degree of $poly(\cdots)$ is one, is the additive Schwarz method, in which the operator has the form

$$T_{asm} = Q_0 + Q_1 + \cdots + Q_J.$$

The second example is the so-called multiplicative Schwarz operator

$$T_{msm} = I - E_{J+1},$$

where $I$ is the identity matrix and $E_{J+1} = (I - Q_0)(I - Q_1) \cdots (I - Q_J)$. The degree of this polynomial depends on the number of colors, and the exact form of the polynomial

depends on how the subregions are colored. The third example, which is a hybrid additive and multiplicative Schwarz algorithm (AMSM), was introduced in [6], and

$$T_{amsm} = \omega P_0 + I - E_J, \tag{9}$$

where $\omega$ is a balancing parameter and $E_J = (I - Q_1) \cdots (I - Q_J)$, without containing the coarse operator term. Numerical experiments suggest that $\omega = 1$ is usually a good choice, although the corresponding theory is yet to be established. The algorithm can be viewed as a combination of the additive and multiplicative Schwarz methods. There are two major advantages. First, it converges faster than the additive Schwarz algorithm because of the extra local data dependency. Secondly, it is more parallelizable than the multiplicative Schwarz algorithm since the global coarse problem can now be solved simultaneously with the rest of the local problems.

It is important to note that even if the original equation (3) is not well-conditioned, the transformed systems can be uniformly well-conditioned and, more importantly, the transformed system can be so arranged that a highly parallelizable algorithm can be developed for solving it.

**3.4. Using inexact subdomain solvers.** Using an inexact solver for the interior subproblems, or an exact solver for approximate interior subproblems, can significantly reduce the overall computational complexity. This is, in fact, one of the major advantages of domain decomposition methods, in that they allow the use of fast solvers designed for special differential operators on regions of special shape. A somewhat disappointing experimental observation is that inexact solutions seem not to work well for the coarse grid solver. In fact, the existing theory for MSM [12], as well as the theory for ASM [11], requires an exact solve on the coarse grid.

There are essentially two ways to introduce an inexact subproblem solver. The first method involves an approximation to the differential operator $L$. In each subdomain $\Omega_k^{'}$, $k \neq 0$, $L$ is replaced by a certain spectrally equivalent differential operator $L_k$, which is usually chosen to have constant coefficients, or has other special properties, so that a fast solver, such as an FFT-based method, can be used to solve the corresponding discretized problem. As an example, we mention that if $L$ is a general operator, defined as in (2), then $L_k$ can be defined as

$$L_k u = - \sum_{i,j=1}^{d} c_k \frac{\partial}{\partial x_i} \frac{\partial u}{\partial x_j},$$

where $c_k$ is an averaged eigenvalue of the matrix $\{a_{ij}(x_k)\}$ and $x_k$ is a fixed point in $\Omega_k^{'}$. In this case, the subdomain matrix $B_i$, used in any of the Schwarz algorithms discussed in the previous subsections, can be replaced by a discretization of $L_k$ with a zero Dirichlet boundary condition on $\partial\Omega_k^{'}$.

The second class of inexact subproblem solvers can be defined at the algebraic level. We assume that the matrices $B_k$ have already been obtained by the discretization of certain differential equations on $\Omega_k^{'}$. In this case, using inexact solver is understood as solving the subdomain linear system

$$B_k x = b,$$

inexactly. Here $x, b \in R^{n_k}$. For examples, the above linear system can be "solved", by (1) a few multigrid cycles [29]; or (2) a few Gauss-Seidel (SOR, SSOR, Jacobi) iterations; or (3) replacing $B_k$ with its ILU [30] or ILUT factorization [34], etc.

**3.5. Algebraic Schwarz algorithms.** According to [10], the previously studied Schwarz framework can also be extended for solving general sparse linear systems. The fundamental principle underlying this extension is to replace the *domain of definition* of the problem by the *adjacency graph* of the sparse matrix, i.e., the graph that represents its non-zero pattern. We note that by switching from a domain to a graph the concept of Euclidean distance, which plays an important role in the optimality analysis of these domain decomposition methods, is lost. It was shown in [10] that, mostly by means of numerical experiments, that the efficiency of the overlapping methods can be preserved to some extent with certain well-balanced overlapping graph decomposition.

Suppose $B = \{b_{ij}\}$ is an $n \times n$ sparse matrix. To describe a model algebraic Schwarz algorithm, let us define the graph $G = (W, E)$, where the set of vertices $W = \{1, \cdots, n\}$, represents the $n$ unknowns and the edge set $E = \{(i,j) \mid b_{ij} \neq 0\}$ represents the pairs of vertices that are coupled by a nonzero element in $B$. Let us assume that the non-zero pattern is symmetric, and therefore the adjacency graph $G$ is undirected. For the remaining discussion, we assume that the graph partitioning has been applied and has resulted in a number $N$ of subsets $W_i$ whose union is $W$,

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

We will denote by $N_i$ the vector space spanned by the set $W_i$ in $R^n$ and by $m_i$ its dimension. For each subspace $N_i$ we define a corresponding submatrix. In matrix terms, this is defined by the sub-identity matrix $I_i$ of size $n \times n$ whose diagonal elements are set to one if the corresponding node belongs to $W_i$ and to zero otherwise. With this we define the matrix,

$$B_i = I_i B I_i \ ,$$

which is an extension to the whole subspace, of the restriction of $B$ to $N_i$. This is sometimes termed the *section* of $B$ on $N_i$. Its action on a vector is to project it on $N_i$, then apply $B$ to the result and finally project the result back onto $N_i$. Note that although $B_i$ is not invertible, we can invert its restriction to the subspace spanned by $W_i$, and define

$$B_i^{-1} \equiv I_i \left( (B_i)_{|N_i} \right)^{-1} I_i$$

With this definition of $B_i^{-1}$, the Schwarz algorithms can be defined the same as in the previous subsections. The only missing piece is the coarse preconditioner. As indicated in [10], without further geometric information of the problem, to define a coarse preconditioner is generally very difficult.

**4. Convergence theory.** We now discuss very briefly a theory that can provide with us some understanding of these Schwarz preconditioners. All the discussions are based on an assumption that there is an underling finite element space. For simplicity, we consider only the piecewise linear finite element case. Let $b(u,v)$ be the bilinear form associated with the Dirichlet problem (1). The convergence of MSR has been proved in [12], under certain assumptions. The rate of convergence is

$$\|u^k - u\|_A \leq \left( \sqrt{1 - \frac{C_{\text{MSR}}}{(J+1)^2}} \right)^k \|u^0 - u\|_A,$$

where $C_{\rm MSR} > 0$ is a constant independent of $h$, $H$ and $J$. The estimate holds in both two- and three-dimensional spaces. Here $H$ is the diameter of the subdomains. The assumptions include: (1) the overlap is uniform and must be $\mathcal{O}(H)$; (2) $H$ must be sufficiently small; and (3) the number of colors, $J$, must be independent of the size of the subregions $H$. The same estimate, with a different constant, holds for MSR with either exact or spectrally equivalent inexact solvers.

For the accelerated version MSM, under the same assumptions, we have that there exist two constants $C_{\rm MSM} > 0$ and $c_{\rm MSM} > 0$, independent of both $h$ and $H$, such that the transformed system is uniformly bounded:

$$\|(I - E_{J+1})x\|_A \leq C_{\rm MSM}\|x\|_A, \quad \forall x \in R^n, \tag{10}$$

and the symmetric part of the transformed system is positive definite in the inner product $(A\cdot,\cdot)$:

$$(A(I - E_{J+1})x, x) \geq c_{\rm MSM}\|x\|_A^2, \quad \forall x \in R^n. \tag{11}$$

For the additive Schwarz algorithm, it was shown [5, 11] that, in the piecewise linear finite element case, the preconditioner $M_{\rm ASM}^{-1}$ is optimal under the same first two assumptions made for MSM in the sense that there exist two constants $C_{\rm ASM} > 0$ and $c_{\rm ASM} > 0$, which may be different for exact and inexact subdomain solvers and are independent of both $h$ and $H$, such that the preconditioned linear system is uniformly bounded:

$$\|M_{\rm ASM}^{-1}Bx\|_A \leq C_{\rm ASM}\|x\|_A, \quad \forall x \in R^n \tag{12}$$

and the symmetric part of the preconditioned linear system is positive definite in the inner product $(A\cdot,\cdot)$

$$(AM_{\rm ASM}^{-1}Bx, x) \geq c_{\rm ASM}\|x\|_A^2, \quad \forall x \in R^n. \tag{13}$$

Similar boundedness results hold also for the operator (9) with certain well chosen parameter $\omega$, see [6]. The extensions of the above results to the unstructured grids cases can be found in [7].

In the case $B = A$, which means that the original elliptic operator is symmetric positive definite, the left-preconditioned system is symmetric positive definite in the $(A\cdot,\cdot)$ inner product; thus one can use a CG method. In the nonsymmetric case, the preconditioned system is nonsymmetric regardless of inner product. Therefore, instead of the $A$-inner product, we usually use the Euclidean inner product for practical implementations. By giving up the symmetry requirement of the preconditioned system, we could also use ASM as a right-preconditioner. Neither of the pair of estimates (12) and (13) has been proved in the $L^2$ norm, but in the numerical experiments section, variability in ASM convergence rates measured (as is customary) with respect to $L^2$ residuals clearly diminishes as mesh and subdomain parameters are both refined, leading us to conjecture that analogous results hold.

We remark that the bounds (10), (11), (12) and (13) can be used to estimate, theoretically, the number of iterations for some of the Krylov space iteration methods, such as GMRES. As is well-known, the GMRES method, introduced in [35], is mathematically equivalent to the generalized conjugate residual (GCR) method [20] and can be used to solve the linear system of algebraic equations:

$$Px = b, \tag{14}$$

where $P$ is a nonsingular matrix, which may be nonsymmetric or indefinite, and $b$ is a given vector in $R^n$. In this paper, $P$ is one of the transformed systems $T = poly(P_0, \cdots, P_N)$. According to the theory of [20, 5], the rate of convergence of the GMRES method can be estimated by the ratio of the minimal eigenvalue of the symmetric part of the operator to the norm of the operator. Those two quantities are defined by $c_P = \inf_{x \neq 0}(x, Px)_A/(x,x)_A$ and $C_P = \sup_{x \neq 0} \|Px\|_A/\|x\|_A$, where $(\cdot,\cdot)_A$ is our $A$-inner product on $R^n$ that induces the norm $\|\cdot\|_A$. Following [20], the rate of convergence can be characterized, not necessarily tightly, as follows: If $c_P > 0$, which means that the symmetric part of the operator $P$ is positive definite with respect to the inner product $[\cdot,\cdot]$, then the GMRES method converges and at the $m^{th}$ iteration, the residual is bounded as

$$\|r_m\|_A \ \leq \ (1 \ - \ \frac{c_P^2}{C_P^2})^{m/2} \ \|r_0\|_A,$$

where $r_m = b - Px_m$. The algorithm is parameter-free and quite robust. Its main disadvantage is its linear-in-$m$ memory requirement. To fit the available memory, one is sometimes forced to use the $k$-step restarted GMRES method [35].

**5. Numerical experiments.** We present a few numerical experiments in this section to illustrate the convergence behavior of some of the Schwarz algorithms. For comparison, We also include some results obtained by using ILU($k$) as preconditioners. Two test problems will be considered. A more complete comparison of overlapping Schwarz algorithms with other domain decomposition algorithms can be found in [8]. Some three dimensional experiences with overlapping Schwarz algorithms can be found in [22].

**Example 1**. $Lu = -\triangle u + \delta u_x + \delta u_y$.

**Example 2**. $Lu = -((1 + \frac{1}{2}\sin(50\pi x)u_x)_x - ((1 + \frac{1}{2}\sin(50\pi x)\sin(50\pi y)u_y)_y$
$+ 20\sin(10\pi x)\cos(10\pi y)u_x - 20\cos(10\pi x)\sin(10\pi y)u_y - 70u.$

In all the tests, $\Omega = [0,1] \times [0,1]$, and homogeneous Dirichlet boundary conditions are prescribed on $\partial\Omega$. GMRES is used as the Krylov iterative method, and the iteration is stopped when the initial residual is reduced by $10^{-5}$. A one-point-per-subdomain coarse grid solver is used in all the tests. All the subdomain problems are solved exactly. For **Example 1**, a uniform $128 \times 128$ grid is used on $\Omega$, the number of subdomains is $64 = 8 \times 8$, and the overlap is $4h$ ($h = 1/128$). The first order terms are discretized by two schemes as indicated in Table 1. For **Example 2**, we test a few different fine mesh sizes, as given in Table 2, and the overlapping size is always set to be 25% of the size of the unextended subdomain in both $x$ and $y$ directions.

**6. Conclusions.** In this chapter, we discussed a family of parallel overlapping Schwarz type domain decomposition algorithms in the framework of preconditioned Krylov space iterative methods. The preconditioners, which are optimal in the sense of possessing mesh and subdomain parameter independent convergence rates, are constructed by using a multiple discretization of the partial differential equation, in local subdomains and also on a coarser grid. When the subdomains are properly colored, the algorithms are fully parallel since sub-problems defined on different subdomains can be mapped on to different processors and solved independently at each Krylov iteration.

TABLE 1
*Iteration count for solving* **Example 1**.

| Methods | $8 \times 8 = 64$ subdomains | | | | | |
|---|---|---|---|---|---|---|
| **Central-difference Method** | | | | | | |
| $\delta =$ | 1 | 5 | 10 | 50 | 100 | 150 |
| MS-Richardson | 6 | 6 | 6 | 10 | 35 | $\infty$ |
| MS-GMRES | 4 | 4 | 4 | 5 | 7 | 9 |
| AMS-GMRES | 7 | 8 | 9 | 13 | 17 | 20 |
| AS-GMRES | 10 | 11 | 11 | 15 | 20 | 23 |
| ILU(0)-GMRES | 60 | 84 | 81 | 59 | 41 | 27 |
| ILU(1)-GMRES | 38 | 53 | 51 | 34 | 22 | 15 |
| ILU(2)-GMRES | 31 | 46 | 42 | 28 | 19 | 13 |
| **Upwind-difference Method** | | | | | | |
| $\delta =$ | 10 | 50 | 100 | 500 | 1000 | 10000 |
| MSR | 9 | 12 | 12 | 12 | 12 | 12 |
| MS-GMRES | 7 | 7 | 6 | 6 | 6 | 6 |
| AMS-GMRES | 9 | 11 | 12 | 12 | 12 | 12 |
| AS-GMRES | 14 | 15 | 16 | 17 | 17 | 18 |
| ILU(0)-GMRES | 82 | 61 | 50 | 23 | 16 | 6 |
| ILU(1)-GMRES | 51 | 36 | 28 | 12 | 9 | 4 |
| ILU(2)-GMRES | 42 | 30 | 24 | 11 | 8 | 4 |

TABLE 2
*Iteration count for solving* **Example 2**.

| | $H = 1/8$ | | |
|---|---|---|---|
| $h =$ | 1/32 | 1/64 | 1/128 |
| MS-Richardson | $\infty$ | $\infty$ | $\infty$ |
| MS-GMRES | 16 | 15 | 15 |
| AS-GMRES | 29 | 26 | 25 |
| ILU(0)-GMRES | 44 | 78 | 312 |
| ILU(1)-GMRES | 28 | 44 | 99 |
| ILU(2)-GMRES | 22 | 36 | 76 |

[1] R. E. Bank, J. F. Bürgler, W. Fichtner, and R. K. Smith, *Some upwinding techniques for the finite element approximation of convection-diffusion equations*, Numer. Math., 58 (1990), pp. 185 – 202.

[2] P. E. Bjørstad and M. D. Skogen, *Domain decomposition algorithms of Schwarz type, designed for massively parallel computers*, in Fifth International Symposium on Domain Decomposition Methods for Partial Differential Equations, T. F. Chan, D. E. Keyes, G. A. Meurant, J. S. Scroggs, and R. G. Voigt, eds., Philadelphia, PA, 1992, SIAM.

[3] J. H. Bramble, Z. Leyk, and J. E. Pasciak, *Iterative schemes for non-symmetric and indefinite elliptic boundary value problems*, Math. Comp., 60 (1993), pp. 1–22.

[4] J. H. Bramble, J. E. Pasciak, J. Wang, and J. Xu, *Convergence estimates for product iterative methods with applications to domain decomposition*, Math. Comp., 57 (1991), pp. 1–21.

[5] X.-C. Cai, *Some Domain Decomposition Algorithms for Nonselfadjoint Elliptic and Parabolic Partial Differential Equations*, PhD thesis, Courant Institute of Mathematical Sciences, September 1989. Tech. Rep. 461, Department of Computer Science, Courant Institute.

[6] ———, *An optimal two-level overlapping domain decomposition method for elliptic problems in two and three dimensions*, SIAM J. Sci. Comput., 14 (1993), pp. 239–247.

[7] ———, *The use of pointwise interpolation in domain decomposition methods with non-nested meshes*, SIAM J. Sci. Comput., 16 (1995). To appear.

[8] X.-C. Cai, W. D. Gropp, and D. E. Keyes, *A comparison of some domain decomposition and ILU preconditioned iterative methods for nonsymmetric elliptic problems*, Numer. Lin. Alg. Applics, (1994). To appear.

[9] X.-C. Cai, W. D. Gropp, D. E. Keyes, and M. D. Tidriri, *Newton-Krylov-Schwarz methods in CFD*, in Proceedings of the International Workshop on the Navier-Stokes Equations, Notes in Numerical Fluid Mechanics, R. Rannacher, ed., Vieweg Verlag, Braunschweig, 1994. To appear.

[10] X.-C. Cai and Y. Saad, *Overlapping domain decomposition algorithms for general sparse matrices*, Tech. Report Preprint 93-027, Army High Performance Computing Research Center, University of Minnesota, 1993.

[11] X.-C. Cai and O. Widlund, *Domain decomposition algorithms for indefinite elliptic problems*, SIAM J. Sci. Statist. Comput., 13 (1992), pp. 243–258.

[12] ———, *Multiplicative Schwarz algorithms for nonsymmetric and indefinite elliptic problems*, SIAM J. Numer. Anal., 30 (1993), pp. 936–952.

[13] T. F. Chan and T. P. Mathew, *Domain decomposition preconditioners for convection diffusion problems*, in Sixth Conference on Domain Decomposition Methods for Partial Differential Equations, A. Quarteroni, J. Périaux, Y. A. Kuznettsov, and O. Widlund, eds., Providence, RI, 1993, AMS.

[14] T. F. Chan and T. P. Mathew, *Domain decomposition algorithms*, Acta Numerica, 1994, pp. 61–143.

[15] M. Dryja, *An additive Schwarz algorithm for two- and three-dimensional finite element elliptic problems*, in Domain Decomposition Methods, T. Chan, R. Glowinski, J. Périaux, and O. Widlund, eds., Philadelphia, PA, 1989, SIAM.

[16] M. Dryja, B. F. Smith, and O. B. Widlund, *Schwarz analysis of iterative substructuring algorithms for problems in three dimensions*, SIAM J. Numer. Anal., 31 (1994). To appear.

[17] M. Dryja and O. B. Widlund, *An additive variant of the Schwarz alternating method for the case of many subregions*, Tech. Report 339, also Ultracomputer Note 131, Department of Computer Science, Courant Institute, 1987.

[18] ———, *Towards a unified theory of domain decomposition algorithms for elliptic problems*, in Third International Symposium on Domain Decomposition Methods for Partial Differential Equations, held in Houston, Texas, March 20-22, 1989, T. Chan, R. Glowinski, J. Périaux, and O. Widlund, eds., SIAM, Philadelphia, PA, 1990.

[19] ———, *Multilevel additive methods for elliptic finite element problems*, in Parallel Algorithms for Partial Differential Equations, Proceedings of the Sixth GAMM-Seminar, Kiel, January 19–21, 1990, W. Hackbusch, ed., Braunschweig, Germany, 1991, Vieweg & Son.

[20] S. C. Eisenstat, H. C. Elman, and M. H. Schultz, *Variational iterative methods for nonsymmetric systems of linear equations*, SIAM J. Numer. Anal., 20 (1983), pp. 345–357.

[21] C. Farhat and M. Lesoinne, *Automatic partitioning of unstructured meshes for the parallel*

*solution of problems in computational mechanics*, Int. J. Numer. Meth. Engrg., 36 (1993), pp. 745–764.

[22] W. D. Gropp and B. F. Smith, *Experiences with domain decomposition in three dimensions: Overlapping Schwarz methods*, in Sixth Conference on Domain Decomposition Methods for Partial Differential Equations, A. Quarteroni, J. Périaux, Y. A. Kuznettsov, and O. Widlund, eds., Providence, RI, 1993, AMS.

[23] C. Johnson, U. Nävert, and J. Pitkäranta, *Finite element methods for linear hyperbolic problems*, Comp. Meth. Appl. Mech. Eng., 45 (1984), pp. 285–312.

[24] D. E. Keyes and W. D. Gropp, *Domain-decomposable preconditioners for second-order upwind discretizations of multicomponent systems*, in Fourth International Symposium on Domain Decomposition Methods for Partial Differential Equations, R. Glowinski, Y. A. Kuznetsov, G. Meurant, J. Périaux, and O. Widlund, eds., Philadelphia, PA, 1991, SIAM.

[25] P. L. Lions, *On the Schwarz alternating method. I.*, in First International Symposium on Domain Decomposition Methods for Partial Differential Equations, R. Glowinski, G. H. Golub, G. A. Meurant, and J. Périaux, eds., Philadelphia, PA, 1988, SIAM.

[26] J. W. H. Liu, *A graph partitioning algorithm by node separators*, ACM Transactions on Mathematical Software, 15 (1989), pp. 198–219.

[27] J. Mandel, *Hybrid domain decomposition with unstructured subdomains*, in Sixth Conference on Domain Decomposition Methods for Partial Differential Equations, A. Quarteroni, J. Périaux, Y. A. Kuznettsov, and O. Widlund, eds., Providence, RI, 1993, AMS.

[28] T. A. Manteuffel and S. V. Parter, *Preconditioning and boundary conditions*, SIAM J. Numer. Anal., 27 (1990), pp. 656–694.

[29] S. F. McCormick, *Multigrid Methods*, SIAM, 1987.

[30] J. A. Meijerink and H. A. V. der Vorst, *Guidelines for the usage of incomplete decompositions in the solving sets of linear equations as they occur in practical problems*, J. Comp. Phys., 44 (1981), pp. 134–155.

[31] S. V. Nepomnyaschikh, *Domain Decomposition and Schwarz Methods in a Subspace for the Approximate Solution of Elliptic Boundary Value Problems*, PhD thesis, Computing Center of the Siberian Branch of the USSR Academy of Sciences, Novosibirsk, USSR, 1986.

[32] S. V. Parter and S.-P. Wong, *Preconditioning second order elliptic operators: Condition numbers and the distribution of the singular values*, J. Sci. Comput., 6 (1991), pp. 129–157.

[33] A. Pothen, H. D. Simon, and K.-P. Liou, *Partitioning sparse matrices with eigenvectors of graphs*, SIAM J. Matrix Anal. Appl., 11 (1990), pp. 430–452.

[34] Y. Saad, *ILUT: a dual threshold incomplete ILU factorization*, Tech. Report 92-38, Minnesota Supercomputer Institute, University of Minnesota, Minneapolis, 1992.

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

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

[37] H. A. Schwarz, *Gesammelte Mathematische Abhandlungen*, vol. 2, Springer, Berlin, 1890, pp. 133–143. First published in Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich, volume 15, 1870, pp. 272–286.

[38] B. F. Smith, P. E. Bjørstad, and W. D. Gropp, *Domain Decomposition and Multilevel Methods for Elliptic PDEs: Algorithms, Implementations and Theory*, Cambridge University Press, 1995. To appear.

[39] J. Xu, *A new class of iterative methods for nonselfadjoint or indefinite problems*, SIAM J. Numer. Anal., 29 (1992), pp. 303–319.

[40] X. Zhang, *Studies in domain decomposition: multilevel methods and the biharmonic Dirichlet problem*, PhD thesis, Courant Institute, New York University, New York City, 1991.

[41] ———, *Multilevel Schwarz methods*, Numer. Math., 63 (1992), pp. 521–539.