Acta Numerica (1994), *pp.* 61–143

Domain decomposition algorithms

Tony F. Chan
Department of Mathematics,
University of California at Los Angeles,
Los Angeles, CA 90024, USA
Email: chan@math.ucla.edu.

Tarek P. Mathew
Department of Mathematics,
University of Wyoming,
Laramie, WY 82071-3036, USA
Email: mathew@corral.uwyo.edu.

Domain decomposition refers to divide and conquer techniques for solving partial differential equations by iteratively solving subproblems defined on smaller subdomains. The principal advantages include enhancement of parallelism and localized treatment of complex and irregular geometries, singularities and anomalous regions. Additionally, domain decomposition can sometimes reduce the computational complexity of the underlying solution method.

In this article, we survey *iterative* domain decomposition techniques that have been developed in recent years for solving several kinds of partial differential equations, including elliptic, parabolic, and differential systems such as the Stokes problem and mixed formulations of elliptic problems. We focus on describing the salient features of the algorithms and describe them using easy to understand matrix notation. In the case of elliptic problems, we also provide an introduction to the convergence theory, which requires some knowledge of finite element spaces and elementary functional analysis.

* The authors were supported in part by the National Science Foundation under grant ASC 92-01266, by the Army Research Office under contract DAAL03-91-G-0150 and subcontract under DAAL03-91-C-0047, and by the Office for Naval Research under contract ONR N00014-92-J-1890.

CONTENTS

1. Introduction

Domain decomposition (DD) methods are techniques for solving partial differential equations based on a decomposition of the spatial domain of the problem into several subdomains. Such reformulations are usually motivated by the need to create solvers which are easily parallelized on coarse grain parallel computers, though sometimes they can also reduce the complexity of solvers on sequential computers. These techniques can often be applied directly to the partial differential equations, but they are of most interest when applied to discretizations of the differential equations (either by finite difference, finite element, spectral or spectral element methods). The primary technique consists of solving subproblems on various subdomains, while enforcing suitable continuity requirements between adjacent subproblems, till the local solutions converge (within a specified accuracy) to the true solution.

In this article, we focus on describing *iterative* domain decomposition algorithms, particularly on the formulation of preconditioners for solution by conjugate gradient type methods. Though many fast *direct* domain decomposition solvers have been developed in the engineering literature, see Kron (1953) and Przemieniecki (1963) (these are often called *substructuring or tearing* methods), the more recent developments have been based on the iterative approach, which is potentially more efficient in both time and storage. The earliest known iterative domain decomposition technique was proposed in the pioneering work of H. A. Schwarz in 1870 to prove the existence of harmonic functions on irregular regions which are the union of overlapping subregions. Variants of Schwarz's method were later studied by Sobolev (1936), Morgenstern (1956) and Babuška (1957). See also Courant

and Hilbert (1962). The recent interest in domain decomposition was initiated in studies by Dinh, Glowinski and Périaux (1984), Dryja (1984), Golub and Mayers (1984), Bramble, Pasciak and Schatz (1986b), Bjørstad and Widlund (1986), Lions (1988), Agoshkov and Lebedev (1985) and Marchuk, Kuznetsov and Matsokin (1986), where the primary motivation was the inherent parallelism of these methods. There are not many general references that provide an overview of the field, but here are a few: discussions in Keyes and Gropp (1987), Canuto, Hussaini, Quarteroni and Zang (1988), Xu (1992a), Dryja and Widlund (1990), Hackbusch (1993), Le Tallec (1994) and the books of Lebedev (1986), Kang (1987) and Lu, Shih and Liem (1992) and the forthcoming book by Smith, Bjørstad and Gropp (1994). The best source of references remains the collection of conference proceedings: Glowinski, Golub, Meurant and Périaux (1988), Chan, Glowinski, Périaux and Widlund (1989, 1990), Glowinski, Kuznetsov, Meurant, Périaux and Widlund (1991), Chan, Keyes, Meurant, Scroggs and Voigt (1992a), Quarteroni (1993).

This article is conceptually organized in three parts. The first part (Sections 1 through 5) deals with second-order self-adjoint elliptic problems. The algorithms and theory are most mature for this class of problem and the topics here are treated in more depth than in the rest of the article. Most domain decomposition methods can be classified as either an overlapping or a nonoverlapping subdomain approach, which we shall discuss in Sections 2 and 3 respectively. A basic theoretical framework for studying the convergence rates will be summarized in Section 4. Some practical implementation issues will be discussed in Section 5. The second part (Sections 6–8) considers algorithms that are not, strictly speaking, domain decomposition methods, but that can be studied by the general framework set up in the first part. The key idea here is to extend the concept of the subdomains to that of subspaces. The topics include multilevel preconditioners (Section 6), locally refined grids (Section 7) and fictitious domain methods (Section 8). In the last part (Sections 9–12), we consider domain decomposition methods for more general problems, including convection–diffusion problems (Section 9), parabolic problems (Section 10), mixed finite element methods and the Stokes problems (Section 11). In Section 12, we provide references to algorithms for the biharmonic problem, spectral element methods, indefinite problems and nonconforming finite element methods. Due to space limitation, and the fact that both the theory and algorithms are generally less well developed for these problems, we do not treat Parts II and III in as much depth as in Part I. Our aim is instead to highlight some of the key ideas, using the framework and terminology developed in Part I, and to provide a guide to the vast developing literature.

We present the methods in algorithmic form, expressed in matrix notation, in the hope of making the article accessible to a broad spectrum of readers.

Given the space limitation, most of the theorems (especially those in Parts II and III) are stated without proofs, with pointers to the literature given instead. We also do not cover nonlinear problems or specific applications (e.g. CFD) of domain decomposition algorithms.

In the rest of this section, we introduce the main features of domain decomposition procedures by describing several algorithms based on the simpler case of *two subdomain* decomposition for solving the following general second-order self-adjoint, coercive elliptic problem:

$$Lu \equiv -\nabla \cdot (a(x,y)\nabla u) = f(x,y), \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \partial\Omega. \tag{1.1}$$

We are particularly interested in the solution of its discretization (by either finite elements or finite differences) which yields a large sparse symmetric positive definite linear system:

$$Au = f. \tag{1.2}$$

1.1. Overlapping subdomain approach

Overlapping domain decomposition algorithms are based on a decomposition of the domain Ω into a number of overlapping subregions. Here, we consider the case of two overlapping subregions $\{\hat{\Omega}_1, \hat{\Omega}_2\}$ which form a *covering* of Ω; see Figure 1. We shall let $\Gamma_i, i = 1, 2$ denote the part of the boundary of Ω_i which is in the interior of Ω.

The basic Schwarz alternating algorithm to solve (1.1) starts with any suitable initial guess u^0 and constructs a sequence of improved approximations $u^1, u^2, \ldots$. Starting with the kth iterate u^k, we solve the following two subproblems on $\hat{\Omega}_1$ and $\hat{\Omega}_2$ successively with the most current values as boundary condition on the artificial interior boundaries:

$$\begin{cases} Lu_1^{k+1} &= f, & \text{on } \hat{\Omega}_1, \\ u_1^{k+1} &= u^k|_{\Gamma_1} & \text{on } \Gamma_1, \\ u_1^{k+1} &= 0, & \text{on } \partial\hat{\Omega}_1 \backslash \Gamma_1, \end{cases}$$

and

$$\begin{cases} Lu_2^{k+1} &= f, & \text{on } \hat{\Omega}_2, \\ u_2^{k+1} &= u_1^{k+1}|_{\Gamma_2} & \text{on } \Gamma_2, \\ u_2^{k+1} &= 0, & \text{on } \partial\hat{\Omega}_2 \backslash \Gamma_2. \end{cases}$$

The iterate u^{k+1} is then defined by

$$u^{k+1}(x,y) = \begin{cases} u_2^{k+1}(x,y) & \text{if } (x,y) \in \hat{\Omega}_2 \\ u_1^{k+1}(x,y) & \text{if } (x,y) \in \Omega \backslash \hat{\Omega}_2. \end{cases}$$

It can be shown that in the norm induced by the operator L, the iterates $\{u^k\}$ converge geometrically to the true solution u on Ω, i.e.

$$\|u - u^k\| \leq \rho^k \|u - u^0\|,$$

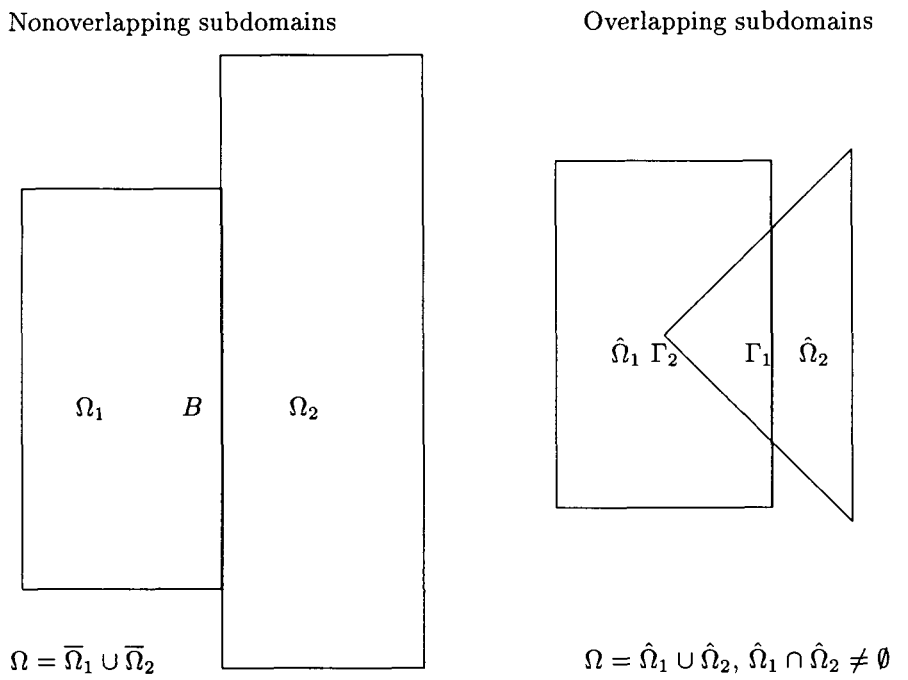

Fig. 1. Two subdomain decompositions.

where $\rho < 1$ depends on the choice of $\hat{\Omega}_1$ and $\hat{\Omega}_2$.

The above Schwarz procedure extends almost verbatim to discretizations of (1.1). We shall describe the discrete algorithm in matrix notation. Corresponding to the subregions $\{\hat{\Omega}_1, \hat{\Omega}_2\}$, let $\{\hat{I}_1, \hat{I}_2\}$ denote the indices of the nodes in the interior of domain $\hat{\Omega}_1$ and interior of $\hat{\Omega}_2$ respectively. Thus $\hat{I}_1$ and $\hat{I}_2$ form an overlapping set of indices for the unknown vector u. Let $\hat{n}_1$ be the number of indices in $\hat{I}_1$, and let $\hat{n}_2$ be the number of indices in $\hat{I}_2$. Due to overlap, $\hat{n}_1 + \hat{n}_2 > n$, where n is the number of unknowns in Ω.

Corresponding to each region $\hat{\Omega}_i$, we define a rectangular $n \times \hat{n}_i$ extension matrix R_i^T whose action extends by zero a vector of nodal values in $\hat{\Omega}_i$. Thus, given a subvector x_i of length $\hat{n}_i$ with nodal values at the interior nodes on $\hat{\Omega}_i$ we define:

$$(R_i^T x_i)_k = \begin{cases} (x_i)_k & \text{for } k \in \hat{I}_i \\ 0 & \text{for } k \in I - \hat{I}_i, \text{ where } I = \hat{I}_1 \cup \hat{I}_2. \end{cases}$$

The entries of the matrix R_i^T are ones or zeros. The transpose R_i of this extension map R_i^T is a restriction matrix whose action restricts a full vector x of length n to a vector of size $\hat{n}_i$ by choosing the entries with indices $\hat{I}_i$ corresponding to the interior nodes in $\hat{\Omega}_i$. Thus, $R_i x$ is the subvector

of nodal values of x in the interior of $\hat{\Omega}_i$. The local subdomain matrices (corresponding to the discretization on $\hat{\Omega}_i$) are, therefore,

$$A_1 = R_1 A R_1^T, \quad A_2 = R_2 A R_2^T,$$

and these are principal submatrices of A.

The discrete version of the Schwarz alternating method, described earlier, to solve $Au = f$, starts with any suitable initial guess u^0 and generates a sequence of iterates $u^0, u^1, \ldots$ as follows

$$u^{k+1/2} = u^k + R_1^T A_1^{-1} R_1 (f - Au^k), \tag{1.3}$$

$$u^{k+1} = u^{k+1/2} + R_2^T A_2^{-1} R_2 (f - Au^{k+1/2}). \tag{1.4}$$

Note that this corresponds to a generalization of the block Gauss–Seidel iteration (with overlapping blocks) for solving (1.1). At each iteration, two subdomain solvers are required (A_1^{-1} and A_2^{-1}). Defining

$$P_i \equiv R_i^T A_i^{-1} R_i A, \quad i = 1, 2,$$

the convergence is governed by the iteration matrix $(I - P_2)(I - P_1)$, hence this is often called a *multiplicative* Schwarz iteration. With sufficient overlap, it can be proved that the above algorithm converges with a rate independent of the mesh size h (unlike the classical block Gauss–Seidel iteration).

We note that P_1 and P_2 are symmetric with respect to the A inner product (see Section 4), but not so for the iteration matrix $(I - P_2)(I - P_1)$. A *symmetrized* version can be constructed by iterating one more half-step with A_1^{-1} after equation (1.4). The resulting iteration matrix becomes $(I-P_1)(I-P_2)(I - P_1)$ which is symmetric with respect to the A inner product and therefore conjugate gradient acceleration can be applied.

An analogous block Jacobi version can also be defined:

$$u^{k+1/2} = u^k + R_1^T A_1^{-1} R_1 (f - Au^k), \tag{1.5}$$

$$u^{k+1} = u^{k+1/2} + R_2^T A_2^{-1} R_2 (f - Au^k). \tag{1.6}$$

This version is more parallelizable because the two subdomain solves can be carried out concurrently. Note that by eliminating $u^{k+1/2}$, we obtain

$$u^{k+1} = u^k + (R_1^T A_1^{-1} R_1 + R_2^T A_2^{-1} R_2)(f - Au^k).$$

This is simply a Richardson iteration on $Au = f$ with the following *additive Schwarz preconditioner* for A:

$$M_{\text{as}}^{-1} = R_1^T A_1^{-1} R_1 + R_2^T A_2^{-1} R_2.$$

The preconditioned system can be written as

$$M_{\text{as}}^{-1} A = P_1 + P_2,$$

which is symmetric with respect to the A inner product and can also be used

with conjugate gradient acceleration. Again, for suitably chosen overlap (see Section 1), the condition number of the preconditioned system is bounded independently of h (unlike classical block Jacobi).

1.2. Nonoverlapping subdomain approach

Nonoverlapping domain decomposition algorithms are based on a partition of the domain Ω into various nonoverlapping subregions. Here, we consider a model partition of Ω into two nonoverlapping subregions Ω_1 and Ω_2, see Figure 1, with interface $B = \partial\Omega_1 \cap \partial\Omega$ (separating the two regions). Let $u = (u_1, u_2, u_B)$ denote the solution u restricted to Ω_1, Ω_2 and B respectively. Then, u_1, u_2 satisfy the following local problems:

$$\left\{\begin{array}{rcll} Lu_1 &=& f & \text{in } \Omega_1 \\ u_1 &=& 0 & \text{on } \partial\Omega_1 \backslash B \\ u_1 &=& u_B & \text{on } B \end{array}\right. \quad \text{and} \quad \left\{\begin{array}{rcll} Lu_2 &=& f & \text{in } \Omega_2 \\ u_2 &=& 0 & \text{on } \partial\Omega_2 \backslash B \\ u_2 &=& u_B & \text{on } B \end{array}\right. \tag{1.7}$$

as well as the following *transmission boundary condition* on the continuity of the flux across B:

$$\boldsymbol{n}_1 \cdot (a\nabla u_1) = -\boldsymbol{n}_2 \cdot (a\nabla u_2) \text{ on } B,$$

where each $\boldsymbol{n}_i$ is the outward pointing normal vector to B from Ω_i. (We omit derivation of the above, but note that it can be obtained by applying integration by parts to the weak form of the problem.) Thus, if the value u_B of the solution u on B is known, the local solutions u_1 and u_2 can be obtained at the cost of solving two subproblems on Ω_1 and Ω_2 in parallel.

The main task in nonoverlapping domain decomposition is to determine the interface data u_B. To this end, an equation satisfied by u_B can be obtained by using the transmission boundary conditions. Let g denote arbitrary Dirichlet boundary data on B. Define $E_1 g$ and $E_2 g$ as solutions of the following local problems, on Ω_1 and Ω_2 respectively:

$$\left\{\begin{array}{rcll} L(E_1 g) &=& f & \text{in } \Omega_1 \\ E_1 g &=& 0 & \text{on } \partial\Omega_1 \backslash B \\ E_1 g &=& g & \text{on } B \end{array}\right. \quad \text{and} \quad \left\{\begin{array}{rcll} L(E_2 g) &=& f & \text{in } \Omega_2 \\ E_2 g &=& 0 & \text{on } \partial\Omega_2 \backslash B \\ E_2 g &=& g & \text{on } B. \end{array}\right. \tag{1.8}$$

Then, by construction the boundary values of $E_1 g$ and $E_2 g$ match on B (and equal g). However, in general the *flux* of the two local solutions will not match on B, i.e.

$$\boldsymbol{n}_1 \cdot (a\nabla E_1 g) \neq -\boldsymbol{n}_2 \cdot (a\nabla E_2 g) \text{ on } B,$$

unless $g = u_B$. Define the following affine linear mapping T which maps the boundary data g on B to the jump in the flux across B:

$$T : g \longrightarrow \boldsymbol{n}_1 \cdot (a\nabla E_1 g) + \boldsymbol{n}_2 \cdot (a\nabla E_2 g)\,.$$

Thus, the boundary value u_B of the true solution u, satisfies the equation

$$Tu_B = 0. \tag{1.9}$$

The map T is referred to as a Steklov–Poincaré operator, and is a pseudo-differential operator (Agoshkov, 1988; Quarteroni and Valli, 1990). A property of the map T (or a linear map derived from T since it is *affine* linear) is that it is symmetric, and positive definite with respect to the L^2 inner product on B. The discrete versions of system (1.9) can therefore be solved by preconditioned conjugate gradient methods.

We now consider the corresponding algorithm for solving the linear system $Au = f$. Based on the partition $\Omega = \Omega_1 \cup \Omega_2 \cup B$, let $I = I_1 \cup I_2 \cup I_3$ denote a partition of the indices in the linear system, where I_1 and I_2 consists of the indices of nodes in the interior of Ω_1 and Ω_2, respectively, while I_3 consists of the nodes on the interface B. Correspondingly, the unknowns u can be partitioned as $u = [u_1, u_2, u_3]^T$ and $f = [f_1, f_2, f_3]^T$, and the linear system (1.2) takes the following block form:

$$\begin{bmatrix} A_{11} & 0 & A_{13} \\ 0 & A_{22} & A_{23} \\ A_{13}^T & A_{23}^T & A_{33} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix}. \tag{1.10}$$

Here, the blocks A_{12} and A_{21} are zero only under the assumption that the nodes in Ω_1 are not directly coupled to the nodes in Ω_2 (except through nodes on B), and this assumption holds true for finite element and low-order finite difference discretizations.

As in the continuous case, the problem $Au = f$ can be reduced to an equivalent system for the unknowns u_3 on the interface B. If u_3 is known, then u_1 and u_2 can be determined by using the first two block rows of (1.10):

$$u_1 = A_{11}^{-1}(f_1 - A_{13}u_3) \quad \text{and} \quad u_2 = A_{22}^{-1}(f_2 - A_{23}u_3).$$

Substituting for u_1 and u_2 in the third block row of (1.10), we obtain a reduced problem for the unknowns u_3:

$$Su_3 = \tilde{f}_3, \tag{1.11}$$

where $S \equiv \left(A_{33} - A_{13}^T A_{11}^{-1} A_{13} - A_{23}^T A_{22}^{-1} A_{23}\right)$ and $\tilde{f}_3 \equiv f_3 - A_{13}^T A_{11}^{-1} f_1 - A_{23}^T A_{22}^{-1} f_2$. The matrix S is referred to as the *Schur complement* of A_{33} in A, and the equation $Su_3 - \tilde{f}_3 = 0$ is a discrete approximation of the Steklov–Poincaré equation $Tu_B = 0$, enforcing the transmission boundary condition. The Schur complement S also plays a key role in the following block LU factorization of (1.10)

$$\begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ A_{13}^T A_{11}^{-1} & A_{23}^T A_{22}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & 0 & A_{13} \\ 0 & A_{22} & A_{23} \\ 0 & 0 & S \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix}, \tag{1.12}$$

from which (1.11) can also be derived.

Solving (1.11) by direct methods can be expensive since the Schur complement S is dense and, moreover, computing it requires as many solves of each A_{ii} system as there are nodes on B.

Therefore, it is common practice to solve the Schur complement system *iteratively* via preconditioned conjugate gradient methods. Each matrix–vector multiplication with S involves two subdomain solvers (A_{12}^{-1} and A_{22}^{-1}) which can be performed in parallel. It can be shown that the condition number of S is $\mathcal{O}(h^{-1})$ (which is better than that of A but can still be large) and therefore a good preconditioner is needed. Note that an advantage of the nonoverlapping approach over the overlapping approach is that the iterates are shorter vectors.

1.3. Main features of domain decomposition algorithms

The two preceding algorithms extend naturally to the case of many subdomains. However, a straightforward extension will not be *scalable*, i.e. the convergence rate will deteriorate as the number of subdomains increase. This is necessarily so because in the above algorithms, the only mechanism for sharing information is local, i.e. either through the interface or the overlapping regions. However, for elliptic problems the domain of dependence is global (i.e. the Green function is nonzero throughout the domain) and some way of transmitting *global* information is needed to make the algorithms scalable. One of the most commonly used mechanisms is to use *coarse spaces*, e.g. solving an appropriate problem on a coarser grid. This will be described in detail later.

In this sense, many of the domain decomposition algorithms can be viewed as a two-scale procedure, i.e. there is a fine grid with size h on which the solution is sought and on which the subdomain problems are solved, as well as a coarse grid with mesh size H which provides the global coupling between distant subdomains. The goal is to design the appropriate interaction of these two mechanisms so that the resulting algorithm has a convergence rate that is as insensitive to h and H as possible. In fact, in the literature on domain decomposition, a method is called *optimal* if its convergence rate is *independent* of h and H.

In practice, however, an *optimal* preconditioner does not necessarily provide the least execution time or minimal computational complexity. To achieve a computationally efficient algorithm requires paying attention to other factors, in addition to h and H. First of all, even though the number of iterations required by an optimal method can be bounded independent of h and H, one still has to ensure that it is not large. Second, each iteration step must not cost too much to implement. In addition, it would be desirable for the convergence rate to be insensitive to the variations in the coefficients

of the elliptic problem, as well as the aspect ratios of the subdomains. We shall touch on some of these issues later.

We summarize here the key features of domain decomposition algorithms that we have introduced in this section, and which we shall study in some detail in the rest of this article:

1 domain decomposition as preconditioners with conjugate gradient acceleration;
2 overlapping versus nonoverlapping subdomain algorithms;
3 nonoverlapping algorithms involve solving a Schur complement system, using interface preconditioners;
4 additive versus multiplicative algorithms;
5 optimal preconditioners require solving a coarse problem;
6 the goal of achieving a convergence rate and efficiency independent of h, H, coefficients and geometry.

Notation We use the notation cond $(M^{-1}A)$ to denote the condition number of the preconditioned system $M^{-1/2}AM^{-1/2}$, where M is symmetric and positive definite. We call a preconditioner M *spectrally equivalent* to A if cond $(M^{-1}A)$ is bounded independently of the mesh sizes h and H, whichever is appropriate.

2. Overlapping subdomain algorithms

We now describe Schwarz algorithms based on *many overlapping subregions* to solve (1.1). We first discuss a commonly used technique for constructing an overlapping decomposition of Ω into p subregions $\hat{\Omega}_1, \ldots, \hat{\Omega}_p$. To this end, let $\Omega_1, \ldots, \Omega_p$ denote a nonoverlapping partition of Ω. For instance, each subregion Ω_i may be chosen as *elements* from a coarse finite element triangulation τ^H of Ω of mesh size H. Next, we extend each nonoverlapping region Ω_i to $\hat{\Omega}_i$, consisting of all points in Ω within a distance of βH from Ω_i where β ranges from 0 to $\mathcal{O}(1)$. See Figure 2 for an illustration of a two-dimensional rectangular region Ω partitioned into sixteen overlapping subregions.

Once the extended subdomains $\hat{\Omega}_i$ are defined, we define *restriction* maps R_i, *extension* maps R_i^T, and local matrices A_i corresponding to each subregion $\hat{\Omega}_i$ as follows. Let A be $n \times n$ and let $\hat{n}_i$ be the number of interior nodes in $\hat{\Omega}_i$. For each $i = 1, \ldots, p$, let $\hat{I}_i$ denote the indices of the nodes lying in the interior of $\hat{\Omega}_i$. Thus $\{\hat{I}_1, \ldots, \hat{I}_p\}$ form an overlapping collection of index sets. For each region $\hat{\Omega}_i$ let R_i denote the $n \times \hat{n}_i$ restriction matrix (whose entries consist of 1's and 0's) that restricts a vector x of length n to $R_i x$ of length $\hat{n}_i$, by choosing the subvector having indices in $\hat{I}_i$ (corresponding to the interior nodes in $\hat{\Omega}_i$). The transpose R_i^T of R_i is referred to as an *extension* or *interpolation* matrix, and it extends subvectors of length $\hat{n}_i$ on

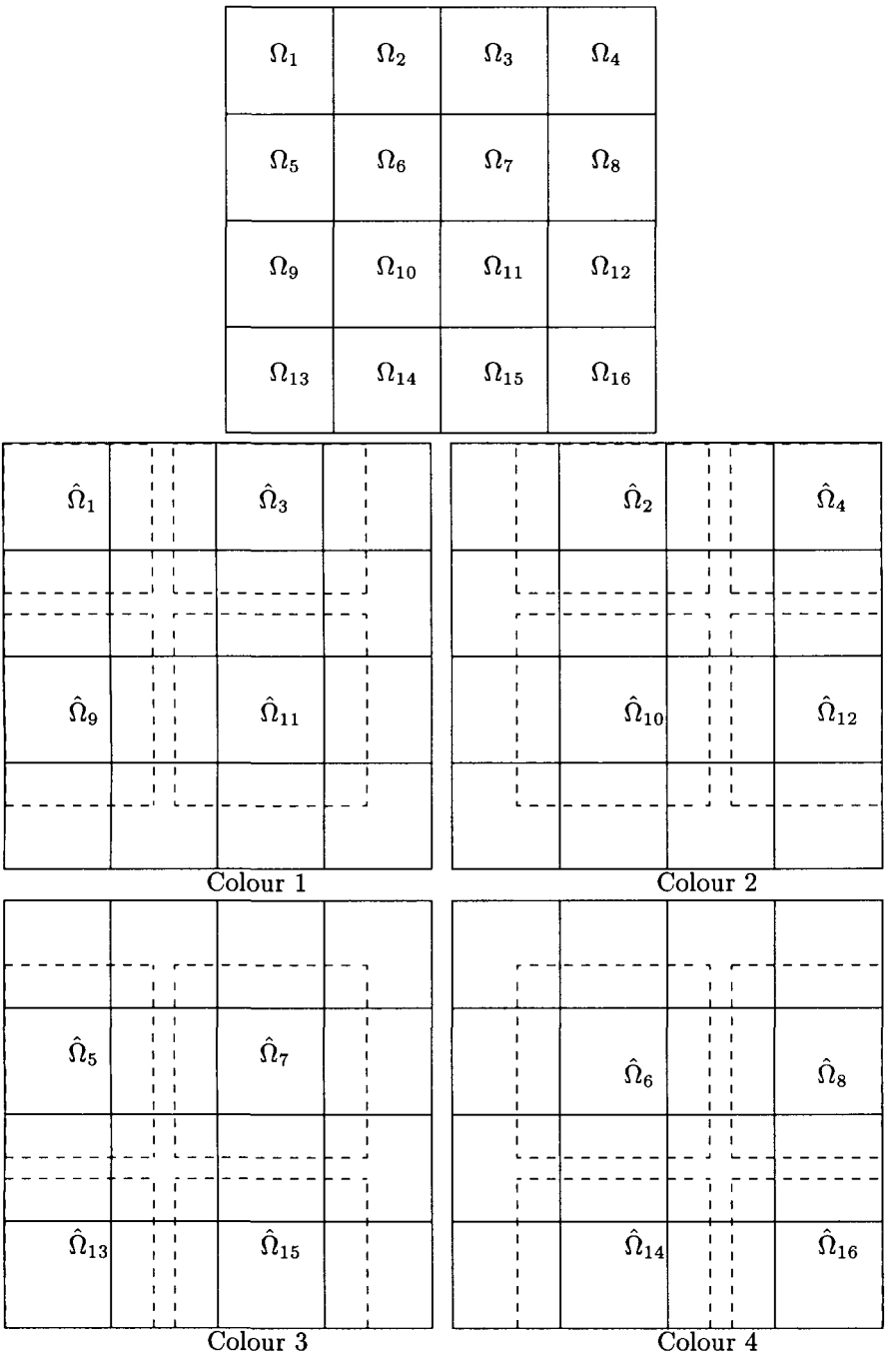

Fig. 2. Nonoverlapping subdomains Ω_i, overlapping subdomains $\hat{\Omega}_i$, 4 colours.

$\hat{\Omega}_i$ to vectors of length n using extension by zero to the rest of Ω. Finally, we let $A_i = R_i A R_i^T$, which is the local stiffness matrix corresponding to the subdomain $\hat{\Omega}_i$. Since R_i and R_i^T have entries of 1's and 0's, each A_i is a principal submatrix of A.

2.1. Additive Schwarz algorithms

The most straightforward generalization of the two subdomain additive Schwarz preconditioners described in Section 1 to the many subdomain case is the following:

$$M_{\text{as},1}^{-1} = \sum_{i=1}^{p} R_i^T A_i^{-1} R_i.$$

Since the action of each term $R_i^T A_i^{-1} R_i z$ can be computed on separate processors, this immediately leads to coarse grain parallelism. The actions of R_i^T and R_i are *scatter–gather* operations, respectively, and it is not necessary to store the extension and restriction matrices.

The preconditioner $M_{\text{as},1}$ is a straightforward generalization of the standard *block Jacobi* preconditioner to include overlapping blocks. However, the algorithm is not scalable because the convergence rate of this preconditioned iteration deteriorates as the number of subdomains p increases (i.e. as H decreases).

Theorem 1 There exists a positive constant C independent of H and h (but possibly dependent on the coefficients a) such that:

$$\text{cond}\,(M_{\text{as},1}^{-1} A) \leq C H^{-2} \left(1 + \beta^{-2}\right).$$

Proof. See Dryja and Widlund (1992a; 1989b). □

This deterioration in the convergence rate can be removed at a small cost by introducing a mechanism for global communication of information. There are several possible techniques for this, and here we will describe the most commonly used mechanism which is suitable only when the fine grid τ^h is a refinement of the coarse mesh τ^H. Accordingly, let R_H^T denote the standard interpolation map of coarse grid functions to fine grid functions (as in two-level *multigrid* methods). In the finite element context, R_H^T simply interpolates the nodal values from the coarse grid vertices to all the vertices on the fine grid, say by piecewise linear interpolation. Its transpose R_H is thus a weighted restriction map. If there are n_c coarse grid interior vertices, then R_H^T will be an $n \times n_c$ matrix. Indeed, if $\psi_1, \ldots, \psi_{n_c}$ are n_c column vectors representing the coarse grid nodal basis functions on the fine grid, then

$$R_H^T = [\psi_1, \ldots, \psi_{n_c}].$$

Corresponding to the coarse grid triangulation τ^H, let A_H denote the coarse grid discretization of the elliptic problem, i.e. $A_H = R_H A R_H^T$. Then, the improved *additive Schwarz* preconditioner $M_{\text{as},2}$ is defined by

$$M_{\text{as},2}^{-1} = R_H^T A_H^{-1} R_H + \sum_{i=1}^{p} R_i^T A_i^{-1} R_i = \sum_{i=0}^{p} R_i^T A_i^{-1} R_i, \tag{2.1}$$

where we have let $R_0 = R_H$ and $A_0 = A_H$. The convergence rate using this preconditioner is independent of H (for sufficient overlap).

Theorem 2 There exists a positive constant C independent of H, h (but possibly dependent on the variation in the coefficients a) such that

$$\text{cond}\,(M_{\text{as},2}^{-1} A) \leq C\left(1 + \beta^{-1}\right).$$

Proof. See Dryja and Widlund (1992a; 1989b), Dryja, Smith and Widlund (1993) and Theorems 14 and 16 in Section 4. □

2.2. Multiplicative Schwarz algorithms

The *multiplicative Schwarz* algorithm for many overlapping subregions can be analogously defined. Starting with an iterate u^k, we compute u^{k+1} as follows

$$u^{k+(i+1)/(p+1)} = u^{k+i/(p+1)} + R_i^T A_i^{-1} R_i (f - A u^{k+i/(p+1)}), \quad i = 0, 1, \ldots, p.$$

Theorem 3 The error $\|u - u^k\|$ in the kth iterate of the above multiplicative Schwarz algorithm satisfies

$$\|u - u^k\| \leq \rho^k \|u - u^0\|,$$

where $\rho < 1$ is independent of h and H, and depends only on β and the coefficients a, and $\|\cdot\|$ is the A-norm.

Proof. See Bramble, Pasciak, Wang and Xu (1991) and Theorems 15 and 16. □

As for the additive Schwarz algorithm, if the coarse grid correction is dropped, then the convergence rate of the multiplicative algorithm will deteriorate as $\mathcal{O}(H^{-2})$ when $H \to 0$.

The multiplicative algorithm as stated above has less parallelism than the additive version. However, this can be improved through the technique of multicolouring, as follows. Each subdomain is identified with a colour such that subdomains of the same colour are disjoint. The multiplicative Schwarz algorithm then iterates sequentially through the different colours, but now all the subdomain systems of the same colour can be solved in parallel. Typically, only a small number of colours is needed, see Figure 2 for an example. We caution that the convergence rate of the multicoloured

algorithm can depend on the ordering of the subdomains in the iteration and the increased parallelism may result in slower convergence (well known for the classical pointwise Gauss–Seidel method). However, this effect is less noticeable when a coarse grid solve is used.

The convergence bounds we have stated for both the additive and multiplicative Schwarz algorithms are valid in both two and three dimensions, but with possible *dependence* on the variation in the *coefficients* a. For large jumps in the coefficients, the convergence rate can deteriorate, but with maximum possible deterioration stated below.

Theorem 4 Assume that the coefficients a are constant (or mildly varying) within each coarse grid element. Then, for the additive Schwarz algorithm in two dimensions,

$$\text{cond}\,(M_{\text{as},2}^{-1}A) \leq C\,(1 + \log(H/h))\,,$$

and in three dimensions,

$$\text{cond}\,(M_{\text{as},2}^{-1}A) \leq C\,(H/h)\,,$$

where C is independent of the jumps in the coefficients and the mesh parameters H and h, but dependent on the overlap parameter β.

Proof. See Dryja and Widlund (1987) and Dryja *et al.* (1993). □

Corresponding results exist for the multiplicative Schwarz algorithms and the deterioration in the convergence rate can be improved by the use of alternative *coarse spaces*, see preceding reference.

For a numerical study of Schwarz methods, see Gropp and Smith (1992).

3. Nonoverlapping subdomain algorithms

As we saw in Section 2, there are two kinds of coupling mechanisms present in an optimal Schwarz type algorithm based on many overlapping subregions: local coupling between adjacent subdomains provided by the overlapped regions, and global coupling between distant subdomains provided by the coarse grid problem. In the case of nonoverlapping approach, the Schur complement system represents the coupling between the nodes on the interface B and in order to obtain optimal convergence rates, a coarse grid solve is still needed. However, since there is no overlap between neighbouring subdomains, the local coupling must be provided by some other mechanism. The most often used method is to use *interface preconditioners*, i.e. an effective approximation to the part of the Schur complement matrix S that corresponds to the unknowns on the interface separating two neighbouring subdomains. (In two dimensions, the interface is an edge and in three dimensions it is a face.) We shall first describe such interface preconditioners in Section 3.1 in the context of two subdomain decomposition (where it is

the only preconditioner needed). The case of many subregions is discussed in Section 3.2.

3.1. Two nonoverlapping subdomains: interface preconditioners

Consider the same setting as in Section 1, with Ω partitioned into two subdomains Ω_1 and Ω_2 separated by an interface B. We need a preconditioner M for the Schur complement $S \equiv A_{33} - A_{13}^T A_{11}^{-1} A_{13} - A_{23}^T A_{22}^{-1} A_{23}$.

(1) Exact eigen-decomposition of S: In some special cases, an exact eigen-decomposition of S can be derived from which the action of S^{-1} can be computed efficiently. For example, consider the five-point discretization of $-\Delta$ on a uniform grid of size h on the rectangular domain $\Omega = [0,1] \times [0, l_1 + l_2]$, which is partitioned into two subdomains $\Omega_1 = [0,1] \times [0, l_1]$ and $\Omega_2 = [0,1] \times [l_1, l_1 + l_2]$ with interface $B = \{(x,y) : y = l_1,\ 0 < x < 1\}$. We assume that the grid is $n \times (m_1 + 1 + m_2)$ with $l_i = (m_i + 1)h$, for $i = 1, 2$ and $h = 1/(n+1)$. It was shown by Bjørstad and Widlund (1986) and Chan (1987) that

$$S = F \Lambda F,$$

where F is the orthogonal sine transform matrix:

$$(F)_{ij} = \sqrt{\frac{2}{n+1}} \sin\left(\frac{ij\pi}{n+1}\right),$$

Λ is a diagonal matrix with elements given by

$$(\Lambda)_i = \left(\frac{1+\gamma_i^{m_1+1}}{1-\gamma_i^{m_1+1}} + \frac{1+\gamma_i^{m_2+1}}{1-\gamma_i^{m_2+1}}\right) \sqrt{\sigma_i + \sigma_i^2/4},$$

where

$$\sigma_i = 4 \sin^2\left(\frac{i\pi}{2(n+1)}\right) \quad \text{and} \quad \gamma_i = (1 + \sigma_i/2 - \sqrt{\sigma_i + \sigma_i^2/4})^2.$$

If m_1, m_2 are large enough, then two good approximations to S are:

$$M_{\mathrm{GM}} = F(\Sigma + \Sigma^2/4)^{1/2} F, \quad \text{and} \quad M_{\mathrm{D}} = F \Sigma^{1/2} F,$$

where $\Sigma = \mathrm{diag}(\sigma_i)$. M_{D} was first used by Dryja (1982) in a more general setting. The improved preconditioner M_{GM} was later proposed by Golub and Mayers (1984).

Note that all the above preconditioners can be solved in $\mathcal{O}(n \log(n))$ operations using the Fast Sine Transform and it is easy to show that they are *spectrally equivalent* to S. In theory, this is true for *any* second-order elliptic operator. However, these preconditioners can be sensitive to the *aspect ratios* l_1 and l_2 and the coefficients (in the case of variable coefficients) on the subdomains. To apply this class of preconditioners to domains more general

than a rectangle, and to provide some adaptivity to aspect ratios, Chan and Resasco (1985; 1987) suggested using the exact eigen-decomposition of a rectangle which approximates the given domain and shares the same interface. Exact eigen-decompositions have also been derived by Resasco (1990) for three-dimensional problems and unequal mesh sizes in each subdomain, and by Chan and Hou (1991) for five point stencils approximating general second-order constant coefficient elliptic problems (which provides some adaptivity to the coefficients).

(2) The Neumann–Dirichlet preconditioner (See Bjørstad and Widlund (1984), Bjørstad and Widlund (1986), Bramble *et al.* (1986b), Marini and Quarteroni (1989).) To describe this method, it is convenient to first write S in a form which reflects the contributions from Ω_1 and Ω_2 more explicitly. In either finite difference or finite element methods, the term A_{33} can be written as

$$A_{33} = A_{33}^{(1)} + A_{33}^{(2)},$$

where $A_{33}^{(i)}$ corresponds to the contribution to A_{33} from subdomain Ω_i (assuming the coefficients are zero on the adjacent subdomain). For instance, in the case of finite elements, $A_{33}^{(i)}$ is obtained by integrating the weak form on Ω_i. We can now write

$$S = S^{(1)} + S^{(2)},$$

where

$$S^{(i)} = A_{33}^{(i)} - A_{i3}^T A_{ii}^{-1} A_{i3}, \quad i = 1, 2.$$

Due to symmetry, $S^{(1)} = S^{(2)} = \frac{1}{2}S$ if the two subdomain problems are symmetric about the interface. This motivates the use of either $S^{(1)}$ or $S^{(2)}$ as a preconditioner for S even if the two subdomains are not equal. For example, a right-preconditioned system using $M_{\mathrm{ND}} = S^{(1)}$ has the form $(S^{(1)} + S^{(2)}){S^{(1)}}^{-1} = I + S^{(2)}{S^{(1)}}^{-1}$. It can be shown that the action of ${S^{(1)}}^{-1}$ on a vector v can be obtained by solving a problem on Ω_1 with v as Neumann boundary condition on the interface and extracting the solution values (Dirichlet values) on the interface:

$${S^{(1)}}^{-1} v = \begin{bmatrix} 0 & I \end{bmatrix} \begin{bmatrix} A_{11} & A_{13} \\ A_{13}^T & A_{33}^{(1)} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ v \end{bmatrix}.$$

It is proved in Bjørstad and Widlund (1986) that this preconditioner is spectrally equivalent to S.

(3) The Neumann–Neumann preconditioner One may notice a lack of symmetry in the Neumann–Dirichlet preconditioner in the choice of which subdomain to solve the Neumann problem on. The Neumann–Neumann

preconditioner, first proposed by Bourgat, Glowinski, Le Tallec and Vidrascu (1989), is completely symmetric with respect to the two subdomains. Here the inverse of the preconditioner is given by

$$M_{\mathrm{NN}}^{-1} = \tfrac{1}{4} {S^{(1)}}^{-1} + \tfrac{1}{4} {S^{(2)}}^{-1}.$$

Obviously, the action of M_{NN}^{-1} requires solving a Neumann problem on each of the two subdomains. In addition to the added symmetry, this preconditioner is also more directly generalizable to the case of many subdomains and to three dimensions (see Sections 3.6 and 3.9).

(4) Probing preconditioner This purely algebraic technique, first proposed by Chan and Resasco (1985) and later refined in Keyes and Gropp (1987) and Chan and Mathew (1992), is motivated by the observation that the entries of the rows (and columns) of the matrix S often decay rapidly away from the main diagonal. This decay is faster than the decay of the Green function of the original elliptic operator. The idea in the probing preconditioner is to efficiently compute a banded approximation to S. Note that this would be easy if S was known explicitly because we could then simply take the central diagonals of S. However, recall that we want to avoid computing S explicitly. The technique used in probing is to find such an approximation by *probing* the action of S on a few carefully selected vectors. For example, *if* S were tridiagonal, then it can be exactly recovered by its action on the three vectors:

$$\begin{aligned} v_1 &= (1,0,0,1,0,0,\ldots)^T, \\ v_2 &= (0,1,0,0,1,0,\ldots)^T, \\ v_3 &= (0,0,1,0,0,1,\ldots)^T \end{aligned}$$

through a simple recursion. Since S is not exactly tridiagonal, the tridiagonal matrix M_{P} obtained by probing will not be equal to S, but it is often a very good preconditioner. Keyes and Gropp (1987) showed that if S were symmetric, then two probing vectors suffice to compute a symmetric tridiagonal approximation. For more details, see Chan and Mathew (1992), where it is proved that the conditioner number of $M_{\mathrm{P}}^{-1}S$ can be bounded by $\mathcal{O}(h^{-1/2})$ (hence M_{P} is not spectrally equivalent to S) but it adapts very well to the aspect ratios and the coefficient variations of the subdomains. It would seem ideal to combine the advantages of the probing technique with a spectrally equivalent technique but this has proved to be elusive.

(5) Multilevel preconditioners These techniques make use of the multilevel elliptic preconditioners to be discussed in Section 6 and adapt them to obtain preconditioners for the Schur complement interface system. We will not describe these methods in detail, but the main idea is simple to understand. If a change of basis from the standard nodal basis to a hierarchical nodal basis is used (assuming that the grid has a hierarchical structure),

then a diagonal scaling often provides an effective preconditioner in the new basis. It can be shown rather easily that the Schur complement of the matrix A in the hierarchical basis is the same as that obtained by representing S with respect to the hierarchical basis on the interface B (i.e. by a multilevel change of basis restricted to the interface). Thus a good multilevel preconditioner for A automatically leads to a good multilevel preconditioner for S. The reader is referred to Smith and Widlund (1990) for using the hierarchical basis method of Yserentant (1986) and Tong, Chan and Kuo (1991) (see also Xu (1989)) for the multilevel nodal basis method of Bramble, Pasciak and Xu (1990). The resulting methods have optimal or almost optimal convergence rates.

3.2. Many nonoverlapping subdomains

Many of the preconditioners described in Section 3.1 for two nonoverlapping subdomains can be extended to the case of many nonoverlapping subregions. However, in the case of many subregions, these preconditioners need to be modified to take account of the more complex geometry of the interface, and to provide global coupling amongst the many subregions.

Let Ω be partitioned into p nonoverlapping regions of size $\mathcal{O}(H)$ with interface B separating them, see Figure 3:

$$\Omega = \Omega_1 \cup \cdots \cup \Omega_p \cup B, \quad \text{where } \Omega_i \cap \Omega_j = \emptyset \text{ for } i \neq j,$$

the interface B is given by: $B = \{\cup_{i=1}^{p} \partial\Omega_i\} \cap \Omega$. For $i = 1, \ldots, p$, let I_i denote the indices corresponding to the nodes in the interior of subdomain Ω_i, and let $I = \cup_{i=1}^{p} I_i$ denote the indices all nodes lying in the interior of subdomains. To minimize notation, we will use B to denote not only the interface, but also the indices of the nodes lying on B. Then, corresponding to the permuted indices $\{I, B\}$, the vector u can be partitioned as $u = [u_I, u_B]^T$, and $f = [f_I, f_B]^T$, and equation (1.2) can be written in block form as follows

$$\begin{bmatrix} A_{II} & A_{IB} \\ A_{IB}^T & A_{BB} \end{bmatrix} \begin{bmatrix} u_I \\ u_B \end{bmatrix} = \begin{bmatrix} f_I \\ f_B \end{bmatrix}. \tag{3.1}$$

For five-point stencils in two dimensions and seven-point stencils in three dimensions, A_{II} will be block diagonal, since the interior nodes in each subdomain will be decoupled from the interior nodes in other subdomains:

$$A_{II} = \text{blockdiag}\,(A_{ii}) = \begin{bmatrix} A_{11} & & 0 \\ & \ddots & \\ 0 & & A_{pp} \end{bmatrix}. \tag{3.2}$$

As in Section 1, the unknowns u_I can be eliminated resulting in a reduced system for u_B (the unknowns on B). We use the following block LU

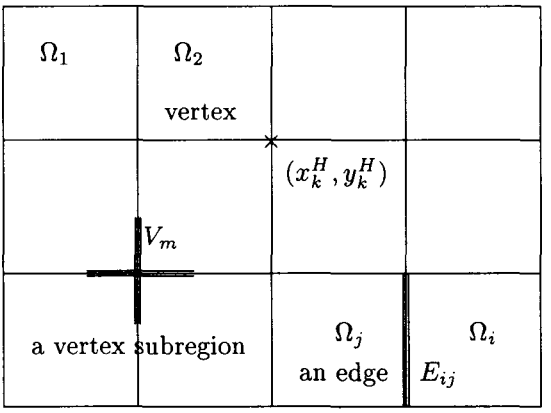

Fig. 3. A partition of Ω into 12 subdomains.

factorization of A:

$$A \equiv \begin{bmatrix} A_{II} & A_{IB} \\ A_{IB}^T & A_{BB} \end{bmatrix} = \begin{bmatrix} I & 0 \\ A_{IB}^T A_{II}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{II} & 0 \\ 0 & S \end{bmatrix} \begin{bmatrix} I & A_{II}^{-1} A_{IB} \\ 0 & I \end{bmatrix}, \tag{3.3}$$

where the Schur complement matrix S is defined by

$$S = A_{BB} - A_{IB}^T A_{II}^{-1} A_{IB}.$$

Consequently, solving $Au = f$ based on the LU factorization above requires computing the action of A_{II}^{-1} twice, and S^{-1} once.

By eliminating u_I, we obtain

$$Su_B = \tilde{f}_S, \tag{3.4}$$

where $\tilde{f}_B \equiv f_B - A_{IB} A_{II}^{-1} f_I$. The Schur complement S in the case of many subdomains has similar properties to the two subdomain case. Here we only note that the condition number of S is approximately $\mathcal{O}(H^{-1}h^{-1})$ in the case of many subdomains, an improvement over the $\mathcal{O}(h^{-2})$ growth for A. The rest of this section will be devoted to the description of various preconditioners M for S in two and three dimensions.

3.3. Two-dimensional case: block Jacobi preconditioner M_1 For S

Here, we describe a block diagonal preconditioner M_1 which reduces the condition number of S from $\mathcal{O}(H^{-1}h^{-1})$ to $\mathcal{O}\left(H^{-2}\log^2(H/h)\right)$ (without involving global communication of information). A variant of this preconditioner was proposed by Bramble, Pasciak and Schatz (1986a), see also Widlund (1988), Dryja *et al.* (1993).

The preconditioner M_1 will correspond to an additive Schwarz preconditioner for S corresponding to a partition of the interface B into subregions. The interface B is partitioned as a union of *edges* E_i for $i = 1, \ldots, m$, and *vertices* V of the subdomains, see Figure 3:

$$B = \{E_1 \cup \cdots \cup E_m\} \cup V,$$

where the edges $E_i = \partial\Omega_j \cap \partial\Omega_l$ form the common boundary of two subdomains (excluding the endpoints). With duplicity of notation, we also denote by E_i the indices of the nodes lying on edge E_i, and use V to denote the indices of the vertices V. Corresponding to this ordering of indices, we partition $u_B = [u_{E_1}, \ldots, u_{E_m}, u_V]$, and obtain a block partition of S:

$$S = \begin{bmatrix} S_{E_1E_1} & S_{E_1E_2} & \cdots & S_{E_1E_m} & S_{E_1V} \\ S_{E_1E_2}^T & S_{E_2E_2} & \cdots & S_{E_2E_m} & S_{E_2V} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ S_{E_1\ E_m}^T & S_{E_2E_m}^T & \cdots & S_{E_mE_m} & S_{E_mV} \\ S_{E_1V}^T & S_{E_2V}^T & \cdots & S_{E_mV}^T & S_{VV} \end{bmatrix}.$$

Note that $S_{E_iE_j} = 0$ if E_i and E_j are not part of the same subdomain.

A block diagonal (Jacobi) preconditioner for S is:

$$M_1 = \begin{bmatrix} S_{E_1E_1} & 0 & \cdots & \cdots & 0 \\ 0 & S_{E_2E_2} & \ddots & & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & S_{E_mE_m} & 0 \\ 0 & \cdots & \cdots & 0 & S_{VV} \end{bmatrix}.$$

The preconditioner M_1 can also be described in terms of restriction and extension maps. For each edge E_i, let R_{E_i} denote the pointwise restriction map from B onto the nodes on E_i, and let $R_{E_i}^T$ denote the corresponding extension map. Similarly, let R_V denote the pointwise restriction map onto the vertices V, and let R_V^T denote extension by zero of nodal values on V to B. Then the block Jacobi preconditioner is defined by

$$M_1^{-1} \equiv \sum_{i=1}^{m} R_{E_i}^T S_{E_iE_i}^{-1} R_{E_i} + R_V^T S_{VV}^{-1} R_V.$$

Since this preconditioner does not involve global coupling between subdomains, its convergence rate deteriorates as $H \to 0$.

Theorem 5 There exists a constant C independent of H and h (but may depend on the coefficient a), such that

$$\text{cond}\,(M_1^{-1}S) \leq CH^{-2}\left(1 + \log^2(H/h)\right).$$

Proof. See Bramble *et al.* (1986a), Widlund (1988), Dryja *et al.* (1993). □

Since the $S_{E_iE_i}$s are not explicitly constructed, computing the action of $S^{-1}_{E_iE_i}$ poses a problem (similarly for S_{VV}). Fortunately, each $S_{E_iE_i}$ and S_{VV} can be replaced by efficient approximations. For example, the block entries $S_{E_iE_i}$ can be replaced by any suitable two subdomain interface preconditioner $M_{E_iE_i}$ discussed in Section 3.1, for instance:

$$M_{E_iE_i} \equiv \alpha_{E_i} F \Lambda^{1/2} F,$$

where α_{E_i} represents the average of the coefficient a in the two subdomains adjacent to E_i. Alternatively, the action of $S^{-1}_{E_iE_i}$ can be computed exactly, using

$$S^{-1}_{E_iE_i} z_{E_i} = \begin{bmatrix} 0 & 0 & I \end{bmatrix} A^{-1}_{\Omega_j \cup \Omega_k \cup E_i} \begin{bmatrix} 0 & 0 & z_{E_i} \end{bmatrix}^T, \tag{3.5}$$

where $E_i = \partial\Omega_j \cap \partial\Omega_k$, and $A_{\Omega_j \cup \Omega_k \cup E_i}$ is the 3×3 block partitioned stiffness matrix corresponding to the region $\Omega_j \cup \Omega_k \cup E_i$. Note that this involves solving a problem on $\Omega_j \cup \Omega_k \cup E_i$. The matrix S_{VV} may be approximated by the diagonal matrix A_{VV} (the principal submatrix of A corresponding to nodes on V).

3.4. Two-dimensional case: the Bramble–Pasciak–Schatz (BPS) preconditioner M_2 for S

The H^{-2} factor in the condition number of the block Jacobi preconditioner M_1 can be removed by incorporating some mechanism for global coupling, such as through a coarse grid problem based on the coarse triangulation $\{\Omega_i\}$. Accordingly, let R_H^T denote an interpolation map (say piecewise linear interpolation) from the nodal values on V (vertices of subdomains) onto all the nodes on B. Then, R_H can be viewed as the weighted restriction map from B onto V. Note that the range of R_H^T here is B instead of the whole domain.

A variant M_2 of the preconditioner proposed by Bramble *et al.* (1986a) is a simple modification of M_1:

$$M_2^{-1} = \sum_{i=1}^{m} R_{E_i}^T S^{-1}_{E_iE_i} R_{E_i} + R_H^T A_H^{-1} R_H, \tag{3.6}$$

where A_H is the coarse grid discretization as in Section 2. With the global communication of information, the rate of convergence of the algorithm becomes logarithmic in H/h.

Theorem 6 There exists a constant C independent of H, h such that

$$\text{cond}\,(M_2^{-1} S) \leq C \left(1 + \log^2(H/h)\right).$$

In case the coefficients a are constant in each subdomain Ω_i, then C is also independent of a.

Proof. See Bramble *et al.* (1986a), Widlund (1988) and Dryja *et al.* (1993). □

As for the preconditioner M_1 to efficiently implement this algorithm, it is necessary to replace the subblocks $S_{E_iE_i}$ by suitable preconditioners, such as those described for the two subdomain case in Section 3.1, see also Chan, Mathew and Shao (1992b).

3.5. Two-dimensional case: vertex space preconditioner M_3 for S

The logarithmic growth $(1 + \log(H/h))^2$ in the condition number of the preceding preconditioner M_2 can be eliminated at additional cost, by modifying the BPS algorithm to result in the *vertex space preconditioner* proposed by Smith (1990, 1992).

The basic idea is to include additional overlap between the subblocks used in the BPS preconditioner M_2. Recall that the Schur complement S is not block diagonal in the permutation $[E_1, \ldots, E_m, V]$, since adjacent edges are coupled, with $S_{E_iE_j} \neq 0$ whenever edges E_i and E_j are part of the boundary of the same subdomain Ω_i. This coupling was ignored in the preceding two preconditioners, and resulted in the logarithmic growth factor in the condition number. By introducing overlapping subblocks, one can provide sufficient approximation of this coupling, resulting in *optimal* convergence bounds.

Overlap in the decomposition of interface

$$B = \{E_1 \cup \cdots \cup E_m\} \cup V,$$

can be obtained by introducing *vertex regions* $\{VS_1, \ldots, VS_q\}$ centred about each vertex in V (assume there are q subdomain vertices):

$$B \subset \{E_1 \cup \cdots \cup E_m\} \cup V \cup \{VS_1 \cup \cdots VS_q\}.$$

The vertex regions VS_k are illustrated in Figure 3, and are defined as the cross shaped regions centred at each subdomain vertex (x_k^H, y_k^H) containing segments of length βH of all the edges E_i that emanate from it. Such vertex spaces were used earlier by Nepomnyaschikh (1984; 1986).

Corresponding to this overlapping cover of B, we denote the indices of the nodes that lie on E_i by E_i, the indices of the vertices by V, and the indices of the vertex region VS_i by VS_i. Thus

$$E_1 \cup \cdots \cup E_m \cup V \cup VS_1 \cdots \cup VS_q$$

form an overlapping collection of indices of all unknowns on B. As with the restriction and extension maps for the BPS, we let R_{VS_i} denote the restriction of full vectors to subvectors corresponding to the indices in VS_i. Its

transpose $R_{VS_i}^T$ denotes the extension by zero of subvectors with indices VS_i to full vectors. The principal submatrix of S corresponding to the indices VS_i will be denoted $S_{VS_i} = R_{VS_i} S R_{VS_i}^T$. The *vertex space* preconditioner M_3 is an additive Schwarz preconditioner defined on this overlapping partition:

$$M_3^{-1} = \sum_{i=1}^{m} R_{E_i}^T S_{E_i E_i}^{-1} R_{E_i} + R_H^T A_H^{-1} R_H + \sum_{i=1}^{q} R_{VS_i}^T S_{VS_i}^{-1} R_{VS_i}. \tag{3.7}$$

In general, the matrices S_{VS_i} are dense and expensive to compute. However, sparse approximations can be computed efficiently using the probing technique or modifications of Dryja's interface preconditioner by Chan *et al.* (1992b). Alternately, using the following approximation:

$$S_{VS_i}^{-1} z_{VS_i} \approx \begin{bmatrix} 0 & I \end{bmatrix} \begin{bmatrix} A_{\Omega_{VS_i}} & A_{\Omega_{VS_i}, VS_i} \\ A_{\Omega_{VS_i}, VS_i}^T & A_{VS_i, VS_i} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ z_{VS_i} \end{bmatrix},$$

the action of $S_{VS_i}^{-1}$ can be approximated by solving a Dirichlet problem on a domain Ω_{VS_i} of diameter $2\beta H$ which contains VS_i and which is partitioned into a small number (four for rectangular regions) subregions by the interface VS_i.

The convergence rate of the vertex space preconditioned system is optimal in H and h (but may depend on variations in the coefficients).

Theorem 7 There exists a constant C_0 independent of H, h and β such that

$$\text{cond}\,(M_3^{-1} S) \leq C_0 (1 + \beta^{-1}),$$

where C_0 may depend on the variations in a. There also exists a constant C_1 independent of H, h, and the jumps in a (provided a is constant on each subdomain Ω_i) but can depend on β such that

$$\text{cond}\,(M_3^{-1} S) \leq C_1 (1 + \log(H/h)).$$

Proof. See Smith (1992), Dryja *et al.* (1993) and also Section 4. □

Thus, in the presence of large jumps in the coefficient a, the condition number bounds for the vertex space algorithm may deteriorate to $(1 + \log(H/h))$, which is the same growth as for the *BPS* preconditioner.

3.6. Two-dimensional case: Neumann–Neumann preconditioner M_4 for S

The Neumann–Neumann preconditioner for S in the case of many subdomains is a natural extension of the Neumann–Neumann algorithm for the case of two subregions, described in Section 3.1. This preconditioner was originally proposed by Bourgat *et al.* (1989), and extended by De Roeck (1989), De Roeck and Le Tallec (1991), Le Tallec, De Roeck and Vidrascu

(1991), Dryja and Widlund (1990; 1993a,b), Mandel (1992) and Mandel and Brezina (1992). There are several versions of the Neumann–Neumann algorithm, with the differences arising in the choice of a mechanism for global communication of information. We follow here a version due to Mandel and Brezina (1992), referred to as the *balancing domain decomposition* preconditioner.

Neumann–Neumann refers to the process of solving Neumann problems on each subdomain Ω_i during each preconditioning step. For each subdomain boundary $\partial\Omega_i$, let $R_{\partial\Omega_i}$ denote the pointwise restriction map (matrix) from nodes on B into nodes on $\partial\Omega_i \cap B$. Its transpose $R^T_{\partial\Omega_i}$ denotes an extension by zero of nodal values in $\partial\Omega_i \cap B$ to the rest of B. Corresponding to subdomain Ω_i, we denote the stiffness matrix of the Neumann problem by

$$A^{(i)} \equiv \begin{bmatrix} A^{(i)}_{II} & A^{(i)}_{IB} \\ A^{(i)^T}_{IB} & A^{(i)}_{BB} \end{bmatrix},$$

where $A^{(i)}_{II}$ is a principal submatrix of A corresponding to the nodes in the interior of Ω_i, $A^{(i)}_{IB}$ is a submatrix of A corresponding to the coupling between nodes in the interior of Ω_i and the nodes on the interface B restricted to $\partial\Omega_i$, and $A^{(i)}_{BB}$ corresponds to the coupling between the nodes on $\partial\Omega_i$ with contributions from Ω_i (in the finite element case, $A^{(i)}_{BB}$ is obtained by integrating the weak form on Ω_i for all the basis functions corresponding to the nodes on $\partial\Omega_i$).

For each subdomain Ω_i, we let $S^{(i)}$ denote the Schur complement with respect to the nodes on $\partial\Omega_i \cap B$ of the local stiffness matrix $A^{(i)}$:

$$S^{(i)} = A^{(i)}_{BB} - A^{(i)T}_{IB} A^{(i)^{-1}}_{II} A^{(i)}_{IB}. \tag{3.8}$$

The natural extension of the two subdomain Neumann–Neumann preconditioner is simply $\tilde{M}_4$:

$$\tilde{M}_4^{-1} = \sum_{i=1}^{p} R^T_{\partial\Omega_i} D_i \left(S^{(i)}\right)^{-1} D_i R_{\partial\Omega_i}, \tag{3.9}$$

where D_i is a diagonal weighting matrix. Note that $(S^{(i)})^{-1}v$ can be computed by a Neumann solve with v as Neumann data (see Section 3.1). This preconditioner is highly parallelizable, but it has two potential problems:

- The matrix $S^{(i)}$ is singular for interior subdomains since it corresponds to a Neumann problem on Ω_i. Accordingly, a compatibility condition must be satisfied, and additionally, the solution of the singular system will not be unique.
- There is no mechanism for global communication of information, and hence the condition number of the preconditioned system deteriorates at least as H^{-2}.

One way to rectify these two defects is the *balancing* procedure of Mandel and Brezina (1992). The residual is projected onto a subspace which automatically satisfies the compatibility conditions for each of the singular systems (as many as p constraints). Additionally, in a post processing step, a constant is added to the solution of each local singular system so that the residual remains in the appropriate subspace. This procedure also provides a mechanism for global communication of information. We omit the technical details, and refer the reader to Mandel and Brezina (1992). The singularity of the local Neumann problems also arises in a related method by Farhat and Roux (1992) where the interface compatibility conditions are enforced by a Lagrange multiplier approach.

The modified Neumann–Neumann preconditioner M_4 (with balancing) satisfies:

Theorem 8 There exists a constant C independent of H and h and the jumps in the coefficients a such that

$$\operatorname{cond}(M_4^{-1}S) \leq C\,(1+\log(H/h))^2\,.$$

Proof. See De Roeck and Le Tallec (1991), Mandel and Brezina (1992), Dryja and Widlund (1993a). □

The Neumann–Neumann preconditioner has several attractive features:

- the subregions Ω_i need not be triangular or rectangular; they can have general shapes;
- no explicit computation of the entries of S;
- the rate of convergence is logarithmic in H/h and insensitive to large jumps in the coefficients a.

However, the Neumann–Neumann preconditioner requires twice as many subdomain solves per step as a multiplication with S.

3.7. Three-dimensional case: vertex space preconditioner M_1 for S

Constructing effective preconditioners for the Schur complement matrix S is more complicated in three dimensions. These difficulties arise in part from the increased dimension of the boundaries of three-dimensional regions, and is also, technically, from a weaker Sobolev inequality in three dimensions.

As in the two-dimensional case, we assume that Ω is partitioned into p nonoverlapping subregions with interface B:

$$\Omega = \Omega_1 \cup \cdots \cup \Omega_p \cup B, \quad \text{where } B = (\cup_{i=1}^{p} \partial\Omega_i) \cap \Omega.$$

For most of the three-dimensional algorithms we will describe, it will be assumed that the $\{\Omega_i\}$ consist of either tetrahedrons or cubes and form a coarse triangulation of Ω having mesh size H. The boundary $\partial\Omega_i$ of

each tetrahedron or cube can be further partitioned into *faces, edges* and *vertices.* The *faces* F_{ij} = interior of $\partial\Omega_i \cap \partial\Omega_j$ are assumed to be open two-dimensional surfaces. The *edges* E_k are one-dimensional curves defined to be the intersection of the boundaries of two faces: $E_k = \partial F_{ij} \cap \partial F_{ln}$ excluding the endpoints. Finally, the *vertices* V are point sets which are the endpoints of edges.

As a prelude, we describe two preconditioners M_{1a} and M_{1b} related to the vertex space preconditioner M_1. Corresponding to the partition of B into faces, edges and subdomain vertices, we permute the unknowns on B as $x_B = [x_F, x_E, x_V]^T$, where F denote all the nodes on the faces, E corresponds to all the nodes on the edges E, while V denotes all the subdomain vertices. Thus, the matrix S has the following block form:

$$S = \begin{bmatrix} S_{FF} & S_{FE} & S_{FV} \\ S_{FE}^T & S_{EE} & S_{EV} \\ S_{FV}^T & S_{EV}^T & S_{VV} \end{bmatrix}.$$

The first preconditioner M_{1a} will be a block diagonal approximation of the above block partition of S, with the inclusion of a coarse grid model for global communication of information, see Dryja *et al.* (1993). Accordingly, for each of the subregions of B, let R_{F^i}, R_{E_k} and R_V denote the pointwise restriction map from B onto the nodes on face F_i, edge E_k and subdomain vertices V, respectively. Their transposes correspond to extensions by zero onto all other nodes on B. The principal submatrices of S corresponding to the nodes on F_i, E_k and V will be denoted by $S_{F_iF_i}$, $S_{E_kE_k}$ and S_{VV}, respectively. For the coarse grid problem, let R_H^T denote the interpolation map from the subdomain vertices V to all nodes on B. Then, its transpose R_H denotes a weighted restriction map onto the subdomain vertices V. The coarse grid matrix is then given by $A_H = R_H A R_H^T$.

In terms of the restriction and extension maps given above, M_{1a} is defined by

$$M_{1a}^{-1} = \sum_i R_{F_i}^T S_{F_iF_i}^{-1} R_{F_i} + \sum_k R_{E_k}^T S_{E_kE_k}^{-1} R_{E_k} + R_H^T A_H^{-1} R_H.$$

We note that the coupling terms $S_{F_iF_j}$ and $S_{E_iE_j}$ between adjacent faces and edges have been dropped. For finite element and finite difference discretizations, the blocks $S_{E_iE_i}$ can be shown to be well conditioned (indeed, for seven-point finite difference approximations on three-dimensional rectangular subdomains, $S_{E_iE_i} = A_{E_iE_i}$, since boundary data on the edges do not influence the solution in the interior of the region). Consequently, $S_{E_iE_i}$ may be effectively replaced by a suitably scaled multiple of the identity matrix $M_{E_iE_i}$:

$$S_{E_iE_i} \approx M_{E_iE_i} = h\sigma_{E_i} I_{E_i},$$

where σ_{E_i} represents the average of the coefficients a in the subdomains ad-

jacent to edge E_i. The action of $S_{F_iF_i}^{-1}$ can be approximated by analogues of the two-dimensional interface preconditioners from Section 3.1 or by solving a Dirichlet problem using a principal submatrix of A corresponding to nodes on a region Ω_{F_i} partitioned by face F_i.

A related preconditioner M_{1b} can be obtained at a small additional cost. For this, we note that the principal submatrix S_{VV} of S (corresponding to the nodes on the subdomain vertices V) can be replaced by a suitably scaled diagonal matrix M_{VV}:

$$S_{VV} \approx M_{VV} \equiv h \operatorname{diag}(\sigma_{V^k}),$$

where σ_{V^k} is the average of the coefficients a in the subdomains adjacent to vertex V_k. The preconditioner M_{1b} is defined by

$$M_{1b}^{-1} = \sum_i R_{F_i}^T S_{F_iF_i}^{-1} R_{F_i} + \sum_k R_{E_k}^T S_{E_kE_k}^{-1} R_{E_k} + R_H^T A_H^{-1} R_H + R_V^T M_{VV}^{-1} R_V.$$

The following are condition number bounds for the two preconditioners given above.

Theorem 9 The preconditioner M_{1a} results in condition number of

$$\operatorname{cond}\left(M_{1a}^{-1}S\right) \leq C_1 \frac{H}{h}\left(1+\log(H/h)\right)^2,$$

where C_1 is independent of H, h and jumps in the coefficients a. The preconditioner M_{1b} results in improved condition number with respect to mesh parameters:

$$\operatorname{cond}\left(M_{1b}^{-1}S\right) \leq C_2\left(1+\log(H/h)\right)^2,$$

where the coefficient C_2 may depend on the coefficients a.

Proof. See Dryja *et al.* (1993). □

We note that for smooth coefficients, M_{1b} is preferable to M_{1a} with improved condition number where the factor H/h has been eliminated.

The vertex space preconditioner of Smith (1992) in three dimensions corresponds to an *additive Schwarz* preconditioner for S, based on a suitable decomposition of the interface B into overlapping subregions and a coarse grid model. Accordingly, for each edge E_j, let $\hat{E}_j$ denote an extension consisting of all nodes on adjacent faces F_{ik} (but not adjacent edges or subdomain vertices) within a distance of βH from E^j. Similarly, corresponding to each subdomain vertex V_l, let $\hat{V}_l$ denote the vertex region consisting of all nodes in B within a distance of βH from vertex V_l. An overlapping partition of the interface B is then obtained:

$$B \subset (\cup_i F_i) \cup \left(\cup_k \hat{E}_k\right) \cup \left(\cup_l \hat{V}_l\right).$$

Corresponding to each overlapping subregion of the interface, define the

pointwise restriction and extension maps as follows. Let $R_{\hat{E}_k}$, $R_{\hat{V}_l}$ and R_{F_i} denote the pointwise restriction map from B onto the nodes on $\hat{E}_k$, $\hat{V}_l$ and F_i, respectively. Their transposes correspond to an extension by zero onto the rest of the nodes on B. Accordingly, let $S_{F_iF_i}$, $S_{\hat{E}_k\hat{E}_k}$ and $S_{\hat{V}_l\hat{V}_l}$ denote the principal submatrices of S corresponding to the nodes on F_i, $\hat{E}_k$ and $\hat{V}_l$ respectively. As for the preconditioners M_{1a} and M_{1b}, R_H^T and R_H will denote the coarse grid interpolation map and weighted restriction map, respectively. The coarse grid discretization matrix is obtained by $A_H = R_H A R_H^T$.

The vertex space preconditioner M_1 is defined by

$$M_1^{-1} = \sum_i R_{F_i}^T S_{F_iF_i}^{-1} R_{F_i} + \sum_k R_{\hat{E}_k}^T S_{\hat{E}_k}^{-1} R_{\hat{E}_k} + \sum_l R_{\hat{V}_l}^T S_{\hat{V}_l\hat{V}_l}^{-1} R_{\hat{V}_l} + R_H^T A_H^{-1} R_H. \tag{3.10}$$

As in the two-dimensional case, the action of the inverses $S_{F_iF_i}^{-1}$, $S_{\hat{V}_l\hat{V}_l}^{-1}$ and $S_{\hat{E}_k\hat{E}_k}^{-1}$ can be approximated without explicit construction of S. These approximations can be obtained by solving linear systems with principal submatrices of A as coefficient matrices, corresponding to subregions Ω_{F_i}, $\Omega_{\hat{E}_k}$ and $\Omega_{\hat{V}^l}$ containing F_i, $\hat{E}_k$ and $\hat{V}^l$ respectively, see Dryja *et al.* (1993), or by extensions of techniques in Chan *et al.* (1992b).

The rate of convergence of the *vertex space* preconditioner is independent of H and h, provided β is uniformly bounded. However, it may depend on the variation in the coefficients a.

Theorem 10 There exists a constant C, independent of H and h, but depending on the coefficients a such that

$$\text{cond}\,(M_1^{-1}S) \leq C(1 + \log^2(\beta^{-1})).$$

Proof. See Smith (1990) and Dryja and Widlund (1992b). □

3.8. Three-dimensional case: wirebasket preconditioners for S

Wirebasket algorithms were originally introduced in Bramble, Pasciak and Schatz (1989) (see also Dryja (1988)), and later modified and generalized by Smith (1991) and Dryja *et al.* (1993). These preconditioners for S involve computations on a *wirebasket* region W of B, and have almost optimal convergence rates with respect to mesh parameters and coefficients a (in case the coefficients are constant or mildly varying within each subdomain). The theoretical basis for the wirebasket method is an alternate coarse grid space based on a wirebasket region, which replaces the standard coarse grid problem. The interpolation map onto the wirebasket based coarse space has the favourable theoretical property that its bounds are independent of the variations in the coefficients and only mildly dependent on the mesh parameters (unlike the standard interpolation map onto the coarse grid).

We describe here a parallel wirebasket algorithm due to Smith (1991), see also Dryja *et al.* (1993).

The wirebasket preconditioners for S are based on a partition of the interface $B = F \cup W$ into *faces* F and a *wirebasket* W. As for the vertex space preconditioner described earlier, F will denote the collection of all the faces F_i. For each subdomain boundary $\partial\Omega_i$, define the ith wirebasket $W^{(i)}$ to consist of the union of all the edges and subdomain vertices lying on $\partial\Omega_i$:

$$W^{(i)} \equiv \bigcup_{E_k \subset \partial\Omega_i} E_k \bigcup_{V_j \subset \partial\Omega_i} V_j.$$

The *wirebasket* of B is defined to be the union of all the subdomain wirebaskets:

$$W \equiv \bigcup_{i=1}^{p} W^{(i)}.$$

Corresponding to the partition of the nodes $B = F \cup B$, the unknowns can be permuted: $x_B = [x_F, x_W]^T$, and the matrix S has the following block partition:

$$S = \begin{bmatrix} S_{FF} & S_{FW} \\ S_{FW}^T & S_{WW} \end{bmatrix}.$$

As for the vertex space algorithm, R_{F_i} will denote the pointwise restriction map onto nodes on F_i. Its transpose $R_{F_i}^T$ will denote extension by zero of nodal values on F_i to all the nodes on B. Next, corresponding to the wirebasket region W, there will be two kinds of restriction (and extension) maps, namely a pointwise restriction map R_W and a weighted restriction map $\hat{\mathcal{R}}_W$. For each i, the pointwise restriction map $R_{W^{(i)}}$ will restrict nodal values on B onto nodal values on the ith wirebasket $W^{(i)}$. Its transpose $R_{W^{(i)}}^T$ denotes the extension of nodal values on $W^{(i)}$ by zero to all nodes on B. Given a grid function u_W on W, the *wirebasket interpolation map* $\hat{\mathcal{R}}_W^T u_W$ extends the nodal values of u_W on W to the nodes on the faces as follows. On all the interior nodes on face F_i, the interpolant $\hat{\mathcal{R}}_W^T u_W$ is a constant equal to the average value of u_W on the boundary ∂F_i of face F_i:

$$\hat{\mathcal{R}}_W^T u_W = \begin{cases} u_W & \text{nodes } \in W \\ \text{average}(u_W)|_{\partial F_j} & \text{nodes } \in F_j. \end{cases}$$

Thus, its transpose $\hat{\mathcal{R}}_W$ will be a weighted restriction, mapping vectors u_B on B into vectors on W as follows:

$$\left(\hat{\mathcal{R}}_W u_B\right)_i = (u_B)_i + \sum_{k: i \in \partial F_k} \sum_{j \in F_k} \frac{(u_B)_j}{\dim(\partial F_k)}.$$

Next, let $z_{W^{(i)}}$ denote the vector whose entries are 1's for all indices on the ith wirebasket $W^{(i)}$. For $i = 1, \ldots, p$, define $\mathcal{B}^{(i)} = \rho_i (1 + \log(H/h))\, hI$

to be a diagonal matrix of the same size as the number of nodes on $W^{(i)}$, with $\rho_i = a|_{\Omega_i}$. Then, the matrix $\mathcal{B}$ is defined on the wirebasket W as a sum of the local matrices $\mathcal{B}^{(i)}$:

$$\mathcal{B} \equiv \sum_{i=1}^{p} R_{W^{(i)}}^T \mathcal{B}^{(i)} R_{W^{(i)}}.$$

Since $\mathcal{B}$ is the sum of several diagonal matrices, it will also be a diagonal matrix.

The wirebasket preconditioner M_2 of Smith (1991) has the following additive form:

$$M_2^{-1} = \sum_{i=1}^{m} R_{F_i}^T S_{F_i F_i}^{-1} R_{F_i} + \hat{\mathcal{R}}_W^T M_{WW}^{-1} \hat{\mathcal{R}}_W, \tag{3.11}$$

where the matrix M_{WW} is defined by its quadratic form:

$$u_W^T M_{WW} u_W = \sum_{i=1}^{p} \min_{\omega_i} \left(R_{W^{(i)}} u_W - \omega_i z_{W^{(i)}}\right)^T \mathcal{B}^{(i)} \left(R_{W^{(i)}} u_W - \omega_i z_{W^{(i)}}\right).$$

The terms $\omega_i z_{W^{(i)}}$ and the minimization are there to ensure that the local Schur complement $S^{(i)}$ and $M_2^{(i)}$ have the same null space spanned by $z_{W^{(i)}}$ (which in the case of scalar problems is $[1, \ldots, 1]^T$, but for systems such as elasticity, there may be several linearly independent null vectors).

The ease of inversion of M_{WW} is of course crucial to the efficiency of the preconditioner M_2. The linear system

$$M_{WW} x_W = f_W,$$

is equivalent, due to positive definiteness, to the following minimization problem:

$$\min_{x_W} \tfrac{1}{2} x_W^T M_{WW} x_W - x_W^T f_W,$$

and by substituting the quadratic form for M_{WW}, we obtain

$$\min_{x_W} \frac{1}{2} \sum_{i=1}^{p} \min_{\omega_i} \left(R_{W^{(i)}} x_W - \omega_i z_{W^{(i)}}\right)^T \mathcal{B}^{(i)} \left(R_{W^{(i)}} x_W - \omega_i z_{W^{(i)}}\right) - x_W^T f_W.$$

Differentiating with respect to all unknowns in x_W and with respect to $\omega_1, \ldots, \omega_p$, the following equivalent linear system is obtained:

$$\begin{cases} z_{W^{(i)}}^T \mathcal{B}^{(i)} \left(R_{W^{(i)}} x_W - \omega_i z_{W^{(i)}}\right) &= 0 \text{ for } i = 1, \ldots, p, \\ \mathcal{B} x_W - \sum_{i=1}^{p} \omega_i R_{W^{(i)}}^T \mathcal{B}^{(i)} z_{W^{(i)}} &= f_W. \end{cases}$$

If $\omega_1, \ldots, \omega_p$ are known, then x_W can be determined by solving the second block row (which is a diagonal system):

$$x_W = \mathcal{B}^{-1} \left(f_W + \sum_{i=1}^{p} \omega_i R_{W^{(i)}}^T \mathcal{B}^{(i)} z_{W^{(i)}} \right).$$

Substituting this into the first block row, we obtain

$$\left(z_{W^{(i)}}^T \mathcal{B}^{(i)} z_{W^{(i)}}\right) \omega_i - z_{W^{(i)}}^T \mathcal{B}^{(i)} R_{W^{(i)}} \mathcal{B}^{-1} \sum_{j=1}^{p} \omega_j R_W^T(j) \mathcal{B}^{(j)} z_{W^{(j)}} = z_W^T(i) \mathcal{B}^{(i)} R_{W^{(i)}} \mathcal{B}^{-1} f_W.$$

Note that this $p \times p$ coefficient matrix for $\omega_1, \ldots, \omega_p$ can be computed, and it can be verified that it will be sparse. The resulting system for $\omega_1, \ldots, \omega_p$ can be solved using any sparse direct solver.

The convergence rate of this additive wirebasket algorithm of Smith (1991) is logarithmic in the number of unknowns per subdomain.

Theorem 11 If the coefficients a are mildly varying within each subdomain, there exists a constant C independent of H, h and a such that

$$\text{cond}\,(M_2^{-1} S) \leq C(1 + \log(H/h))^2.$$

Proof. See Smith (1991), Dryja *et al.* (1993). □

For alternate wirebasket algorithms, we refer the reader to Bramble *et al.* (1989), Mandel (1989a), Dryja *et al.* (1993). The latter contains a wirebasket algorithm with condition number bounded by $1 + \log(H/h)$.

3.9. Three dimensions: Neumann–Neumann preconditioner M_3 for S

The Neumann–Neumann preconditioner for S in three dimensions is identical in form to the two-dimensional Neumann–Neumann preconditioner described earlier, and so the algorithm will not be repeated here. We mention here that an attractive feature of the Neumann–Neumann algorithm in three dimensions is that it does not require distinction between various subregions of the boundary $\partial\Omega_i$ of each subdomain (such as faces, edges, vertices and wirebaskets). Additionally, the almost optimal convergence rates are also valid for three-dimensional problems, see De Roeck and Le Tallec (1991), Dryja and Widlund (1990; 1993a), Mandel and Brezina (1992).

4. Introduction to the convergence theory

In this section, we provide a brief introduction to a theoretical framework for studying the convergence rates of the Schwarz (overlapping) and Schur complement (nonoverlapping) based domain decomposition methods discussed in this article (the Schwarz framework can also be used for analysing multilevel methods). Since the convergence rates of preconditioned conjugate gradient methods depend on the quotient of the extreme eigenvalues of the preconditioned matrix $M^{-1}A$ (which is assumed to be symmetric, positive definite in a suitable inner product), this theoretical framework involves techniques for estimating and bounding the extreme eigenvalues of the resulting preconditioned matrices. Additionally, in case of unaccelerated

iterations based on matrix splittings, the framework provides a technique for estimating the spectral radius or norm of the error propagation matrix.

A prominent feature of the Schwarz algorithms that simplifies their convergence analysis is that the preconditioned matrices (or the error propagation matrices in case of unaccelerated iterations) can be expressed as sums (or products) of orthogonal projection matrices. The abstract framework described here, originated and evolved from convergence studies of the classical Schwarz alternating algorithm in a variational framework, see Lions (1988), Sobolev (1936), Babuška (1957) and Morgenstern (1956), with extensions and applications in the finite element context by Widlund (1988), Dryja and Widlund (1987; 1989b; 1990; 1993a), Matsokin and Nepomnyaschikh (1985), Nepomnyaschikh (1986), Bramble *et al.* (1991), Xu (1992a), and others. Nonvariational theories, in particular ones based on the maximum principle, have also been used to study domain decomposition methods, Miller (1965), Tang (1988), Lions (1989), Chan, Hou and Lions (1991a).

4.1. Abstract framework for additive and multiplicative Schwarz algorithms

Recall that the preconditioned system $M^{-1}A$ of the additive Schwarz preconditioner M is defined by

$$M^{-1}A = \sum_{i=0}^{p} R_i^T A_i^{-1} R_i A = \sum_{i=0}^{p} P_i,$$

where $P_i \equiv R_i^T A_i^{-1} R_i A$. (We have, for convenience, denoted the coarse grid problem $R_H^T A_H^{-1} R_H$ by $R_0^T A_0^{-1} R_0$.) When A is symmetric positive definite, the matrices P_i are orthogonal projection matrices in the A inner product, since

$$P_i P_i = R_i^T \left(A_i^{-1} R_i A R_i^T \right) A_i^{-1} R_i A = R_i^T A_i^{-1} R_i A = P_i,$$

and

$$AP_i = AR_i^T A_i^{-1} R_i A = P_i^T A.$$

Thus, the extreme eigenvalues of $M^{-1}A$ can be estimated by finding upper and lower bounds for the spectra of the sums of the orthogonal projections P_i. We describe the abstract framework for doing this in the following.

Let V be a Hilbert space with inner product $a(.\,,.)$ and let $V_0, \ldots, V_p$ be subspaces $V_i \subset V$. (In the matrix case, $a(u,v) \equiv u^T Av$.) For $i = 0, \ldots, p$, let P_i denote the orthogonal projection from V into V_i, i.e.

$$P_i u \in V_i \text{ satisfies } \; a(P_i u, v) = a(u, v) \; \forall v \in V_i.$$

Let N_c denote the minimum number of distinct colours so that the spaces $V_1, \ldots, V_p$ of the same colour are mutually orthogonal in the $a(.\,,.)$ inner product (note that the subspaces corresponding to disjoint subdomains will

be mutually orthogonal, for domain decomposition algorithms). Then the following upper bound holds for the spectra of the additive operator $P_0 + \cdots + P_p$.

Theorem 12 $\lambda_{\max}(P_0 + \cdots + P_p) \leq N_c + 1$.

Proof. Recall that the spectral radius of any matrix A satisfies $\rho(A) \leq \|A\|$, and for orthogonal matrices the norm $\|P_i\| \leq 1$. Thus, an upper bound of $p+1$ is trivially obtained since the norm of each projection P_i is bounded by 1, and the sum of $p+1$ such projections gives a bound of $p+1$. The improved upper bound of $N_c + 1$ is obtained by noting that the sum of projections of the same colour, equals a projection onto the sum of the subspaces of the same colour. Consequently, there are only N_c projections for the colours, and projection P_0 onto the coarse grid. The result thus follows. □

A lower bound for a sum of the projections can be obtained, provided the spaces V_i satisfy the following property with constant C_0 that can be estimated.

Partition property of V_i *For any $u \in V$, there exists a constant $C_0 \geq 1$, such that the partition: $u = u_0 + \cdots + u_p$, where $u_i \in V_i$, satisfies*

$$\sum_{i=0}^{p} a(u_i, u_i) \leq C_0 a(u, u).$$

The lower bound for the sum of the projections can be estimated based on C_0, in a result described in Lions (1988), see also Dryja and Widlund (1987; 1989b). Similar ideas were developed earlier by Matsokin and Nepomnyaschikh (1985).

Theorem 13 Suppose the subspaces V_i for $i = 0, \ldots, p$, satisfy the partition property with constant $C_0 \geq 1$. Then,

$$\lambda_{\min}(P_0 + \cdots + P_p) \geq 1/C_0.$$

Proof. We shall use the Rayleigh quotient characterization:

$$\lambda_{\min}(P_0 + \cdots + P_p) = \min_{u \neq 0} \sum_{i=0}^{p} a(P_i u, u)/a(u, u).$$

For arbitrary $u \in V$, consider

$$a(u, u) = \sum_{i=0}^{p} a(u_i, u), \qquad \text{where } u = u_0 + \cdots + u_p.$$

Since P_i are projections, we obtain that $a(u_i, u) = a(u_i, P_i u)$. Now, applying

the Schwarz inequality, we obtain

$$\sum_{i=0}^{p} a(u_i, u) = \sum_{i=0}^{p} a(u_i, P_i u) \leq \left(\sum_{i=0}^{p} a(u_i, u_i) \right)^{1/2} \left(\sum_{i=0}^{p} a(P_i u, P_i u) \right)^{1/2}.$$

By the partition property, we obtain that

$$a(u, u) \leq C_0^{1/2} a(u, u)^{1/2} \left(\sum_{i=0}^{p} a(P_i u, P_i u) \right)^{1/2}.$$

After cancellation this becomes

$$a(u, u)^{1/2} \leq C_0^{1/2} \left(\sum_{i=0}^{p} a(P_i u, P_i u) \right)^{1/2} = C_0^{1/2} \left(\sum_{i=0}^{p} a(P_i u, u) \right)^{1/2},$$

where the last equality follows since $a(P_i u, P_i u) = a(P_i u, u)$. Squaring both sides, the result gives a lower bound for the Rayleigh quotient. □

Combining the upper and lower bounds, we obtain:

Theorem 14 The condition number cond $(M^{-1}A)$ of the additive Schwarz preconditioned system is bounded by $(N_c + 1)C_0$.

Next, we estimate the convergence rate of the unaccelerated multiplicative Schwarz method. Analogous to the two subdomain case presented in Section 1, it can be easily derived that the error $e^n = u - u^n$ satisfies

$$e^{n+1} = (I - P_p) \cdots (I - P_0) e^n.$$

Thus:

$$\|e^n\| \leq \|(I - P_p) \cdots (I - P_0)\| \|e^n\|.$$

Clearly, $\|(I - P_p) \cdots (I - P_0)\| \leq 1$ in the norm generated by bilinear form $a(.\,, .)$, since the $(I - P_i)$ are also orthogonal projections with norms bounded by 1. Moreover, it is strictly less than 1 whenever $V = V_0 + \cdots + V_p$. More precisely, we have:

Theorem 15 Let V_i satisfy the partition property with constant C_0. Then the error propagation map of the multiplicative Schwarz iteration satisfies

$$\|(I - P_p) \cdots (I - P_0)\| \leq 1 - c/C_0 < 1,$$

where c is a constant that depends only on N_c but independent of p.

Proof. See Bramble *et al.* (1991). A precise expression for $0 < c < C_0$ is also given in Xu (1992a), Wang (1993), Cai and Widlund (1993). □

For the Schwarz algorithms based on the subdomains illustrated in Figure 2, the number of colours is $N_c = 4$. Analogous subdomain partitions in

three dimensions yield $N_c = 8$. More generally, for most domain decomposition algorithms, N_c is a fixed number, independent of the number of subdomains. (However, for multilevel methods, N_c equals the number of levels, and then the colouring assumption must be replaced by a weaker assumption, see Bramble, Pasciak, Wang and Xu (1991), Xu (1992a), Yserentant (1986) and Griebel and Oswald (1993).) Thus, the rate of convergence depends critically on the *partition constant* C_0 and this will be estimated for finite element spaces in the next section.

4.2. A partition lemma for finite element spaces

In this section, following Dryja and Widlund (1987; 1992b) and Bramble *et al.* (1991), we describe a technique for estimating the partition constant C_0 for the basic overlapping Schwarz algorithms of Section 2.

Let $V^h(\Omega)$ denote the space of finite element functions defined on a quasi-uniform triangulation $\tau^h(\Omega)$, and let $V_i \equiv V^h(\Omega_i) \cap H_0^1(\hat{\Omega}_i)$ denote the finite element functions in $V^h(\Omega)$ which vanish outside $\hat{\Omega}_i$. Additionally, let $V_0 = V^H(\Omega)$ denote the space of finite element functions based on the coarse triangulation $\tau^H(\Omega)$ consisting of nonoverlapping elements $\Omega_1, \ldots, \Omega_p$.

We then have the following partition lemma.

Theorem 16 Let $a(.\,,.)$ denote the bilinear form associated with the elliptic problem in R^d for $d \leq 3$. The subspaces V_i defined above satisfy that for any $u \in V^h(\Omega)$, there exists $u_i \in V_i$ with

$$u = \sum_{i=0}^{p} \text{ and } \sum_{i=0}^{p} a(u_i, u_i) \leq C\left(1 + \beta^{-2}\right) a(u, u), \tag{4.1}$$

where C is a constant independent of H and h, but which depends on the coefficients.

Proof. We outline the proof only for the case of continuous piecewise linear finite element functions. Let $u_0 = Q_0 u_h$, where Q_0 is the L^2 orthogonal projection onto V_0. Then, by the H^1 stability of the L^2 projection, see Xu (1989) and Bramble and Xu (1991), we have

$$|u_0|_{H^1(\Omega)} \leq C|u_h|_{H^1(\Omega)}, \tag{4.2}$$

for some constant C independent of H and h. By using the equivalence between $a(.\,,.)$ and the H^1 norm, it follows from (4.2) that

$$a(u_0, u_0) \leq C a(u_h, u_h). \tag{4.3}$$

Let I_H denote the finite element interpolation map onto the coarse space $V^H(\Omega)$. By using the best approximation property of Q_0 and applying the standard finite element interpolation error bound for $(u_h - I_H u_h)$ we obtain

$$\|u_h - u_0\|_{L^2(\Omega)} \leq \|u_h - I_H u_h\|_{L^2(\Omega)} \leq CH|u_h|_{H^1(\Omega)}. \tag{4.4}$$

Next, let $\chi_1, \ldots, \chi_p$ be a partition of unity, subordinate to the covering $\hat{\Omega}_1, \ldots, \hat{\Omega}_p$, satisfying:

$$0 \leq \chi_i \leq 1, \ \chi_i \in C_0^\infty\left(\hat{\Omega}_i\right), \text{ with } \sum_{i=1}^{p} \chi_i = 1, \text{ and } |\nabla \chi_i|_\infty \leq C\beta^{-1}H^{-1}.$$

Note that such a partition of unity exists due to the overlapping cover. We then define the following partition of $u_h - u_0$:

$$u_i = I_h\left(\chi_i\left(u_h - u_0\right)\right), \text{ for } i = 1, \ldots, p, \tag{4.5}$$

where I_h is the finite element interpolation onto $V^h(\Omega)$. We note that without the interpolation, the terms $\chi_i\left(u_h - u_0\right)$ will not be in the finite element space, since the product with χ_i is not piecewise polynomial. By linearity of the interpolant I_h, and the partition of unity, it follows that

$$u_1 + \cdots + u_p = u_h - u_0.$$

We now estimate the partition constant C_0 in several steps. To simplify the notation, C will denote a generic constant below. For each element $e \in \tau^h$, let $0 \leq \theta_e \leq 1$ be a constant such that $\|\chi_i - \theta_e\|_{L^\infty(e)} = \mathcal{O}(h/H)$ (e.g. $\theta_e = \chi_i(x_0)$ where x_0 is the centre of the element). Then, in element e we have

$$\begin{aligned} u_i &\equiv I_h\left(\chi_i(u_h - u_0)\right) \\ &= I_h\left((\chi_i - \theta_e)(u_h - u_0)\right) + I_h\left(\theta_e(u_h - u_0)\right) \\ &= I_h\left((\chi_i - \theta_e)(u_h - u_0)\right) + \theta_e\left(u_h - u_0\right), \end{aligned}$$

since θ_e is constant in element e.

By applying the triangle inequality to the gradient of the above expression, and using that $\theta_e \leq 1$, we obtain

$$|u_i|_{H^1(e)} \equiv \|\nabla u_i\|^2_{L^2(e)} \leq 2\|\nabla I_h(\chi_i - \theta_e)(u_h - u_0)\|^2_{L^2(e)} + 2\|\nabla(u_h - u_0)\|^2_{L^2(e)}.$$

By applying an inverse inequality (which states that $|v_h|_{H^1} \leq Ch^{-1}\|v_h\|_{L^2}$ for any finite element function v_h), and the fact that $\|I_h\left(fv_h\right)\|_{L^2(e)} \leq \|f\|_{L^\infty(e)}\|v_h\|_{L^2(e)}$ for any continuous function f, the first term on the right-hand side can be bounded by

$$Ch^{-2}\|I_h(\chi_i - \theta_e)(u_h - u_0)\|^2_{L^2(e)} \leq Ch^{-2}\|\chi_i - \theta_e\|^2_{L^\infty(e)}\|I_h(u_h - u_0)\|^2_{L^2(e)}.$$

Since $\|\chi_i - \theta_e\|_{L^\infty(e)} = \mathcal{O}(h/H)$, this in turn can be bounded by

$$Ch^{-2}\left(\frac{h}{\beta H}\right)^2 \|u_h - u_0\|^2_{L^2(e)}.$$

Combining the above, we obtain

$$|u_i|_{H^1(e)} \leq \frac{C}{\beta^2 H^2}\|u_h - u_0\|^2_{L^2(e)} + 2|u_h - u_0|^2_{H^1(e)}.$$

Summing over all i and noting that only a finite number of u_i (bounded by the minimum number of colors N_c) is nonzero on the element e, we obtain

$$\sum_{i=1}^{p} |u_i|_{H^1(e)} \leq \left(\frac{C}{\beta^2 H^2} \|u_h - u_0\|^2_{L^2(e)} + 2|u_h - u_0|^2_{H^1(e)} \right) N_c.$$

Summing over all elements e in Ω, we obtain

$$\sum_{i=1}^{p} |u_i|_{H^1(\Omega)} \leq \left(\frac{C}{\beta^2 H^2} \|u_h - u_0\|^2_{L^2(\Omega)} + C|u_h - u_0|^2_{H^1(\Omega)} \right) N_c.$$

Applying (4.4) to the first term and the triangle inequality to the second term on the right, we have

$$\sum_{i=1}^{p} |u_i|_{H^1(\Omega)} \leq C\beta^{-2} |u_h|^2_{H^1(\Omega)} + C|u_0|^2_{H^1(\Omega)}.$$

Using the H^1 stability of Q_0 in the second term on the right and the equivalence between the H^1 norm and the $a(.\,,.)$ norm, we obtain

$$\sum_{i=1}^{p} a(u_i, u_i) \leq C \left(1 + \beta^{-2}\right) a(u_h, u_h).$$

Adding (4.3), we obtain (4.1). □

Here, C is independent of h, H and β, but may depend on the coefficients, since we used the equivalence between the $a(.\,,.)$ norm and the H^1 norm. For an improved bound of $C\,(1 + \beta^{-1})$ and for bounds which are valid independently of the jumps in the coefficients, we refer the reader to Dryja and Widlund (1992b).

4.3. Theory for Schur complement based methods

The convergence rate of Schur complement based methods depends on the spectrum of the preconditioned Schur matrix $M^{-1}S$. In this section, we will describe some techniques for estimating the extreme eigenvalues of some preconditioned Schur systems (mainly in two dimensions).

First, we prove the following equivalence between S and A. Given a vector x_B on the boundary B (see Section 3 for notation), define the discrete *harmonic extension* $Ex_B \equiv -A_{II}^{-1} A_{IB} x_B$. Then we have the following fundamental result:

Lemma 1

$$A[Ex_B, x_B]^T = [0, Sx_B]^T;$$

and

$$x_B^T S x_B = [Ex_B, x_B] A [Ex_B, x_B]^T.$$

Proof. Direct computation from the block factorization of A. □

Thus, the action of S on x_B can be obtained by first computing Ex_B, followed by a matrix product of A with $[Ex_B, x_B]^T$, and restricting the result to the nodes on the interface B.

This lemma provides a framework for constructing suitable preconditioners M for S: *if M is a matrix defined for vectors x_B, such that the M energy of x_B (i.e. $x_B^T M x_B$) approximates the A energy of the* discrete harmonic extension $[Ex_B, x_B]^T$, *then M can be used as a preconditioner for S, provided M can be easily inverted.*

Theorem 17 (**Trace Theorem**) There exists a continuous linear map $\gamma : H^1(\Omega) \longrightarrow L^2(\partial\Omega)$ such that $\gamma u = u|_{\partial\Omega}$ for smooth functions $u \in C^\infty(\overline{\Omega})$. Furthermore

$$\|\gamma u\|_{H^{1/2}(\partial\Omega)} \leq C\|u\|_{H^1(\Omega)},$$

for some positive constant C, where $H^{1/2}(\partial\Omega)$ is a fractional Sobolev norm.

Proof. See Nečas (1967) and Lions and Magenes (1972). □

The map γ is often referred to as the *trace* map. $H^{1/2}(\partial\Omega)$ is a fractional index Sobolev space which can be defined by interpolation between $H^1(\partial\Omega)$ and $H^0(\partial\Omega) = L^2(\partial\Omega)$ (we omit this description; see Lions and Magenes (1972)).

Using the trace theorem, we can prove the following fundamental property of harmonic functions.

Lemma 2 Let L be a second-order uniformly elliptic operator and u be a function defined on any region D, such that $Lu = 0$ in the interior of D. Then the $H^1(D)$ semi-norm of u on D is equivalent to the $H^{1/2}(\partial D)$ semi-norm of u on the boundary ∂D, i.e. there exist positive constants c and C such that

$$c|u|^2_{H^{1/2}(\partial D)} \leq |u|^2_{H^1(D)} \leq C|u|^2_{H^{1/2}(\partial D)}, \text{ for all } u \in H^1(D).$$

Proof. The left inequality follows from the trace theorem. The right inequality follows from elliptic regularity for harmonic functions, see Lions and Magenes (1972), Nečas (1967) for a proof. □

The corresponding result also holds for *discrete harmonic* functions, with constants c and C independent of mesh size h.

Theorem 18 If $u_h \in V^h(D)$ is a finite element function defined on a region D, such that u_h is discrete harmonic in D, then there exist constants c and C, independent of h such that

$$c|u_h|^2_{H^{1/2}(\partial D)} \leq |u_h|^2_{H^1(D)} \leq C|u_h|^2_{H^{1/2}(\partial D)}.$$

Proof. The left inequality follows from the trace theorem (as in the continuous case). The right inequality can be proved by using an *extension* theorem for finite element functions (which extends finite element functions defined on the boundary of a domain into the interior, such that the H^1 norm of the extension is bounded in terms of the $H^{1/2}$ norm of the boundary data), with a constant C independent of the mesh size h. Such an extension theorem was established by Widlund (1987), Bramble *et al.* (1986b), and Bjørstad and Widlund (1986). □

Thus, if a matrix M is the matrix representation of the bilinear form given by the $H^{1/2}(\partial D)$ inner product restricted to the finite element space $V^h(\partial D)$, then M is spectrally equivalent to S, the Schur complement obtained if $B = \partial\Omega$. The matrix M can be obtained by interpolation as follows.

4.4. Interface preconditioners for two-dimensional problems

Let K_B denote the discretization of $-\Delta$ on edge B, with zero boundary conditions on the vertices ∂B. Additionally, let M_B denote the mass matrix representing the $L^2(B)$ inner product on B. Then, the matrix representation J_B of the $H^{1/2}(B)$ bilinear form (or more precisely, the $H_{00}^{1/2}(B)$ bilinear form, see Lions and Magenes (1972)) is obtained by matrix interpolation between K_B and M_B as follows

$$J_B^{1/2} = [M_B, K_B]_{1/2} \equiv M_B^{1/2} \left(M_B^{-1/2} K_B M_B^{-1/2} \right)^{1/2} M_B^{1/2},$$

see Bjørstad and Widlund (1986) and Bramble *et al.* (1986b). Since M_B is spectrally equivalent to a scaled identity matrix, $J_B^{1/2}$ can be replaced by a scaled version of $K_B^{1/2}$, which is precisely Dryja's preconditioner M_{D} as presented in Section 3.1.

Theorem 19 For a two subdomain partition, the condition number of the preconditioned Schur matrix $J_B^{-1/2} S$ is bounded by a constant C independent of h.

Proof. By construction, $J_B^{1/2}$ is the matrix representation of the $H_{00}^{1/2}(B)$ inner product, therefore

$$u_B^T J_B^{1/2} u_B = \|u_B\|^2_{H_{00}^{1/2}(B)},$$

where we have used u_B to denote both a finite element function and its vector representation. By a variant of Theorem 18, $\|u_B\|^2_{H_{00}^{1/2}(B)}$ is spectrally equivalent to

$$[Eu_B, u_B] A [Eu_B, u_B]^T,$$

which in turn is spectrally equivalent to $u_B^T S u_B$ by Lemma 1. Therefore $J_B^{1/2}$ is spectrally equivalent to S. □

4.5. Many subdomain nonoverlapping algorithms

The theory for estimating the convergence rates of many subdomain preconditioners for S can often be reduced to estimates based on the Schwarz algorithms, see Dryja *et al.* (1993), Dryja and Widlund (1990; 1993a). Here, we sketch some of the basic ideas by considering the vertex space preconditioner M_{vs} of Smith (1992) in two dimensions:

$$M_{vs}^{-1} = \sum_k R_{E_k}^T S_{E_k E_k}^{-1} R_{E_k} + \sum_i R_{VS_i}^T S_{VS_i VS_i}^{-1} R_{VS_i} + R_H^T A_H^{-1} R_H.$$

In the following, we will assume that A_H is replaced by $S_H = R_H S R_H^T$, in which case the above preconditioner becomes an additive Schwarz preconditioner for S, based on an overlapping decomposition of the interface B:

$$B = \bigcup_k \{E_k\} \cup \left\{ \bigcup_l VS_l \right\},$$

and additionally the use of a coarse solver.

The preconditioned Schur matrix $M_{vs}^{-1} S$ can thus be written as a sum of projections, orthogonal in the S based inner product:

$$M_{vs}^{-1} S = \sum_k P_{E_k} + \sum_i P_{VS_i} + P_H,$$

where

$$P_{E_k} \equiv R_{E_k}^T S_{E_k E_k}^{-1} R_{E_k} S, \quad P_{VS_i} \equiv R_{VS_i}^T S_{VS_i VS_i}^{-1} R_{VS_i} S$$

and

$$P_H = R_H^T S_H^{-1} R_H S.$$

The condition number can be estimated in terms of a partition property with constant C_0 and the number of colours N_c.

We now sketch briefly, a technique for reducing this to using a corresponding partition for $V^h(\Omega)$ in the $a(.,.)$ based norm. First, corresponding to each subregion of the interface, we define a decomposition of Ω as follows. Let Ω_{E_k} be a subdomain of size $\mathcal{O}(H)$ containing E_k, and partitioned into two disjoint regions by E_ks (for instance, let Ω_{E_k} be the union of the two subdomains adjacent to E_k). Similarly, for each vertex region VS_i, let Ω_{VS_i} denote a subregion of Ω of size $\mathcal{O}(H)$ containing VS_i, and which is partitioned into a small number of disjoint subregions by VS_i (for instance, let Ω_{VS_i} be a rectangular or quadrilateral patch covering the vertex region VS_i). Then,

- Given u_B defined on B, extend it *discrete harmonically* into the subdomains: $[Eu_B, u_B]^T$.
- Next, partition $[Eu_B, u_B]^T$ (the extension) using the spaces $\{V^h(\Omega_{E_k})\}$, $\{V^h(\Omega_{VS_i})\}$ and coarse space V_0 with a partition constant C_0 that can be estimated by the same partition lemma (which was stated earlier). Thus,

$$[Eu_B, u_B]^T = \tilde{u}_0 + \sum_k \tilde{u}_{E_k} + \sum_i \tilde{u}_{VS_i},$$

with

$$\sum a(\tilde{u}_i, \tilde{u}_i) \leq C_0 a(Eu_B, Eu_B) = C_0 S(u_B, u_B),$$

where $\tilde{u}_i$ denotes the same partition, suitably re-indexed. The last equality follows from the equivalence between the S-energy and the A-energy of discrete harmonic extensions. The constant C_0 is bounded independent of H and h.

- Next, restrict each $\tilde{u}_i$ onto B to obtain a partition for u_B on B.
- Finally, use the equivalence between the S-energy and the A-energy of discrete harmonic extensions with the additional fact that the $a(.\,,.)$ energy of each $\tilde{u}_i$ is greater than the $a(.\,,.)$ energy of the discrete harmonic extension of its values on B.

By combining the results above, the partition constant for the Schur based algorithm can be estimated, see Dryja *et al.* (1993) for the details.

4.6. Summary of convergence bounds

In Table 1, we summarize the known condition number bounds for several of the preconditioners described in Sections 2 and 3. In the last two columns, we list condition number bounds that are most appropriate (tighter) when the coefficients are mildly varying and when the coefficients are discontinuous with possibly large jumps, respectively. $C(a)$ refers to a constant independent of H and h but dependent on the coefficients a, while C refers to a constant independent of H, h and a (provided a is mildly varying in each subdomain Ω_i). For the Schwarz and vertex space algorithms, β refers to the overlap parameter.

5. Some practical implementation issues

The focus of the previous sections were on the development of the basic components of domain decomposition algorithms (at a certain level of abstraction). In order to implement these algorithms efficiently, possibly on a parallel computer, there are other more practical matters to consider as well. In this section, we shall briefly touch on several of these issues.

Table 1. *Upper bounds for condition numbers of various algorithms.*

Algorithm	Eqn	Mild Coeff.	Disc. Coeff.
2D BPS	(3.6)	$C\left(1+\log^2(H/h)\right)$	$C\left(1+\log^2(H/h)\right)$
2D vertex space	(3.7)	$C(a)\left(1+\log^2(\beta^{-1})\right)$	$C(\beta)\left(1+\log(H/h)\right)$
3D vertex space	(3.10)	$C(a)\left(1+\log^2(\beta^{-1})\right)$	$C(\beta)(H/h)$
2D additive Schwarz	(2.1)	$C(a)\left(1+\beta^{-1}\right)$	$C(\beta)\left(1+\log(H/h)\right)$
3D additive Schwarz	(2.1)	$C(a)\left(1+\beta^{-1}\right)$	$C(\beta)(H/h)$
3D wirebasket	(3.11)	$C\left(1+\log^2(H/h)\right)$	$C\left(1+\log^2(H/h)\right)$
2D Neumann–Neumann	(3.9)	$C\left(1+\log^2(H/h)\right)$	$C\left(1+\log^2(H/h)\right)$
3D Neumann–Neumann	(3.9)	$C\left(1+\log^2(H/h)\right)$	$C\left(1+\log^2(H/h)\right)$

5.1. Inexact subdomain solvers

Every step of a domain decomposition iteration normally requires the exact solution of a subdomain problem, and perhaps also a coarse problem. Although this usually costs less than the solution of the original problem on the whole domain, it can still be quite expensive and it is natural to try to use a cheaper approximate solver instead. Also, when the iterates are still far from the true solution, it seems wasteful to solve these subdomain problems exactly. The issue here is how to incorporate these inexact solvers properly into the existing framework.

In most of the domain decomposition algorithms we have introduced so far, the exact solves involving A_i^{-1} and A_H^{-1} can be replaced by inexact solves $\tilde{A}_i^{-1}$ and $\tilde{A}_H^{-1}$, which can be standard elliptic preconditioners themselves (e.g. multigrid, ILU, SSOR, etc.). However, in order to rigorously prove that the conjugate gradient method converges, the inexact solvers $\tilde{A}_i^{-1}$ and $\tilde{A}_H^{-1}$ must be *fixed, linear* operators, e.g. they cannot be a few steps of an adaptive iterative method that depends on the vector being operated on (e.g. a few steps of the conjugate gradient method). In practice, however, solving the local problems approximately with a Krylov space method may work fine.

For the overlapping additive Schwarz methods the modification is straightforward. For example, the *Inexact Solve Additive Schwarz Preconditioner* is simply:

$$\tilde{M}_{\text{as},2}^{-1} z = R_0^T \tilde{A}_H^{-1} R_H z + \sum_{i=1}^{p} R_i^T \tilde{A}_i^{-1} R_i z.$$

We caution, however, that replacing A_i by $\tilde{A}_i$ can potentially lead to divergence in multiplicative Schwarz iteration, unless the spectral radii

$$\rho(\tilde{A}_i^{-1} A_i) < 2,$$

see Bramble *et al.* (1991), Xu (1992a), Cai and Widlund (1993).

The Schur complement methods require more changes to accommodate inexact solves. For example, by replacing A_H^{-1} by $\tilde{A}_H^{-1}$ and $S_{E_iE_i}$ by $\tilde{S}_{E_iE_i}$ in the definitions of the Bramble–Pasciak–Schatz preconditioner M_2 (see (3.6)) and the vertex space preconditioner M_3 (see (3.7)), we can easily obtain relatively ill-conditioned inexact preconditioners $\tilde{M}_2$ and $\tilde{M}_3$ for S. The main difficulty is, however, that the evaluation of the product Sz_B still requires exact subdomain solves using A_{II}^{-1}. One way to get around this is to use an *inner* iteration using $\tilde{A}_i$ as a preconditioner for A_i in order to compute the action of A_{II}^{-1}. An alternative is to perform the iteration on the original system $Au = f$, and construct a preconditioner $\tilde{A}$ for A from the block factorization of A in equation (3.3) by replacing the terms A_{II} and S by $\tilde{A}_{II}$ and $\tilde{S}$, respectively, where $\tilde{S}$ can be either $\tilde{M}_2$ or $\tilde{M}_3$. However, care must be taken to scale $\tilde{A}_H$ and $\tilde{A}_i$ so that they are as close to A_H and A_i as possible respectively – it is not sufficient that the condition number of $\tilde{A}_H^{-1}A_H$ and $\tilde{A}_i^{-1}A_i$ be close to unity, because the scaling of the coupling matrix A_{IB} may be wrong. For more details, the reader is referred to Börgers (1989), Goovaerts (1989) and Goovaerts, Chan and Piessens (1991).

We note that, when set up properly, the use of inexact solvers does not compromise on the accuracy of the final converged solution – only the preconditioner is changed, see Gropp and Smith (1992).

5.2. The choice of the coarse grid size H

Another practical matter in implementing a domain decomposition algorithm is to decide how many subdomains to use, i.e. the coarse scale H. Since most of the domain decomposition algorithms we have described have convergence rates that are bounded independently (or only slightly dependent on) of H, the theory does not lead to a clear choice. If the fine grid is obtained as a refinement of a coarse grid, then H is naturally defined. Moreover, very often the choice of subdomains is dictated by geometric considerations, e.g. if the domain can be naturally decomposed into several subdomains with regular geometry on which fast solvers can be used. Finally, in a parallel setting, it is natural to match the number of subdomains to the number of processors available. The choice of H must take all these factors into account and there are no guidelines that will work in all situations.

However, from a purely computational complexity standpoint, it is possible to make a more rational decision based on minimizing the computational cost. Given h, it has been observed empirically (Keyes and Gropp, 1989; Smith, 1990; Gropp and Smith, 1992) that there often exists an optimal value of H which minimizes the total computational time for solving for the converged solution. A small H provides a better, but more expensive, coarse grid approximation, and requires solving more subdomain problems

Table 2. *Complexity of solvers on an n^3 grid with coarse grid size n_H. (MIC: modified incomplete Cholesky.)*

Basic solver	Complexity	Optimal n_H	Complexity of domain decomposition solver using optimal n_H
Multigrid	$\mathcal{O}(n^3)$	1	$\mathcal{O}(n^3)$
MIC	$\mathcal{O}(n^{3.5})$	$0.61n^{7/8}$	$\mathcal{O}(n^{3.06})$
Nested dissection	$\mathcal{O}(n^6)$	$0.93n^{2/3}$	$\mathcal{O}(n^4)$
Band-Cholesky	$\mathcal{O}(n^7)$	$0.95n^{7/11}$	$\mathcal{O}(n^{4.45})$
Solver n^α	$\mathcal{O}(n^\alpha), \alpha \to \infty$	$n^{1/2}$	$\mathcal{O}(n^{\alpha/2})$

of smaller size. A large H has the opposite effect. If we make the assumption that the *same solver* is used for the subdomain problems as well as for the coarse problems, and that the convergence rate is independent of H (which is true in practice for most optimal methods), then one can derive an asymptotically optimal value of H (Chan and Shao, 1993). For example, on a one-processor architecture, for a model problem on a uniform d-dimensional grid with mesh size h and a solver with complexity $\mathcal{O}(m^\alpha)$ on an m^d grid, the optimal choice is

$$H_{\text{opt}} = \left(\frac{\alpha}{\alpha - d}\right)^{1/(\alpha-d)} h^{\alpha/(2\alpha-d)},$$

and the complexity of the overall domain decomposition solver using H_{opt} is $\mathcal{O}(h^{-\alpha/(2\alpha-d)})$, which can be significantly smaller than $\mathcal{O}(h^{-\alpha})$, the complexity of using the same solver to solve the whole problem without using a domain decomposition method. For example, in three dimensions ($d = 3$), the complexities are summarized in Table 2, where $n \equiv 1/h$.

In a parallel environment, if we assume that each subdomain solve is performed in parallel on the individual processors, and that the coarse solve is performed on one of the processors, either sequentially after or in parallel with the subdomain solves, then it turns out, ignoring communication costs (whether this is valid depends on the problem size and the particular hardware), the optimal value of H is $H_{\text{opt}} = \sqrt{h}$, independent of α and d. The optimal number of processors is $n^{d/2}$, and the execution time using H_{opt} is $\mathcal{O}(n^{\alpha/2})$.

In practice, it may pay to empirically determine a near optimal value of H if the preconditioner is to be re-used many times. The above asymptotic results for the model problem can be used as a guide.

5.3. Partition of the domain

In addition to deciding *how many* subdomains to use, it is also necessary to *identify* them. Very often, the domain Ω is already discretized and the problem is to decompose the grid itself. This can be viewed as a graph partitioning problem. The geometry of the domain can usually provide some guidance, e.g. subdomains with regular geometry are preferable. In a parallel setting, it is also desirable to have connections (i.e. edges) between neighbouring subdomains to be minimized (which would in turn minimize the communication cost) and to have the load (e.g. the number of grid points) in each subdomain balanced. For a structured and quasi-uniformly refined grid, one can often do this decomposition at a coarse level either by inspection or by brute force. For unstructured grids, finding the optimal decomposition is an NP-complete problem. There have been several heuristic approaches proposed, including geometric approaches such as the recursive coordinate bisection method (Fox, 1988; Berger and Bokhari, 1987) and the inertia method (Farhat and Lesoinne, 1993); recursive graph based approaches such as the Kernighan and Lin (1970) exchange method, the minimum bandwidth method and the spectral partitioning method (Pothen, Simon and Liou, 1990); and global minimization techniques such as using simulated annealing (Williams, 1991). These techniques trade off efficiency with the ability to find good partitions, and it is not clear at this point which method is the best. Recent surveys can be found in Simon (1991) and Farhat and Lesoinne (1993).

5.4. Solving the coarse problem in parallel

The most natural way of mapping a domain decomposition algorithm onto a parallel architecture is to map the subdomains to individual processors. In this setting, the solution of the coarse problem often presents some difficulties because the data are scattered among all the processors. If not done carefully, the coarse solve can dominate the execution time of the domain decomposition method. There are several obvious alternatives:

1 keep the data in place and solve it using a parallel method with data exchanges at each step;

2 gather the data in *one* processor, solve there and broadcast the result;

3 gather the data to *all* processors and solve it on all of them in parallel.

According to Gropp (1992), the last two approaches are often better than the first and on typical architectures. For parallel implementations of domain decomposition methods, see Bjørstad and Skogen (1992) and Smith (1993).

5.5. To overlap or not to overlap?

There is no definitive answer to this question but here are some guidelines. First, the overlapping method is generally easier to describe, implement and understand. It is also easier to achieve an optimal convergence rate and often more robust. On the other hand, extra work is performed on the overlapped regions. Moreover, if the coefficients are discontinuous across the subdomains, the extended subdomains must necessarily have discontinuous coefficients, making their solution more problematic. Recently, Bjørstad and Widlund (1989) and Chan and Goovaerts (1992) have shown that there is a fundamental relationship between the two approaches: the overlapping method is equivalent to a nonoverlapping method with a specific interface preconditioner. One can think of the overlapping method implicitly computing the effect of this preconditioner by the extra operations performed on the overlapping region.

6. Multilevel algorithms

In recent years, much research and interest has been focused on the development of multilevel algorithms to solve elliptic problems, that provide alternative preconditioners to the standard multigrid method. These multilevel algorithms include, for instance, the hierarchical basis multigrid method of Yserentant (1986) and Bank, Dupont and Yserentant (1988), the BPX algorithm of Bramble *et al.* (1990), the multilevel algorithms of Axelsson and Vassilevski (1990), and the multilevel additive Schwarz algorithm of Zhang (1992b) (a similar idea was mentioned in the thesis of Xu (1989) and in Wang (1991)). Although strictly speaking these algorithms are not domain decomposition methods, they have similarities with Schwarz type domain decomposition methods (with inexact solves) where different grid levels and subspaces play the role of subregions, see for instance Xu (1992a). Additionally, a convergence theory has been developed that incorporates both multilevel and domain decomposition methods into a unified framework, see Xu (1992a) and Dryja and Widlund (1990).

6.1. Background on multilevel discretizations

Consider the Dirichlet boundary value problem for the elliptic problem (1.1) on Ω. In order to obtain a multilevel discretization of this problem, the domain Ω is first triangulated by a coarse grid $\tau^1(\Omega)$ consisting of elements of diameter h_1. By successive refinement of each element, (say by dividing each element into four pieces in two dimensions, etc) a refined triangulation $\tau^2(\Omega)$ is obtained with a mesh size of $h_2 = h_1/2$, and such that each element of $\tau^1(\Omega)$ is a union of elements of $\tau^2(\Omega)$. This procedure can be repeated a total of $J-1$ times, till the grid size $h_J = h_1/2^{J-1}$ on the finest level

J provides sufficient accuracy. We therefore have J nested triangulations $\tau^1(\Omega), \ldots, \tau^J(\Omega)$ of Ω.

On each grid level i, for $i = 1, \ldots, J$, we define the standard finite element space $V^{h_i}(\Omega) \subset H_0^1(\Omega)$ consisting of continuous piecewise linear functions based on a triangulation $\tau^i(\Omega)$, which vanish on the boundary $\partial\Omega$. Note that

$$V^{h_1}(\Omega) \subset V^{h_2}(\Omega) \subset \cdots \subset V^{h_J}(\Omega).$$

For $i = 1, \ldots, J$, we let A^{h_j} denote the stiffness matrix corresponding to the discretization of the elliptic problem on the jth level based on the finite element space $V^{h_j}(\Omega)$, and let M^{h_j} denote the mass matrix corresponding to the bilinear form generated by the L^2 inner product.

We now describe several multilevel preconditioners that correspond to additive Schwarz (additive subspace) preconditioners with suitably defined restriction maps R_j.

6.2. The hierarchical basis multigrid method

The hierarchical basis method of Yserentant (1986) and Bank *et al.* (1988) is based on a new multilevel *hierarchical* basis for the finite element space. Let I_j denote the standard finite element interpolation map:

$$I_j : V^{h_J}(\Omega) \to V^{h_j}(\Omega),$$

from the fine grid onto the nodal basis functions on grid level j. Then, by telescoping series, we obtain:

$$I_J = I_1 + (I_2 - I_1) + \cdots + (I_J - I_{J-1}).$$

Each of the terms $I_j - I_{j-1}$ represents grid functions on level j which are zero at the nodes corresponding to the coarser grid level $j-1$. The range of these interpolation maps $I_j - I_{j-1}$ (i.e. the new nodes on each level) will correspond to the 'subdomains' in a Schwarz (subspace) method.

The hierarchical basis multigrid preconditioner M for A is an additive subspace (Schwarz) preconditioner of the form:

$$M_{\text{hb}}^{-1} = \sum_{j=1}^{J} R_j^T D_j^{-1} R_j,$$

with restriction map $R_j \equiv I_j - I_{j-1}$, and where the local matrices $A_j = R_j A R_j^T$ are replaced by its diagonal D_j, resulting in an inexact solve. See Bank *et al.* (1988), Xu (1992a) for details. In two dimensions, cond $(M^{-1}A)$ is bounded by $\mathcal{O}(1+\log^2(h))$, but in three dimensions this bound deteriorates to $\mathcal{O}(h^{-1})$, see Yserentant (1986) and Ong (1989).

6.3. The BPX algorithm

The BPX preconditioner of Bramble *et al.* (1990) can also be viewed as an additive subspace (Schwarz) preconditioner:

$$M^{-1} \equiv \sum_{j=1}^{J} R_j^T A_j^{-1} R_j,$$

where R_j^T denotes the interpolation map from the jth grid level to the finest grid, and R_j corresponds to a weighted restriction. Additionally, the exact local matrices $A_j = R_j A R_j^T$ can be further approximated by $ch_j^{d-2} I$ for second-order uniformly elliptic problems without deterioration in the convergence rates. The resulting preconditioner is

$$M_{\text{BPX}}^{-1} \equiv \sum_{j=1}^{J} R_j^T h_j^{2-d} R_j.$$

The convergence rate of the BPX algorithm is optimal.

Theorem 20 There exists a constant C independent of h_i and J such that

$$\text{cond}\,(M_{\text{BPX}}^{-1} A_J) \leq C.$$

Proof. The original convergence bound due to Xu (1989) and Bramble *et al.* (1990) was J^2 (J with full elliptic regularity), i.e. deteriorated mildly with increasing number of levels. A different proof by Zhang (1992b) improved the bound to J. Bounds by Oswald (1991) are optimal, independent of J. For alternative proofs, see Griebel (1991) and Bornemann and Yserentant (1993). □

We note that when implementing the restriction and interpolation maps R_i and R_i^T respectively, it is easier and more efficient to obtain $R_i z$ from $R_{i+1} z$ as in a standard multigrid algorithm.

6.4. Multilevel additive Schwarz algorithm

We note that the above version of the BPX algorithm does not take into account the variation in the coefficients of the elliptic problems. In this section, we describe the multilevel additive Schwarz algorithm of Zhang (1992b; 1991) which generalizes the BPX algorithm by including overlapping subdomains on each grid level, and which takes coefficients into account in the preconditioning.

The multilevel Schwarz algorithm is based on the same J grid levels as the previous algorithms. However, the elements $\{e_{h_j}\}$ on grid level j are decomposed into a collection of N_j overlapping subdomains $\Omega_1^{h_j}, \ldots, \Omega_{N_j}^{h_j}$:

$$\Omega \subset \left(\Omega_1^{h_j} \cup \cdots \cup \Omega_{N_j}^{h_j}\right),$$

where the diameter of each jth level subdomain $\Omega_i^{h_j}$ is $\mathcal{O}(h_{j-1})$ (which is the size of the preceding coarser level). Additionally, it is assumed, that the size of the overlap between the adjacent subregions on grid level j, is βh_{j-1}.

For all the subdomains, on all the grid levels, the following interpolation maps are defined:

$$R^T_{\Omega_i^{h_j}} : V^{h_j}(\Omega_i^{h_j}) \cap H_0^1(\Omega_i^{h_j}) \longrightarrow V^{h_J}(\Omega),$$

where $R^T_{\Omega_i^{h_j}}$ is the extension map from the nodal values on the interior grid points in $\Omega_i^{h_j}$ on the jth grid level to the finest grid level J. Its transpose $R_{\Omega_i^{h_j}}$ is a weighted restriction map onto the interior nodes in subdomain $\Omega_i^{h_j}$ on the jth grid level. The local stiffness matrix corresponding to subregion $\Omega_i^{h_j}$ on the jth grid level is denoted $A_{\Omega_i^{h_j}}$, where

$$A_{\Omega_i^{h_j}} = R_{\Omega_i^{h_j}} A R^T_{\Omega_i^{h_j}},$$

is a principal submatrix of the jth level stiffness matrix A_{h_j}.

The multilevel additive Schwarz preconditioner M_{mlas} is defined by

$$M^{-1}_{\text{mlas}} z = \sum_{j=1}^{J} \sum_{i=1}^{N_j} R^T_{\Omega_i^{h_j}} A^{-1}_{\Omega_i^{h_j}} R_{\Omega_i^{h_j}} z.$$

We note that this corresponds to a sum of additive Schwarz preconditioners on each grid level with suitably chosen subdomain sizes. The convergence rate of the multilevel additive Schwarz algorithm is described in the following theorem.

Theorem 21 Suppose that the mesh sizes satisfy: $h_i/h_{i-1} \leq cr$, where $r < 1$, and that the subregions on grid level j satisfy $\text{Area}(\Omega_i^{h_j}) \approx h_{j-1}$. Then,

$$\text{cond}\,(M^{-1}_{\text{mlas}} A) \leq C(r, a),$$

where the constant $C(r, a)$ can depend on r and the coefficients a, but is independent of J and h_i.

Proof. See Zhang (1992b; 1991). □

Remarks

- The preconditioner M_{mlas} can be obtained as a special case of the BPX preconditioner by choosing the 'smoothing' operator in Xu (1989) to be the additive Schwarz preconditioner. Conversely, the BPX preconditioner for the discrete Laplacian can be obtained as a special case of the multilevel additive Schwarz algorithm by choosing each subdomain

$\Omega_i^{h_j}$ on the jth grid level to contain only one interior point from the jth grid level, i.e. with minimal overlap amongst subdomains on each grid level. In this case, the local matrices $A_{\Omega_i^{h_j}} = ch_j^{d-2}$ ($\Omega \subset R^d$) will be 1×1, and correspond to the diagonal entries of the stiffness matrix on the jth grid level A_{h_j}. Additionally: $\sum_{i=1}^{N_j} R_{\Omega_i^{h_j}} = R_j$, the weighted restriction map onto the jth grid level.

- We note that using the submatrices $A_{\Omega_i^{h_j}}$ on level j provides the scaling based on the coefficients and computing (or approximating) them involves some overhead cost.
- We may skip a few levels of refinement, and the convergence rate will depend only on the ratio of the relevant mesh sizes.
- A multiplicative version has been considered in Wang (1991).

7. Algorithms for locally refined grids

In this section we describe domain decomposition algorithms for solving the linear systems arising from discretizations of elliptic partial differential equations on composite grids obtained by local refinement on subregions of Ω. The discretizations we consider are based on the use of 'slave variables' on the interface separating the different refined regions, see Bramble, Ewing, Pasciak and Schatz (1988), McCormick (1989), Widlund (1989a). Our description will be brief, and our goal is to formulate the problem so that the same domain decomposition methodology of Schwarz methods can be applied. Indeed, a composite grid is the union of various 'subgrids' on different subregions, see Figure 4, and these 'subgrids' correspond to 'subdomains' in a Schwarz method.

7.1. Discretization of elliptic problems on locally refined grids

Consider the elliptic problem (1.1) on a domain Ω, which is triangulated by a quasi-uniform grid $\tau^h(\Omega)$ of mesh size h. The local refinement procedure is applied to a sequence of nested subregions: $\Omega_p \subset \cdots \subset \Omega_2 \subset \Omega_1 \equiv \Omega$. Starting with a quasi-uniform triangulation $\tau^{h_1}(\Omega_1)$ with mesh size h_1, all elements from this triangulation lying in Ω_2 are uniformly refined, for instance with mesh size $h_2 = h/2$ resulting in the local triangulation $\tau^{h_2}(\Omega_2)$. The process is repeated, with successive refinements on each nested subregion, with local triangulations $\tau^{h_i}(\Omega_i)$ for $i = 2, \ldots, p$, where $h_i = h_{i-1}/2$, see Figure 4.

Corresponding to each local grid $\tau^{h_i}(\Omega_i)$ let $V^{h_i}(\Omega_i) \subset H_0^1(\Omega_i)$ denote the space of continuous, piecewise linear finite element functions vanishing outside Ω_i. The composite finite space $V^{h_1,h_2,\ldots,h_p}$ is defined as the sum of the local spaces:

$$V^{h_1,h_2,\ldots,h_p} = V^{h_1}(\Omega_1) + V^{h_2}(\Omega_2) + \cdots + V^{h_p}(\Omega_p).$$

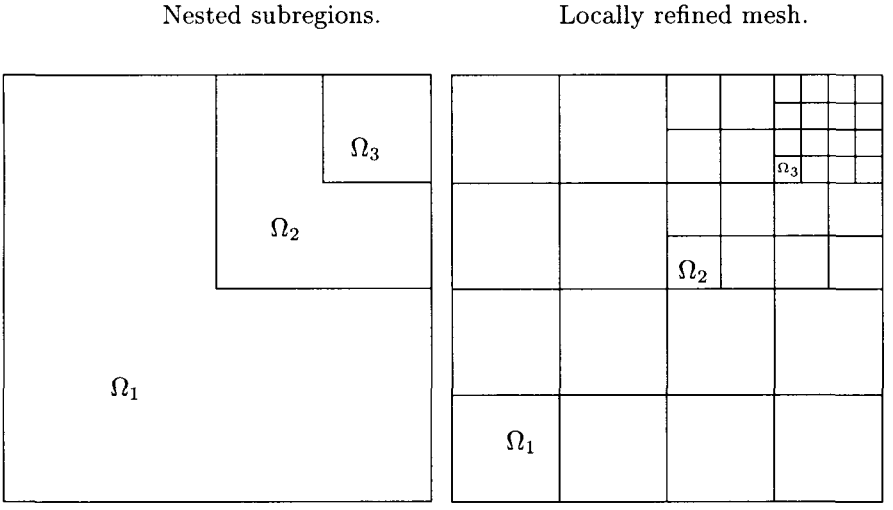

Fig. 4. Nested subregions with repeated local refinement.

The elliptic problem is discretized using the standard Galerkin procedure based on the composite finite element space $V^{h_1,h_2,\ldots,h_p}$ resulting in a linear system

$$Au = f, \tag{7.1}$$

see McCormick (1984), Bramble *et al.* (1988) and Widlund (1989a) for the details.

Throughout the rest of this section, we will use $\Omega_i^{h_i}$ to denote $\tau^{h_i}(\Omega_i)$, the ith refined grid on Ω_i, where for $i = 1$ this corresponds to the initial triangulation of Ω. For $i = 1, \ldots, p$, we let $R^T_{\Omega_i^{h_i}}$ denote the interpolation (extension) map from $V^{h_i}(\Omega_i)$ to the composite grid $V^{h_1,h_2,\ldots,h_p}$ and let $R_{\Omega_i^{h_i}}$ denote the corresponding restriction map. The local stiffness matrices are given by $A_{\Omega_i^{h_i}} = R_{\Omega_i^{h_i}} A R^T_{\Omega_i^{h_i}}$.

7.2. The Bramble–Ewing–Pasciak–Schatz (BEPS) algorithm for solving two-level problems

For the case of just one level of refinement (i.e. $p = 2$), Bramble *et al.* (1988) proposed a preconditioner M_{BEPS} for system (7.1) that corresponds to a symmetrized multiplicative Schwarz preconditioner sweeping over the grids $\Omega_i^{h_i}$ for $i = 2, 1, 2$ respectively, with zero initial iterate. The BEPS preconditioner therefore involves inversion of $A_{\Omega_1^{h_1}}$ once and $A_{\Omega_2^{h_2}}$ twice.

We refer the reader to Bramble *et al.* (1988) for the algorithmic details and the proof of the following convergence theorem.

Theorem 22 There exists a constant C, independent of h_1 and h_2, such that

$$\text{cond}\,(M_{\text{BEPS}}^{-1}A) \leq C.$$

For a more parallelizable variant of the BEPS preconditioner, see Bramble, Ewing, Parashkevov and Pasciak (1992).

7.3. The FAC and AFAC algorithms for composite grids

The FAC (*Fast Adaptive Composite Grid Method*) and AFAC (*Asynchronous Fast Adaptive Composite Grid Method*) algorithms (McCormick, 1984; Mandel and McCormick, 1989; Widlund, 1989b) for solving (7.1) can be viewed as multilevel generalizations of the BEPS algorithm. The FAC algorithm corresponds to a multiplicative Schwarz algorithm based on the 'subproblems' on the refined grids $\Omega_i^{h_i}$ with matrices $A_{\Omega_i^{h_i}}$, restriction and extension maps $R_{\Omega_i^{h_i}}$ and $R^T_{\Omega_i^{h_i}}$ respectively, for $i = 1, \ldots, p$, see McCormick (1989) and Widlund (1989b) for the algorithmic details and the proof of the following convergence theorem.

Theorem 23 The convergence factor ρ of the FAC iteration is independent of the mesh sizes h_i and the number of levels, p, and depends only on the ratio $\max\{h_i/h_{i-1}\}$ and on the ratio of volumes (or areas) $\max\{|\Omega_{i-1}|/|\Omega_i|\}$.

An additive preconditioner M_{FAC} corresponding to the FAC iteration is

$$M_{\text{FAC}}^{-1}f \equiv \sum_{i=1}^{p} R^T_{\Omega_i^{h_i}} A^{-1}_{\Omega_i^{h_i}} R_{\Omega_i^{h_i}} f.$$

The convergence is not as good as the multiplicative version.

Theorem 24 There exists a constant C, independent of the mesh sizes h_i and the number of levels p, such that

$$\text{cond}\,(M_{\text{FAC}}^{-1}A) \leq Cp,$$

Proof. See Widlund (1989b) and McCormick (1989). □

Part of the reason why M_{FAC} is nonoptimal is that some of the grid points in the refined regions are redundantly accounted for by all coarser level terms in the preconditioner. In the AFAC preconditioner (see Mandel and McCormick (1989), Widlund (1989b)), this redundancy is removed explicitly and *optimal* convergence is restored.

We introduce the following additional notation.

- For $i = 2, \ldots, p$, we use $A_{\Omega_i^{h_{i-1}}}$ to denote the stiffness matrix obtained by discretizing the elliptic problem based on the triangulation $\tau^{h_{i-1}}(\Omega_i)$ on Ω_i, i.e. using the space $V^{h_{i-1}}(\Omega_i) \cap H_0^1(\Omega_i)$.
- For $i = 2, \ldots, p$, the following additional *extension* maps will be used:

$$R^T_{\Omega_i^{h_{i-1}}} : V^{h_{i-1}}(\Omega_i) \to V^{h_1, h_2, \ldots, h_p},$$

which denotes extension of interior nodal values on the grid $\tau^{h_{i-1}}(\Omega_i)$ to the composite grid. Its transpose will be a weighted *restriction* map onto the nodes in $\tau^{h_{i-1}}(\Omega_i)$.

The AFAC preconditioner M_{AFAC} is defined by

$$M^{-1}_{\text{AFAC}} \equiv R^T_{\Omega^h} A^{-1}_{\Omega^h} R_{\Omega^h} + \sum_{i=2}^{p} \left(R^T_{\Omega_i^{h_i}} A^{-1}_{\Omega_i^{h_i}} R_{\Omega_i^{h_i}} - R^T_{\Omega_i^{h_{i-1}}} A^{-1}_{\Omega_i^{h_{i-1}}} R_{\Omega_i^{h_{i-1}}} \right).$$

Thus, the AFAC preconditioner requires solving two subproblems (with different grid sizes) on each refined subregion Ω_i.

Theorem 25 There exists a constant C, independent of the mesh sizes h_i and the number of levels p, and dependent only on the ratios of the mesh sizes h_{i-1}/h_i and the ratios of the areas (or volumes) of the refined regions, such that

$$\text{cond}\,(M^{-1}_{\text{AFAC}} A) \leq C.$$

Proof. See Widlund (1989b), Dryja and Widlund (1989a) and McCormick (1989). □

8. Domain imbedding or fictitious domain methods

A dual approach to domain decomposition is the *domain imbedding* or *fictitious domain* method (another name is *capacitance matrix* method), in which problems on irregular domains are imbedded into larger problems on regular domains (such as rectangles or cubes) on which fast solvers are available, and the solution to the original problem is obtained iteratively by solving a sequence of problems on the extended domain. We will follow here the approach of Buzbee, Dorr, George and Golub (1971), Proskurowski and Widlund (1976), O'Leary and Widlund (1979), Börgers and Widlund (1990), Proskurowski and Vassilevski (1994). A rich literature on fictitious domain methods is found in the Soviet literature, and we refer the reader to Astrakhantsev (1978), Lebedev (1986), Marchuk *et al.* (1986) and Finogenov and Kuznetsov (1988), for details and references. Recently, very interesting alternative approaches based on control theory and optimization have been proposed for fictitious domain methods, and we refer the reader to Atamian, Dinh, Glowinski, He and Périaux (1991).

In this section, we briefly describe two examples of domain imbedding methods for solving a coercive (positive definite) Helmholtz problem on a domain Ω_1:

$$-\Delta u + cu = f, \quad \text{in } \Omega_1, \quad \text{where } c \geq 0$$

with either Dirichlet boundary conditions $u = g_D$ or Neumann boundary conditions $\partial u/\partial n = g_N$ on $\partial\Omega_1$. In case $c = 0$, then the Neumann boundary data g_N must satisfy the standard compatibility conditions with f.

We imbed Ω_1 in a regular domain (for instance a rectangle or cube) $\Omega \supset \Omega_1$ and define $\Omega_2 = \Omega - \Omega_1$. The interface separating the two subregions will be denoted by $B = \partial\Omega_1 \cap \partial\Omega_2$ (which may equal $\partial\Omega_1$, in case Ω_1 is completely imbedded in Ω). The extended elliptic problem on Ω, in the above case will be the same Helmholtz problem (assuming that c is constant). We assume that the extended problem on Ω is discretized (by either finite element or finite difference methods) resulting in the linear system $Au = f$. We partition the unknowns as $u = [u_1, u_2, u_3]^T$, where u_1 and u_2 corresponds to the interior nodes in Ω_1 and Ω_2, respectively, while u_3 corresponds to the nodes on the interface B separating the two regions. The extended linear system then has the following block form:

$$\begin{bmatrix} A_{11} & 0 & A_{13} \\ 0 & A_{22} & A_{23} \\ A_{13}^T & A_{23}^T & A_{33} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix}, \tag{8.1}$$

where A_{ii} are the coefficient matrices corresponding to the Dirichlet problem on Ω_i, for $i = 1, 2$, etc.

In the following two subsections, we describe imbedding methods for solving Neumann and Dirichlet problems on Ω_1.

8.1. Preconditioner M_N for the Neumann problem on Ω_1

Here we describe a domain imbedding preconditioner for the Neumann problem on Ω_1, following the development in Börgers and Widlund (1990). Using the block ordering in (8.1), the linear system corresponding to the Neumann problem on Ω_1 is

$$A_N \begin{bmatrix} u_1 \\ u_3 \end{bmatrix} \equiv \begin{bmatrix} A_{11} & A_{13} \\ A_{13}^T & A_{33}^{(1)} \end{bmatrix} \begin{bmatrix} u_1 \\ u_3 \end{bmatrix} = \begin{bmatrix} g_1 \\ g_3 \end{bmatrix},$$

where $A_{33}^{(1)}$ corresponds to the contribution to A_{33} from Ω_1. We note that this matrix may be singular, in case $c = 0$ for the Helmholtz problem, with $[1, \ldots, 1]^T$ in its null space. In such cases, care must be exercised to ensure that the conjugate gradient iterates remain orthogonal to the null space.

The action of the inverse M_N^{-1} of a domain imbedding preconditioner M_N

to the above problem is defined by

$$M_{\mathrm{N}}^{-1} \begin{bmatrix} g_1 \\ g_3 \end{bmatrix} \equiv \begin{bmatrix} I & 0 & 0 \\ 0 & 0 & I \end{bmatrix} A^{-1} \begin{bmatrix} I & 0 \\ 0 & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} g_1 \\ g_3 \end{bmatrix}.$$

This involves the solution of the extended problem with right-hand sides $g_1, 0$ and g_3 on Ω_1, Ω_2 and B respectively,

By using the block factorization of A, it can be easily verified that

$$M_{\mathrm{N}} = \begin{bmatrix} A_{11} & A_{13} \\ A_{13}^T & A_{33}^{(1)} + S^{(2)} \end{bmatrix},$$

where $S^{(2)} = A_{33}^{(2)} - A_{23}^T A_{22}^{-1} A_{23}$ is the Schur complement of the nodes on B with respect to the nodes in the domain Ω_2. Thus, the preconditioner M_{N} is a modification of the Neumann problem, by the addition of the Schur complement to a diagonal block. The convergence rate is optimal.

Theorem 26 The exists a constant C, independent of h, such that

$$\mathrm{cond}\,(M_{\mathrm{N}}^{-1} A_N) \leq C.$$

Proof. See Börgers and Widlund (1990). □

Finally, we note that the problem of choosing a grid on Ω that allows a fast solver, and whose restriction on Ω_1 allows for suitable discretization on Ω_1 is discussed at length in Börgers and Widlund (1990), where a triangulation algorithm is also described. Additionally, we note that exact solvers on Ω may be replaced by suitable inexact solvers, especially based on a topologically equivalent grid, without affecting the optimal convergence rate.

8.2. Capacitance matrix solution of the Dirichlet problem on Ω_1

Here, we consider the solution of the following linear system corresponding to the Dirichlet problem on Ω_1:

$$A_{11} u_1 = f_1.$$

Unfortunately, a straightforward modification of preconditioner M_{N} to the Dirichlet case, i.e.

$$\hat{M}_{\mathrm{D}}^{-1} f_1 \equiv \begin{bmatrix} I & 0 & 0 \end{bmatrix} A^{-1} \begin{bmatrix} I \\ 0 \\ 0 \end{bmatrix} f_1,$$

does not work very well. Indeed, $\mathrm{cond}\,(\hat{M}_{\mathrm{D}}^{-1} A_{11})$ grows as $\mathcal{O}(h^{-1})$, see Börgers and Widlund (1990). An alternative preconditioner based on the

Neumann problem for the exterior domain $\Omega_2 = \Omega - \Omega_1$, is described in the same article.

The solution procedure we describe for the Dirichlet problem will be based on a recently proposed capacitance matrix algorithm of Proskurowski and Vassilevski (1994). The solution of $A_{11}u_1 = f_1$ will be computed in a few stages, just as in Schur complement based domain decomposition methods, and it is based on the following two matrix identities relating the the solution u_1 on Ω_1 to the extended problem on Ω:

Lemma 3 Let the Schur complement of A be

$$S \equiv A_{33} - A_{13}^T A_{11}^{-1} A_{13} - A_{23}^T A_{22}^{-1} A_{23}.$$

Then the following identities hold:

$$(1) \quad A_{11}^{-1} = \begin{bmatrix} I & 0 & 0 \end{bmatrix} A^{-1} \left[\begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & S \end{bmatrix} A^{-1} \right] \times \begin{bmatrix} I & 0 & 0 \end{bmatrix}^T,$$

$$(2) \quad C \equiv S^{-1} = \begin{bmatrix} 0 & 0 & I \end{bmatrix} A^{-1} \begin{bmatrix} 0 & 0 & I \end{bmatrix}^T.$$

Proof. This can be verified directly using the block factorization of A. □

The algorithm is a direct implementation of the first identity.

Capacitance matrix method for solving $A_{11}u_1 = f_1$

1 Solve $A \begin{bmatrix} y_1, & y_2, & y_3 \end{bmatrix}^T = \begin{bmatrix} f_1, & 0, & 0 \end{bmatrix}^T$.

2 Compute $w_3 = Sy_3$ by solving $Cw_3 = y_3$, using a preconditioned conjugate gradient method, with a matrix–vector product involving C computed by identity (2) in the lemma above (requiring solves with A). The inverse of any preconditioner for S (e.g. from Section 3) can be used as a preconditioner for C.

3 Solve $A \begin{bmatrix} v_1, & v_2, & v_3 \end{bmatrix}^T = \begin{bmatrix} 0, & 0, & w_3 \end{bmatrix}^T$.

4 Set $u_1 = y_1 - v_1$.

Theorem 27 For preconditioners M for C such that M^{-1} is a spectrally equivalent preconditioner for S, $\mathrm{cond}\,(M^{-1}C)$ is bounded independent of the mesh size h.

Proof. See Proskurowski and Vassilevski (1994). □

We refer the reader to Atamian *et al.* (1991), and to Proskurowski and Vassilevski (1992) for domain imbedding algorithms for solving indefinite and nonsymmetric problems.

9. Convection–diffusion problems

In this section, we briefly describe some domain decomposition algorithms for solving the nonsymmetric linear systems arising from the discretization of convection–diffusion problems such as

$$-\epsilon\Delta u + \boldsymbol{b}\cdot\nabla u + c_0 u = f, \text{ in } \Omega, \; u = 0, \text{ on } \partial\Omega, \tag{9.1}$$

where $\epsilon > 0$ represents viscosity, $\boldsymbol{b}$ is a vector field and $c_0 \geq 0$. Though such problems are elliptic, they pose some difficulties for iterative solution. In case the diffusion term dominates the convection term, (such as when $\|\boldsymbol{b}\|h/\epsilon \ll 1$) most of the domain decomposition algorithms we have described, including the Schwarz and Schur methods, can be extended to solve the nonsymmetric problem, with suitable modifications such as replacing conjugate gradient methods by GMRES, BiCG, BiCGStab or QMR methods, see Freund, Golub and Nachtigal (1992). However, the convergence rates of the standard algorithms deteriorate as ϵ approaches zero, unless a coarse grid discretization of the original problem is solved exactly on a grid of size H, where $H < H_0$ (a constant), see Cai and Widlund (1992, 1993), Xu (1992b), Xu (1992c), Xu and Cai (1992). This coarse grid condition has been known in the multigrid literature. Additionally, for small diffusion, the solution is more strongly coupled along the characteristics of the convection problem, making the solution procedure sensitive to the ordering of nodes. Thus, the solution of these nonsymmetric problems by standard algorithms poses some difficulties when the convection term dominates.

In Sections 9.1 and 9.2, we describe the extension of several many subdomain overlapping and nonoverlapping algorithms to the nonsymmetric case. Following that, in Sections 9.3 and 9.4, we briefly describe alternative approaches that have been recently proposed by Gastaldi, Quarteroni and Sacchi-Landriani (1990), Glowinski, Périaux and Terrasson (1990b), and Ashby, Saylor and Scroggs (1992) based on two subdomain decompositions that couple elliptic and hyperbolic problems using an asymptotics approach.

Throughout this section, we will assume that problem (9.1) is discretized by a stable scheme (such as upwind finite differences, streamline diffusion finite elements or a scheme based on artificial viscosity), resulting in a linear system:

$$L(\epsilon)u = \epsilon Au + Cu = f, \tag{9.2}$$

where $A = A^T > 0$ is the discretization of the Laplacian, and C corresponds to the discretization of the convection and the $c_0 u$ term.

9.1. Schwarz algorithms for convection–diffusion problems

As in Section 2, let $\hat{\Omega}_1, \ldots, \hat{\Omega}_p$ denote an overlapping covering of Ω, with corresponding restriction and extension maps R_i and R_i^T, respectively. The

coarse grid restriction and extension maps will be denoted by R_H and R_H^T respectively.

A straightforward extension of the additive Schwarz preconditioner for $L(\epsilon)$ is defined by

$$M_{\text{as},1}^{-1} = R_H^T \left(\epsilon A_H + C_H\right)^{-1} R_H + \sum_{i=1}^{p} R_i^T \left(\epsilon A_i + C_i\right)^{-1} R_i,$$

where $\epsilon A_H + C_H = R_H L(\epsilon) R_H^T$ and $\epsilon A_i + C_i = R_i L(\epsilon) R_i^T$ are the coarse grid and local matrices, respectively. The corresponding linear system can be solved by any suitable nonsymmetric conjugate gradient like procedure. In the nonsymmetric case, we also have the following variant:

$$M_{\text{as},2}^{-1} = R_H^T \left(\epsilon A_H + C_H\right)^{-1} R_H + \sum_{i=1}^{p} R_i^T \left(\epsilon A_i\right)^{-1} R_i,$$

where the local convection–diffusion problems are replaced by more easily solvable (symmetric, positive definite) diffusion problems. The following convergence bounds have been established by Cai and Widlund (1993) and Xu and Cai (1992).

Theorem 28 There exists a maximum coarse grid size $H_0(\epsilon, h, \boldsymbol{b}, c_0)$ such that if $H < H_0(\epsilon, h, \boldsymbol{b}, c_0)$, then the rate of convergence of both the additive Schwarz preconditioned systems is independent of $H < H_0$ and h.

An explicit form for $H_0(\epsilon, h, \boldsymbol{b}, c_0)$ has not been derived in the literature (to the knowledge of the authors), but heuristically, it *may* depend on ϵ and h as

$$H_0 \approx \max\left\{\frac{\epsilon}{\|\boldsymbol{b}\|}, h\right\},$$

and this decreases as $\epsilon \to 0$. Consequently, the cost of solving the coarse grid problem can increase with smaller ϵ, and places some limitations on the convergence rate and efficiency of the algorithms, see Cai, Gropp and Keyes (1992).

The *multiplicative Schwarz* method can also be extended to the nonsymmetric case, analogously. However, to ensure convergence without acceleration, care must be exercised so that if approximation of the local problems are used, they must be spectrally close to the true local problems. We refer the reader to Xu (1992b), Cai and Widlund (1993), Xu and Cai (1992) and Wang (1993) for the details.

9.2. Schur complement based algorithms for convection–diffusion problems

As for the symmetric, positive definite case described in Section 3, we partition the domain Ω into p nonoverlapping subregions $\Omega_1, \ldots, \Omega_p$, with in-

terface B. The block form of the system becomes:

$$\begin{bmatrix} L_{II} & L_{IB} \\ L_{IB}^T & L_{BB} \end{bmatrix} \begin{bmatrix} u_I \\ u_B \end{bmatrix} = \begin{bmatrix} f_I \\ f_B \end{bmatrix}, \tag{9.3}$$

where $L_{II} = \epsilon A_{II} + C_{II}$, etc. The Schur complement system is:

$$Su_B = \tilde{f}_B, \text{ where } S = L_{BB} - L_{IB}^T L_{II}^{-1} L_{IB}, \text{ and } \tilde{f}_B = f_B - L_{IB}^T L_{II}^{-1} f_I.$$

The solution procedure is analogous to the symmetric, positive definite case. Once u_B is determined, u_I is obtained as $u_I = L_{II}^{-1}(f_I - L_{IB}u_B)$.

The nonsymmetric Schur complement system can be solved by a preconditioned iterative method (in conjunction with GMRES or suitable algorithms), with any of the preconditioners of Section 3. However, as previously noted, care must be exercised so that the size of the coarse grid problem is sufficiently small with $H < H_0$. For instance, the nonsymmetric BPS preconditioner has the form:

$$M_{\text{BPS}}^{-1} = R_H^T L_H^{-1} R_H + \sum_{i=1}^{n} R_{E_i}^T S_{E_i E_i}^{-1} R_{E_i},$$

where the edge problems $S_{E_i E_i}$ can be replaced by preconditioners applicable in the symmetric, positive definite case, or preferably preconditioners that adapt to the convection term. We refer the reader to Cai and Widlund (1993), D'Hennezel (1992) and Chan and Keyes (1990) for the details.

For a numerical comparison of both Schwarz and Schur complement algorithms, see Cai *et al.* (1992).

9.3. Elliptic–hyperbolic approximation of convection–diffusion problems

Classical asymptotics based studies of singular perturbation problems have much in common with domain decomposition. Typically, the domain is decomposed into two regions, one corresponding to a boundary or interior layer region and referred to as the *inner region*, where the full viscous problem is solved, and an *outer region*, where the inviscid or hyperbolic problem is solved. The inner and outer solutions are required to satisfy certain compatibility conditions on the interface or region of overlap between the two subregions. In problems where asymptotic expansions may not be tractable, an alternative is to use numerical approximations in each of the subregions, and to couple the solutions together using matching conditions. Several detailed and interesting studies have been conducted in the domain decomposition framework, and we provide references to some of the literature.

For second-order scalar elliptic convection diffusion problems, Gastaldi *et al.* (1990) proposed a mixed elliptic–hyperbolic approximation of the convection diffusion problem. The domain is partitioned into two nonoverlapping subregions: Ω_{E}, where the full elliptic problem is solved, and Ω_{H} where the

hyperbolic problem obtained by dropping the viscous term is solved. They proposed new *transmission boundary conditions* coupling the two subproblems, obtained by using a vanishing viscosity procedure. Additionally, a Dirichlet–Neumann type iterative procedure was proposed that solves the resulting mixed, elliptic–hyperbolic approximation of the convection diffusion problem. Theoretical and numerical estimates of the approximation error and convergence rates are provided in Gastaldi *et al.* (1990) and the references contained therein. A detailed theory has now been developed by Quarteroni and Valli (1990) for various heterogeneous approximations, and studies are being conducted for the compressible Navier–Stokes equations.

An alternative approach based on overlapping subregions was used by Glowinski *et al.* (1990b) for coupling the viscous and inviscid compressible Navier–Stokes equations. The domain is decomposed into two overlapping subregions corresponding to viscous and inviscid regions, and a least-squares minimization is applied to a functional of the two solutions on the region of overlap. The resulting least-squares problem is then solved via a nonlinear GMRES procedure.

For alternative studies, more closely aligned with classical boundary layer expansions, we refer the reader to Hedstrom and Howes (1990), Chin, Hedstrom, McGraw and Howes (1986), Gropp and Keyes (1993), and to Garbey (1992) and Scroggs (1989), for studies on conservation laws. An interesting domain decomposition method based on an approximate factorization of the convection–diffusion operator was recently proposed by Nataf and Rogier (1993).

9.4. Block preconditioners for convection–diffusion problems

In this section, we briefly describe an alternative block matrix preconditioner (without coarse grid solves) for the nonsymmetric linear system (9.2). This preconditioner was recently proposed by Ashby *et al.* (1992), motivated by matched asymptotic expansions, and is referred to as the *physically motivated domain decomposition preconditioner.*

We consider a decomposition of Ω into two regions, a hyperbolic region Ω_{H} and an elliptic region Ω_{E}, with an overlap of width equal to one grid size. Corresponding to this partition, the unknowns can be ordered $u = [u_1, u_2]^T$, where u_1 corresponds to the interior unknowns in the hyperbolic region Ω_{H} and u_2 corresponds to the interior unknowns in the elliptic region Ω_{E}. Note that due to one grid overlap, there are no 'boundary unknowns'. The linear system (9.2) then takes on the block form:

$$\begin{bmatrix} \epsilon A_{11} + C_{11} & \epsilon A_{12} + C_{12} \\ \epsilon A_{12}^T + C_{21} & \epsilon A_{22} + C_{22} \end{bmatrix} \begin{bmatrix} u_{\mathrm{H}} \\ u_{\mathrm{E}} \end{bmatrix} = \begin{bmatrix} f_{\mathrm{H}} \\ f_{\mathrm{E}} \end{bmatrix}.$$

Based on the above block partition, the physically motivated domain de-

composition preconditioner M_{pmdd} of Ashby *et al.* (1992) is defined by

$$M_{\text{pmdd}} = \begin{bmatrix} C_{11} & 0 \\ \epsilon A_{12}^T + C_{21} & \epsilon A_{22} + C_{22} \end{bmatrix}. \tag{9.4}$$

It is block lower triangular and inverting it involves inverting the two diagonal blocks. The motivation for setting the diffusion term to zero in the $(1,1)$ block is that it then corresponds to a hyperbolic problem on region Ω_{H} (analogous to asymptotic expansions for singular perturbation problems). For most direction fields $\boldsymbol{b}$, and for upwind finite difference discretizations, C_{11} can be inverted by 'marching along characteristics'. That is, if the subregion Ω_{H} is suitably chosen, the indices of the nodes in Ω_{H} may be reordered to produce a *lower triangular matrix* C_{11}, which can be easily solved since it is sparse. The block $\epsilon A_{22} + C_{22}$ may be more difficult to invert, since in the elliptic region the grid may be refined, and the diffusion term may dominate the convection term. In such cases, it may be suitable to replace $\epsilon A_{22} + C_{22}$ by the symmetric, positive definite matrix ϵA_{22} (or suitable parallelizable preconditioners).

Numerical tests conducted in Chan and Mathew (1993) indicate that on uniform grids, with suitably chosen elliptic and hyperbolic regions, the convergence rate of the M_{pmdd} preconditioned system improves as $\epsilon \to 0$, for fixed mesh size h. However, for fixed ϵ, as $h \to 0$, the convergence rate deteriorates mildly. It is speculated in Chan and Mathew (1993) that this deterioration may be due to the approximation of the elliptic term by a hyperbolic term in Ω_{H}. However, since in general, the mesh size does not need refinement on the hyperbolic region Ω_{H}, but only in the boundary layer region Ω_{E}, the above algorithm may be more robust with respect to local refinement in Ω_{E}.

A variant of this method was studied in Chan and Mathew (1993), and corresponds to a matrix version of the Dirichlet–Neumann preconditioner for the elliptic–hyperbolic approximation of Gastaldi *et al.* (1990), and is based on the use of a Neumann problem on Ω_{E}. In matrix terms, both preconditioners correspond to variants of the classical block Gauss–Seidel preconditioner, i.e. a block lower triangular matrix, whose diagonal blocks are modified to permit ease of solvability.

10. Parabolic problems

In this section, we briefly describe domain decomposition algorithms for solving the linear systems obtained by implicit discretization of parabolic problems. We consider the following model parabolic problem for $(x,t) \in$

$\Omega \times [0,T]$:

$$\begin{cases} u_t &= \nabla \cdot (a \nabla u) + f, & \text{on } \Omega \times [0,T], \\ u(x,0) &= u_0(x), & \text{on } \Omega, \\ u(x,t) &= 0, & \text{on } \partial\Omega \times [0,T]. \end{cases} \tag{10.1}$$

To be specific, we consider a discretization by finite differences in space and backward Euler in time, resulting in

$$\begin{cases} (u^{n+1} - u^n)/\tau = -Au^{n+1} + f^{n+1}, \\ u^0 = u_0^h, \end{cases}$$

where A is a symmetric positive definite matrix corresponding to the discretization of $-\nabla \cdot (a \nabla u)$ and τ is the time step. At each time step, the following linear system must be solved:

$$(I + \tau A)\, u^{n+1} = u^n + \tau f^{n+1}. \tag{10.2}$$

Similar equations are obtained for Crank–Nicholson in time, and finite elements in space. The implicit system (10.2) corresponds to a discretization of the elliptic operator $L(\tau)u = \tau u - \nabla \cdot (a \nabla u)$ and, consequently, most of the domain decomposition algorithms of Sections 2 and 3 are applicable. However, there are some crucial differences that make this system easier to solve: the condition number of $I + \tau A$ is bounded by $\mathcal{O}(\tau h^{-2})$ which can be relatively smaller than cond (A) if τ is small (say $\tau = \mathcal{O}(h)$ or $\tau = \mathcal{O}(h^2)$). Consequently:

- The entries of the Green function $(I + \tau A)^{-1}$ can be shown to decay more rapidly away from the diagonal than the entries of A^{-1}, and so depending on τ, a *coarse grid problem may not be required* for global communication of information.
- It is possible to use just one iteration of the domain decomposition method and still maintain a stable approximation preserving the local truncation error.

In Section 10.1, Schwarz algorithms are described for (10.2), with modifications in the coarse problem. In Sections 10.2 and 10.3, algorithms that require only *one* iteration are described.

10.1. Schwarz preconditioners for parabolic problems

We follow here the development due to Lions (1988) and Cai (1991; 1993). As in Section 2, we decompose Ω into an overlapping covering $\hat{\Omega}_1, \ldots, \hat{\Omega}_p$, with corresponding restriction and extension maps R_i and R_i^T, respectively. Similarly, R_H and R_H^T will denote the restriction and interpolation maps corresponding to the coarse grid. The local submatrices will be denoted $L_i(\tau) \equiv I_i + \tau A_i = R_i\,(I + \tau A)\, R_I^T$, and the coarse grid problem by $L_H(\tau) \equiv$

$R_H (I + \tau A) R_H^T$. We define two additive Schwarz preconditioners for $L(\tau) \equiv I + \tau A$:

$$M_{\text{as},1}^{-1} = \sum_{i=1}^{p} R_i^T L_i(\tau)^{-1} R_i,$$

and

$$M_{\text{as},2}^{-1} = \sum_{i=1}^{p} R_i^T L_i(\tau)^{-1} R_i + R_H^T L_H(\tau)^{-1} R_H.$$

The following convergence results are proved in Cai (1991).

Theorem 29 If $\tau \leq CH^2$, then $\text{cond}\,(M_{\text{as},1}^{-1} L(\tau))$ is bounded by a constant C_1 independent of τ, H and h. For larger τ, $\text{cond}\,(M_{\text{as},2}^{-1} L(\tau))$ is bounded by a constant C_2 independent of τ, H and h.

Thus, if $\tau \leq CH^2$, then a preconditioner without a coarse model may be used effectively. However, if τ is large, a coarse grid correction term must be used in order to maintain a constant rate of convergence. Similar results hold for multiplicative Schwarz methods and for Schur complement based methods. We refer the reader to Cai (1991) for the details.

10.2. One iteration based approximations: overlapping subdomains

As mentioned before, it is possible to obtain approximate solutions w^{n+1} of system (10.2) that are accurate to within the local truncation error of the true numerical solution u^{n+1}:

$$\|w^{n+1} - u^{n+1}\| \leq \mathcal{O}(\epsilon),$$

where $\mathcal{O}(\epsilon)$ is the local truncation error, and which can be constructed by solving only one problem on suitably chosen subdomains. Here, we briefly describe one such algorithm proposed by Kuznetsov (1991; 1988) and Meurant (1991).

Kuznetsov's method is based on the observation that the entries in the ith row of the discrete Green function $G(\tau)$ (where $G(\tau) = (I + \tau A)^{-1}$) decays rapidly as the distance between the nodes $\{x_i\}$ increases, specifically

$$|G_{ij}(\tau)| \leq \epsilon, \text{ when } |x_i - x_j| \geq c\sqrt{\tau}\log(\epsilon^{-1}). \tag{10.3}$$

Thus, if the right-hand side of equation (10.2) has support in a subregion Ω_i, then the solution will decay rapidly with distance with a rate of decay given by (10.3).

Accordingly, let $\Omega_1, \ldots, \Omega_p$ denote a partition of Ω into p *nonoverlapping subregions*, and let $\hat{\Omega}_i \supset \Omega_i$ denote an extension of Ω_i containing all points in Ω within a distance of $c\sqrt{\tau}\log(\epsilon^{-1})$. Thus, $\hat{\Omega}_1, \ldots, \hat{\Omega}_p$ form an overlapping

covering of Ω, as in Schwarz algorithms. To approximately solve

$$(I + \tau A)\, u^{n+1} = g,$$

the right-hand side is first partitioned as

$$g = g_1 + \cdots + g_p, \quad \text{where support}(g_i) \subset \overline{\Omega}_i.$$

(Such a partition can be obtained, for instance, analogously to the construction in the proof of the partition lemma in Theorem 16.) Next, solve the following problem on each extended subdomain $\hat{\Omega}_i$:

$$L_{\hat{\Omega}_i} u_i = g_i, \quad \text{for } i = 1, \ldots, p,$$

where $L_{\hat{\Omega}_i} \equiv R_{\hat{\Omega}_i}(I + \tau A) R_{\hat{\Omega}_i}^T$ denotes the principal submatrix of $I + \tau A$ corresponding to the interior nodes on $\hat{\Omega}_i$. The approximate solution w^{n+1} is defined as

$$w^{n+1} \equiv u_1 + \cdots + u_p.$$

The following error bound is proved in Kuznetsov (1988).

Theorem 30 If the extended subdomains have overlap of size

$$\mathcal{O}(\sqrt{\tau} \log(\epsilon^{-1})),$$

the error satisfies

$$\|w^{n+1} - u^{n+1}\| \leq \mathcal{O}(\epsilon).$$

Thus, for instance, when the time step $\tau = h$ and $\epsilon = h^2$, the overlap should be approximately $\mathcal{O}(\sqrt{h} \log(h))$. Consequently, the extended subdomains must have a minimum overlap of the size prescribed above in order for the truncation error to be acceptable. This provides a constraint on the choice of subdomains. The case of convection diffusion problems is discussed in Kuznetsov (1990).

10.3. Alternative one iteration based approximations

An alternative algorithm that provides an approximate solution of (10.2) was proposed by Dryja (1991) and corresponds to a domain decomposed matrix splitting (fractional step method) involving two nonoverlapping subregions. The resulting scheme can be shown to be unconditionally stable. Unfortunately, the discretization error of the splitting scheme becomes the square root of the discretization error of the original scheme, see Dryja (1991) for the details. It is possible to recover the original discretization error by using an alternative splitting, see Laevsky (1992; 1993).

Kuznetsov (1988) proposed an explicit–implicit scheme to solve parabolic problems based on a partition of Ω into nonoverlapping regions. The boundary value of u^{n+1} on the interface B is first computed using an explicit

method (or even an implicit scheme) in a small neighbourhood of B. Using these boundary values, Dirichlet problems can be solved on each subdomain to provide the solution u^{n+1} on the whole domain Ω. This idea is particularly appealing on grids containing regions of refinement.

Another alternative approach was proposed by Dawson and Du (1991), Dawson, Du and Dupont (1991), in which the domain is partitioned into many nonoverlapping subdomains with interface B. Special basis function are constructed having support in a small 'tube' of width $\mathcal{O}(H)$ containing the interface B. In the first step approximate boundary values are computed on B using these special basis functions (involving some overhead cost). Finally, using these boundary values, the solution u^{n+1} is determined at the interior of the subdomains.

11. Mixed finite elements and the Stokes problem

In this section, we briefly describe some domain decomposition methods for solving the linear systems arising from mixed finite element discretizations of elliptic problems and discretizations of the steady Stokes equations (see Girault and Raviart (1986), Brezzi and Fortin (1991) for details on mixed finite element discretizations). Studies of domain decomposition methods for mixed finite element discretizations of elliptic problems were initiated by Glowinski and Wheeler (1988), while studies of domain decomposition for the Stokes problem were initiated by Lions (1988), Fortin and Aboulaich (1988), Bramble and Pasciak (1988) and Quarteroni (1989).

The mixed formulation of an elliptic problem: $-\nabla \cdot (a\nabla p) = f$ on Ω, with Neumann boundary conditions $\boldsymbol{n} \cdot a\nabla p = g$ on $\partial\Omega$ is given by

$$\begin{cases} a^{-1}\boldsymbol{u} + \nabla p = 0, & \text{in } \Omega, \quad \text{Darcy's law} \\ \nabla \cdot \boldsymbol{u} = f, & \text{in } \Omega, \quad \text{Conservation of mass} \\ \boldsymbol{n} \cdot \boldsymbol{u} = -g, & \text{in } \partial\Omega, \quad \text{Flux boundary condition} \end{cases}$$

where the compatibility condition

$$\int_\Omega f \, \mathrm{d}x + \int_{\partial\Omega} g \, \mathrm{d}s = 0$$

is assumed. The Stokes problem with Dirichlet boundary conditions for the velocity u is

$$\begin{cases} -\nu\Delta\boldsymbol{u} + \nabla p = f, & \text{in } \Omega, \\ \nabla \cdot \boldsymbol{u} = 0, & \text{in } \Omega, \\ \boldsymbol{u} = 0, & \text{on } \partial\Omega. \end{cases}$$

In both problems $\boldsymbol{u}$ refers to the velocity and p to the pressure.

After discretization, both these problems result in linear systems of the

following form:

$$\begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix}, \tag{11.1}$$

where u is the discrete velocity unknowns and p is the discrete pressure unknowns. Note that (11.1) is symmetric but indefinite and cannot be solved directly by the conjugate gradient method. Such systems are usually solved by block matrix and optimization based solution procedures. The square matrix A is symmetric and positive definite for both the Stokes and mixed case. In particular, A is block diagonal in the Stokes case, with diagonal blocks corresponding to discretization of the Laplacian. In the mixed elliptic case, A corresponds to a discretization of a^{-1}, the inverse of the coefficients a of the elliptic problem. The matrix B^T is rectangular and represents a discretization of the gradient, while its transpose B represents a discretization of the divergence operator. In many applications B^T has a null space spanned by $[1, \ldots, 1]^T$.

11.1. Methods based on elimination of the velocity

A simple procedure to solve (11.1) is to eliminate u and solve the reduced system for p:

$$Sp \equiv -BA^{-1}B^T p = g - BA^{-1}f,$$

after which u can be obtained by $u = A^{-1}(f - B^T p)$. Note that the Schur complement S is negative definite and hence a conjugate gradient type method can be used. Each matrix–vector product with S can be computed at the cost of solving a linear system with coefficient matrix A.

For the Stokes problem, it can be shown that S is well conditioned and requires no preconditioning. However, the matrix A is block diagonal with diagonal blocks corresponding to the Laplacian, and domain decomposition preconditioners can be applied to A. We refer the reader to Bramble and Pasciak (1988) for the details, where the Stokes problem is reformulated as a positive definite linear system and, additionally, a nonoverlapping domain decomposition algorithm is described. See Rusten and Winther (1992) and Rusten (1991) for an interesting algorithm for preconditioning the entire system without eliminating either u or p.

For the mixed elliptic case, the operator S is not well conditioned, and the above elimination method is not as attractive, see Wheeler and Gonzalez (1984). However, if a *dual* formulation of the mixed problem is used, see Arnold and Brezzi (1985), then the resulting Schur complement for the pressure becomes a nonconforming finite element discretization of the corresponding elliptic problem for the pressure. Efficient domain decomposition preconditioners have been proposed for such nonconforming discretizations (corresponding to the Schur complement S in the *dual* formulation), see

Cowsar, Mandel and Wheeler (1993), Cowsar (1993), Sarkis (1993) and Meddahi (1993).

11.2. Methods based on divergence free velocities

An alternative to algorithms based on elimination of the velocity are those in which the pressure is implicitly eliminated. These methods are based on the observation that the pressure corresponds to a Lagrange multiplier in the following constrained minimization problem:

$$\min \tfrac{1}{2}u^T Au - u^T f, \quad \text{subject to} \quad Bu = g,$$

see Girault and Raviart (1986), Lions (1988), Glowinski and Wheeler (1988) and Quarteroni (1989). In particular, if the divergence constraint $Bu = g$ can be reduced to $Bu = 0$, (i.e. the feasible set of velocities corresponds to a linear subspace of divergence free velocities), then in this subspace the problem becomes positive definite because

$$\begin{bmatrix} u \\ p \end{bmatrix}^T \begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix} = u^T Au + 2p^T Bu = u^T Au > 0. \tag{11.2}$$

This positive definiteness provides the basis for applying standard conjugate gradient methods to determine the minimum velocity solution within the feasible set of velocities satisfying the divergence constraint.

Based on the space of divergence free velocities, Glowinski and Wheeler (1988) proposed a nonoverlapping domain decomposition method for mixed finite element discretizations of elliptic problems. In subsequent articles, Glowinski, Kinton and Wheeler (1990a) and Cowsar and Wheeler (1991) proposed improved preconditioners for the corresponding Schur complement system. Nonoverlapping algorithms for the Stokes problem were proposed by Bramble and Pasciak (1988) and Quarteroni (1989).

Schwarz alternating algorithms for the Stokes problem were proposed in Lions (1988) and Fortin and Aboulaich (1988), see also Pahl (1993), and are based on implicit elimination of the pressure. They were extended to the case of mixed finite element discretizations of elliptic problems in Mathew (1989; 1993a) and Ewing and Wang (1991). In the following, we briefly describe the basic linear algebraic issues for formulating Schwarz algorithms in the mixed case.

Two issues need to be addressed in order to define a Schwarz method involving subproblems of the form:

$$\begin{bmatrix} * & * & * & * \\ * & A_i & B_i^T & * \\ * & B_i & 0 & * \\ * & * & * & * \end{bmatrix} \begin{bmatrix} * \\ u_i \\ p_i \\ * \end{bmatrix} = \begin{bmatrix} * \\ W_i \\ F_i \\ * \end{bmatrix}, \tag{11.3}$$

after some suitable reordering of (11.1). They are:

1 The submatrix B_i may be singular, depending on the boundary conditions, due to the nonuniqueness of the pressure. In such a case, F_i must have mean value zero if $[1, \ldots, 1]^T$ spans the null space of B_i (as is often the case). This corresponds to the compatibility condition for solvability of the subproblem.
2 When B_i is singular, due to the nonuniqueness of the local pressure solution p_i, its mean value on the subregion is arbitrary and should be suitably prescribed in order to compute a globally defined pressure p_h.

The first difficulty can be handled by reducing the problem to one involving divergence free velocities. The second difficulty can be treated by sequentially modifying the local pressure solutions so that they have the same mean value with adjacent pressures on the regions of overlap. The algorithm can now be outlined:

1 Determine a velocity u^* satisfying $Bu^* = g$. Then, the correction $\tilde{u} = u - u^*$ to the velocity satisfies: $B\tilde{u} = 0$, and all subsequent local subproblems will be compatible with the zero flux data.
2 Next, apply the Schwarz methods to compute the divergence free velocity $\tilde{u}$ by solving local problems which have the same form as (11.1) and which involves local velocities and pressures.
3 Finally, determine a global pressure using the local pressures determined in Step 2.

We refer the reader to Ewing and Wang (1991) and Mathew (1993a) for the details.

For suitable choices of overlapping subregions and a coarse mesh, as in Section 2, the convergence rates of the additive and multiplicative Schwarz algorithms in the mixed elliptic case has been shown to be independent of H and h, see Ewing and Wang (1991) and Mathew (1993b).

12. Other topics

In this section, we provide some references to several domain decomposition procedures that we do not have space to discuss in any details. A good source of references is the set of conference proceedings mentioned in the introduction.

12.1. Biharmonic problem

For conforming finite element discretizations of the biharmonic problem based on *Hermite* finite elements, the additive and multiplicative Schwarz algorithms as well as the multilevel Schwarz algorithms have been developed with optimal convergence rates, see Zhang (1991; 1992a,c). See also Scapini (1990). Nonoverlapping domain decomposition algorithms based on

Hermite elements are discussed in Sun and Zou (1991), Hoffmann and Zou (1992), and Scapini (1991).

Algorithms for finite difference discretizations of the biharmonic equation (such as the 13-point stencil) pose additional difficulties. Indeed, if a nonoverlapping decomposition is used for such discretizations, the *interface* must consist of *two lines* in order to decouple the local subproblems. This requires modifications in the usual construction of interface preconditioners for the Schur complement system. We refer the reader to Tsui (1991) and Chan, Weinan and Sun (1991b). Thus far, to the knowledge of the authors, the Schwarz algorithms have not been studied in the case of finite difference discretizations of the biharmonic problem.

12.2. Spectral, spectral element and p version finite elements

For a general discussion on domain decomposition for spectral methods, we refer the reader to Canuto *et al.* (1988), and for a discussion on spectral element methods to Bernardi, Maday and Patera (1989), Maday and Patera (1989), Bernardi, Debit and Maday (1989), Bernardi and Maday (1992) and Fischer and Rønquist (1993).

The Schwarz algorithm for spectral methods was proposed in Morchoisne (1984), Canuto and Funaro (1988). More recently, Dirichlet–Neumann-type domain decomposition algorithms were proposed by Funaro, Quarteroni and Zanolli (1988). For boundary layer and elliptic–hyperbolic problems, spectral methods are described in Gastaldi *et al.* (1990). Applications and techniques of pseudospectral domain decomposition methods in fluid dynamics are described by Phillips (1992).

The earliest domain decomposition algorithm for p version finite elements was proposed by Babuška, Craig, Mandel and Pitkäranta (1991), in two dimensions. Since then, algorithms similar to the Neumann–Neumann, wirebasket and Schwarz methods have been developed for p version finite elements having almost optimal convergence rates (polylogarithmic growth in p). We refer the reader to Mandel (1989; 1990), for Neumann–Neumann and wirebasket-type algorithms, and to Pavarino (1992, 1993a,b) and Pavarino and Widlund (1993) for Schwarz, local refinement and wirebasket type algorithms for p version finite elements.

12.3. Indefinite Helmholtz problems

The solution of the indefinite Helmholtz problem

$$-\Delta u - k^2 u = f$$

is a difficult problem for large k (by domain decomposition or other methods). For a discussion of Schwarz algorithms for solving indefinite problems, we refer the reader to Cai and Widlund (1992); the convergence rate

depends on the size of the coarse grid used. Nonoverlapping domain decomposition algorithms were recently proposed by Despres (1991), and by Ernst and Golub (1992) (for the complex Helmholtz equation). Alternative approaches based on fictitious domains are described in Proskurowski and Vassilevski (1992) and Atamian *et al.* (1991).

12.4. Nonconforming finite elements

Domain decomposition algorithms (cf Neumann–Neumann and Schwarz) have been developed for solving nonconforming finite element discretizations of elliptic problems, such as the Crouzeix–Raviart elements and dual mixed finite element discretizations, see Arnold and Brezzi (1985). Sarkis (1993) proposed several extensions of the Neumann–Neumann algorithm to the nonconforming case. In the context of dual formulations, related algorithms were independently proposed by Cowsar *et al.* (1993). Versions of the Schwarz algorithm were proposed by Cowsar (1993) and Meddahi (1993).

Acknowledgements The authors wish to thank Olof Widlund for many helpful discussions on various topics presented in this article. We have greatly benefitted from comments and helpful suggestions on an earlier draft of this article by Patrick Ciarlet, Max Dryja, Patrick Le Tallec, Barry Smith, Olof Widlund, Jinchao Xu and Jun Zou. Our sincere thanks to all of them.

REFERENCES

V.I. Agoshkov (1988), 'Poincaré–Steklov operators and domain decomposition methods in finite dimensional spaces', in *First Int. Symp. Domain Decomposition Methods for Partial Differential Equations* (R. Glowinski, G.H. Golub, G.A. Meurant and J. Périaux, eds), SIAM (Philadelphia, PA).

V.I. Agoshkov and V.I. Lebedev (1985), 'The Poincaré–Steklov operators and the domain decomposition methods in variational problems', in *Computational Processes and Systems* Nauka (Moscow) 173–227. In Russian.

D.N. Arnold and F. Brezzi (1985), 'Mixed and nonconforming finite element methods: Implementation, post processing and error estimates', *Math. Model. Numer. Anal.* **19**, 7–32.

S.F. Ashby, P.E. Saylor and J.S. Scroggs (1992), 'Physically motivated domain decomposition preconditioners', in *Proc. Second Copper Mountain Conf. on Iterative Methods*, Vol. 1, Comput. Math. Group, University of Colorado at Denver.

G.P. Astrakhantsev (1978), 'Method of fictitious domains for a second-order elliptic equation with natural boundary conditions', *USSR Comput. Math. Math. Phys.* **18**, 114–121.

C. Atamian, Q.V. Dinh, R. Glowinski, J. He and J. Périaux (1991), 'Control approach to fictitious-domain methods application to fluid dynamics and electromagnetics', in *Fourth Int. Symp. Domain Decomposition Methods for Partial Differential Equations* (R. Glowinski, Y.A. Kuznetsov, G.A. Meurant, J. Périaux and O. Widlund, eds), SIAM (Philadelphia, PA).

O. Axelsson and P. Vassilevski (1990), 'Algebraic multilevel preconditioning methods, II', *SIAM J. Numer. Anal.* **27**, 1569–1590.

I. Babuška (1957), 'Über Schwarzsche Algorithmen in partielle Differentialgleichungen der mathematischen Physik', *ZAMM* **37**(7/8), 243–245.

I. Babuška, A. Craig, J. Mandel and J. Pitkäranta (1991), 'Efficient preconditioning for the p-version finite element method in two dimensions', *SIAM J. Numer. Anal.* **28**(3), 624–661.

R.E. Bank, T.F. Dupont and H. Yserentant (1988), 'The hierarchical basis multigrid method', *Numer. Math.* **52**, 427–458.

M. Berger and S. Bokhari (1987), 'A partitioning strategy for nonuniform problems on multiprocessors', *IEEE Trans. Comput.* **36**, 570–580.

C. Bernardi and Y. Maday (1992), 'Approximations spectral de problèmes aux limites elliptiques', in *Mathematiques et Applications*, Vol. 10, Springer (Paris).

C. Bernardi, Y. Maday and A. Patera (1989), 'A new nonconforming approach to domain decomposition: the mortar element method', *Nonlinear Partial Differential Equations and their Applications* (H. Brezis and J.L. Lions, eds), Pitman (London).

C. Bernardi, N. Debit and Y. Maday (1990), 'Coupling finite element and spectral methods', *Math. Comput.* **54**, 21–39.

P.E. Bjørstad and M. Skogen (1992), 'Domain decomposition algorithms of Schwarz type, designed for massively parallel computers', *Fifth Int. Symp. Domain Decomposition Methods for Partial Differential Equations* (T.F. Chan, D.E. Keyes, G.A. Meurant, J.S. Scroggs and R.G. Voigt, eds), SIAM (Philadelphia, PA).

P.E. Bjørstad and O.B. Widlund (1984), 'Solving elliptic problems on regions partitioned into substructures', *Elliptic Problem Solvers II* (G. Birkhoff and A. Schoenstadt, eds), Academic (London), 245–256.

P.E. Bjørstad and O.B. Widlund (1986), 'Iterative methods for the solution of elliptic problems on regions partitioned into substructures', *SIAM J. Numer. Anal.* **23**(6), 1093–1120.

P.E. Bjørstad and O.B. Widlund (1989), 'To overlap or not to overlap: A note on a domain decomposition method for elliptic problems', *SIAM J. Sci. Stat. Comput.* **10**(5), 1053–1061.

C. Börgers (1989), 'The Neumann–Dirichlet domain decomposition method with inexact solvers on the subdomains', *Numer. Math.* **55**, 123–136.

C. Börgers and O.B. Widlund (1990), 'On finite element domain imbedding methods', *SIAM J. Numer. Anal.* **27**(4), 963–978.

F. Bornemann and H. Yserentant (1993), 'A basic norm equivalence for the theory of multilevel methods', *Num. Math.* **64**, 455–476.

J.-F. Bourgat, R. Glowinski, P. Le Tallec and M. Vidrascu (1989), 'Variational formulation and algorithm for trace operator in domain decomposition calculations', in *Second Int. Conf. on Domain Decomposition Methods* (T. Chan, R. Glowinski, J. Périaux and O. Widlund, eds), SIAM (Philadelphia, PA).

J.H. Bramble and J.E. Pasciak (1988), 'A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems', *Math. Comput.* **50**, 1–18.

J.H. Bramble and J. Xu (1991), 'Some estimates for a weighted L^2 projection', *Math. Comput.* **56**, 463–476.

J.H. Bramble, R.E. Ewing, J.E. Pasciak and A.H. Schatz (1988), 'A preconditioning technique for the efficient solution of problems with local grid refinement', *Comput. Meth. Appl. Mech. Engrg* **67**, 149–159.

J.H. Bramble, R.E. Ewing, R.R. Parashkevov and J.E. Pasciak (1992), 'Domain decomposition methods for problems with partial refinement', *SIAM J. Sci. Comput.* **13(1)**, 397–410.

J.H. Bramble, J.E. Pasciak and A.H. Schatz (1986a), 'The construction of preconditioners for elliptic problems by substructuring, I', *Math. Comput.* **47**, 103–134.

J.H. Bramble, J.E. Pasciak and A.H. Schatz (1986b), 'An iterative method for elliptic problems on regions partitioned into substructures', *Math. Comput.* **46**(173), 361–369.

J.H. Bramble, J.E. Pasciak and A.H. Schatz (1989), 'The construction of preconditioners for elliptic problems by substructuring, IV', *Math. Comput.* **53**, 1–24.

J.H. Bramble, J.E. Pasciak and J. Xu (1990), 'Parallel multilevel preconditioners', *Math. Comput.* **55**, 1–22.

J.H. Bramble, J.E. Pasciak, J. Wang and J. Xu (1991), 'Convergence estimates for product iterative methods with applications to domain decomposition', *Math. Comput.* **57**(195), 1–21.

S.C. Brenner (1993), 'Two-level additive Schwarz preconditioners for nonconforming finite element methods', in *Proc. Seventh Int. Symp. on Domain Decomposition Methods for Partial Differential Equations*, to appear.

F. Brezzi and M. Fortin (1991), *Mixed and Hybrid Finite Element Methods*, Springer (Berlin).

B.L. Buzbee, F. Dorr, J. George and G. Golub (1971), 'The direct solution of the discrete poisson equation on irregular regions', *SIAM J. Numer. Anal.* **11**, 722–736.

X.-C. Cai (1991), 'Additive Schwarz algorithms for parabolic convection–diffusion equations', *Numer. Math.* **60**(1), 41–61.

X.-C. Cai (1993), 'Multiplicative Schwarz methods for parabolic problems', *SIAM J. Sci. Comput.*, to appear.

X.-C. Cai and O. Widlund (1992), 'Domain decomposition algorithms for indefinite elliptic problems', *SIAM J. Sci. Comput.* **13**(1), 243–258.

X.-C. Cai and O. Widlund (1993), 'Multiplicative Schwarz algorithms for some nonsymmetric and indefinite problems', *SIAM J. Numer. Anal.* **30**(4), 936–952.

X.-C. Cai, W.D. Gropp and D.E. Keyes (1992), 'A comparison of some domain decomposition algorithms for nonsymmetric elliptic problems', in *Fifth Int. Symp. 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), SIAM (Philadelphia, PA).

C. Canuto and D. Funaro (1988), 'The Schwarz algorithm for spectral methods', *SIAM J. Numer. Anal.* **25**(1), 24–40.

C. Canuto, M.Y. Hussaini, A. Quarteroni and T.A. Zang (1988), *Spectral Methods in Fluid Dynamics*, Springer (Berlin).

T.F. Chan (1987), 'Analysis of preconditioners for domain decomposition', *SIAM J. Numer. Anal.* **24**(2), 382–390.

T.F. Chan and J. Shao (1993), 'Optimal coarse grid size in domain decomposition', Technical Report 93-24, UCLA CAM Report, Los Angeles, CA 90024-1555.

T.F. Chan and D. Goovaerts (1992), 'On the relationship between overlapping and nonoverlapping domain decomposition methods', *SIAM J. Matrix Anal. Appl.* **13**(2), 663.

T.F. Chan and T.Y. Hou (1991), 'Eigendecompositions of domain decomposition interface operators for constant coefficient elliptic problems', *SIAM J. Sci. Comput.* **12**(6), 1471–1479.

T.F. Chan and D.F. Keyes (1990), 'Interface preconditioning for domain decomposed convection diffusion operators', in *Third Int. Symp. on Domain Decomposition Methods for Partial Differential Equations* (T. Chan, R. Glowinski, J. Périaux and O. Widlund, eds), SIAM (Philadelphia, PA).

T.F. Chan and T.P. Mathew (1992), 'The interface probing technique in domain decomposition', *SIAM J. Matrix Anal. Appl.* **13**(1), 212–238.

T.F. Chan and T. Mathew (1993), 'Domain decomposition preconditioners for convection diffusion problems', in *Domain Decomposition Methods for Partial Differential Equations* (A. Quarteroni, ed.), American Mathematical Society (Providence, RI), to appear.

T.F. Chan and D.C. Resasco (1985), 'A survey of preconditioners for domain decomposition', Technical Report /DCS/RR-414, Yale University.

T.F. Chan and D.C. Resasco (1987), 'Analysis of domain decomposition preconditioners on irregular regions', in *Advances in Computer Methods for Partial Differential Equations - VI* (R. Vichnevetsky and R. Stepleman, eds), IMACS, 317–322.

T.F. Chan, T.Y. Hou and P.L. Lions (1991a), 'Geometry related convergence results for domain decomposition algorithms', *SIAM J. Numer. Anal.* **28**(2), 378.

T.F. Chan, E. Weinan and J. Sun (1991b), 'Domain decomposition interface preconditioners for fourth order elliptic problems', *Appl. Numer. Math.* **8**, 317–331.

T.F. Chan, D.E. Keyes, G.A. Meurant, J.S. Scroggs and R.G. Voigt, eds (1992a), *Fifth Conf. on Domain Decomposition Methods for Partial Differential Equations*, SIAM (Philadelphia, PA).

T.F. Chan, T.P. Mathew and J.-P. Shao (1992b), 'Efficient variants of the vertex space domain decomposition algorithm', Technical Report CAM 92-07, Department of Mathematics, UCLA, to appear in *SIAM J. Sci. Comput.*

T.F. Chan, R. Glowinski, J. Périaux and O. Widlund, eds (1989), *Second Int. Conf. on Domain Decomposition Methods*, SIAM (Philadelphia, PA).

T. Chan, R. Glowinski, J. Périaux and O. Widlund, eds (1990), *Third Int. Symp. on Domain Decomposition Methods for Partial Differential Equations*, SIAM (Philadelphia, PA).

R.C.Y. Chin, G.W. Hedstrom, J.R. McGraw and F.A. Howes (1986), 'Parallel computation of multiple scale problems', in *New Computing Environments: Parallel, Vector and Systolic* (A. Wouk, ed.) SIAM (Philadelphia, PA).

R. Courant and D. Hilbert (1962), *Methods of Mathematical Physics*, Vol. 2, Interscience (New York).

L.C. Cowsar (1993), 'Dual variable Schwarz methods for mixed finite elements', Technical Report TR93-09, Department of Mathematical Sciences, Rice University.

L.C. Cowsar and M.F. Wheeler (1991), 'Parallel domain decomposition method for mixed finite elements for elliptic partial differential equations', in *Fourth Int. Symp. on Domain Decomposition Methods for Partial Differential Equations* (R. Glowinski, Y.A. Kuznetsov, G.A. Meurant, J. Périaux and O. Widlund, eds), SIAM (Philadelphia, PA).

L.C. Cowsar, J. Mandel and M.F. Wheeler (1993), 'Balancing domain decomposition for mixed finite elements', Technical Report TR93-08, Department of Mathematical Sciences, Rice University.

C.N. Dawson and Q. Du (1991), 'A domain decomposition method for parabolic equations based on finite elements', in *Fourth Int. Symp. on Domain Decomposition Methods for Partial Differential Equations* (R. Glowinski, Y.A. Kuznetsov G.A. Meurant, J. Périaux and O. Widlund, eds), SIAM (Philadelphia, PA).

C. Dawson, Q. Du and T.F. Dupont, (1991), 'A finite difference domain decomposition algorithm for numerical solution of the heat equation', *Math. Comput.* **57**, 195.

Y.-H. De Roeck (1989), 'A local preconditioner in a domain-decomposed method', Technical Report TR89/10, Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique, Toulouse, France.

Y.-H. De Roeck and P. Le Tallec (1991), 'Analysis and test of a local domain decomposition preconditioner', in *Fourth Int. Symp. on Domain Decomposition Methods for Partial Differential Equations* (R. Glowinski, Y. Kuznetsov, G. Meurant, J. Périaux and O. Widlund, eds), SIAM (Philadelphia, PA).

B. Despres (1991), 'Methodes de decomposition de domaines pour les problemes de propagation d'ondes en regime harmoniques', PhD thesis, University of Paris IX, Dauphine.

Q. Dinh, R. Glowinski and J. Périaux (1984), 'Solving elliptic problems by domain decomposition methods with applications', in *Elliptic Problem Solvers II* (G. Birkhoff and A. Schoenstadt, eds), Academic (New York), 395–426.

M. Dryja (1982), 'A capacitance matrix method for Dirichlet problem on polygon region', *Numer. Math.* **39**, 51–64.

M. Dryja (1984), 'A finite element-capacitance method for elliptic problems on regions partitioned into subregions', *Numer. Math.* **44**, 153–168.

M. Dryja (1988), 'A method of domain decomposition for 3-D finite element problems', in *First Int. Symp. on Domain Decomposition Methods for Partial Differential Equations* (R. Glowinski, G.H. Golub, G.A. Meurant and J. Périaux, eds) SIAM (Philadelphia, PA).

M. Dryja (1989), 'An additive Schwarz algorithm for two- and three-dimensional finite element elliptic problems', in *Second Int. Conf. on Domain Decomposition Methods* (T. Chan, R. Glowinski, J. Périaux and O. Widlund, eds), SIAM (Philadelphia, PA).

M. Dryja (1991), 'Substructuring methods for parabolic problems', in *Fourth Int. Symp. on Domain Decomposition Methods for Partial Differential Equations*

(R. Glowinski, Y.A. Kuznetsov, G.A. Meurant, J. Périaux and O. Widlund, eds), SIAM (Philadelphia, PA).

M. Dryja and O.B. Widlund (1987), '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.

M. Dryja and O.B. Widlund (1989a), 'On the optimality of an additive iterative refinement method', in *Proc. Fourth Copper Mountain Conf. on Multigrid Methods*, SIAM (Philadelphia, PA) 161–170.

M. Dryja and O.B. Widlund (1989b), 'Some domain decomposition algorithms for elliptic problems', in *Iterative Methods for Large Linear Systems* (L. Hayes and D. Kincaid, eds), Academic (San Diego, CA), 273–291.

M. Dryja and O.B. Widlund (1990), 'Towards a unified theory of domain decomposition algorithms for elliptic problems', in *Third Int. Symp. on Domain Decomposition Methods for Partial Differential Equations* (T. Chan, R. Glowinski, J. Périaux and O. Widlund, eds), SIAM (Philadelphia, PA).

M. Dryja and O.B. Widlund (1992a), 'Additive Schwarz methods for elliptic finite element problems in three dimensions', in *Fifth Conf. 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), SIAM (Philadelphia, PA).

M. Dryja and O.B. Widlund (1992b), 'Domain decomposition algorithms with small overlap', Technical Report 606, Department of Computer Science, Courant Institute, to appear in *SIAM J. Sci. Comput.*

M. Dryja and O.B. Widlund (1993a), 'Schwarz methods of Neumann-Neumann type for three-dimensional elliptic finite element problems', Technical Report 626, Department of Computer Science, Courant Institute.

M. Dryja and O.B. Widlund (1993b), 'Some recent results on Schwarz type domain decomposition algorithms', in *Sixth Conf. on Domain Decomposition Methods for Partial Differential Equations* (A. Quarteroni, ed.), American Mathematical Society (Providence, RI), to appear. Technical Report 615, Department of Computer Science, Courant Institute.

M. Dryja, B.F. Smith and O.B. Widlund (1993), 'Schwarz analysis of iterative substructuring algorithms for elliptic problems in three dimensions', Technical Report 638, Department of Computer Science, Courant Institute. *SIAM J. Numer. Anal.*, submitted.

O. Ernst and G. Golub (1992), 'A domain decomposition approach to solving the helmholtz equation with a radiation boundary condition', Technical Report 92-08, Stanford University, Computer Science Department, Numerical Analysis Project, Stanford, CA 94305.

R.E. Ewing and J. Wang (1991), 'Analysis of the Schwarz algorithm for mixed finite element methods', *RAIRO, Math. Modell. Numer. Anal.* **26**(6), 739–756.

C. Farhat and M. Lesoinne (1993), 'Automatic partitioning of unstructured meshes for the parallel solution of problems in computational mechanics', *Int. J. Numer. Meth. Engrg* **36**, 745–764.

C. Farhat and F.X. Roux (1992), 'An unconventional domain decomposition method for an efficient parallel solution of large scale finite element systems', *SIAM J. Sci. Comput.* **13**, 379–396.

S.A. Finogenov and Y.A. Kuznetsov (1988), 'Two-stage fictitious components method for solving the Dirichlet boundary value problem', *Sov. J. Numer. Anal. Math. Modell.* **3**(4), 301–323.

P.F. Fischer and E.M. Rønquist (1993), 'Spectral element methods for large scale parallel Navier–Stokes calculations', in *Second Int. Conf. on Spectral and High Order Methods for PDE's*. Proc. ICOSAHOM 92, a conference held in Montpellier, France, June 1992. To appear in *Comput. Meth. Appl. Mech. Engrg.*

M. Fortin and R. Aboulaich (1988), 'Schwarz's decomposition method for incompressible flow problems', in *First Int. Symp. on Domain Decomposition Methods for Partial Differential Equations* (R. Glowinski, G.H. Golub, G.A. Meurant and J. Périaux, eds), SIAM (Philadelphia, PA).

G. Fox (1988), 'A review of automatic load balancing and decomposition methods for the hypercube', in *Numerical Algorithms for Modern Parallel Computers* (M. Schultz, ed.), Springer (Berlin), 63–76.

R.W. Freund, G.H. Golub and N. Nachtigal (1992), 'Iterative solution of linear systems', *Acta Numerica*, Cambridge University Press (Cambridge), 57–100.

D. Funaro, A. Quarteroni and P. Zanolli (1988), 'An iterative procedure with interface relaxation for domain decomposition methods', *SIAM J. Numer. Anal.* **25**, 1213–1236.

M. Garbey (1992), 'Domain decomposition to solve layers and singular perturbation problems', in *Fifth Conf. 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), SIAM (Philadelphia, PA).

F. Gastaldi, A. Quarteroni and G. Sacchi-Landriani (1990), 'On the coupling of two-dimensional hyperbolic and elliptic equations: analytical and numerical approach', in *Domain Decomposition Methods for Partial Differential Equations* (T. Chan, R. Glowinski, J. Périaux and O. Widlund, eds), SIAM (Philadelphia, PA) 23–63.

V. Girault and P.-A. Raviart (1986), *Finite Element Approximation of the Navier–Stokes Equations: Theory and Algorithms*, Springer (Berlin).

R. Glowinski and M.F. Wheeler (1988), 'Domain decomposition and mixed finite element methods for elliptic problems', in *First Int. Symp. on Domain Decomposition Methods for Partial Differential Equations* (R. Glowinski, G.H. Golub, G.A. Meurant and J. Périaux, eds), SIAM (Philadelphia, PA).

R. Glowinski, G.H. Golub, G.A. Meurant and J. Périaux, eds (1988), *Proc. First Int. Symp. on Domain Decomposition Methods for Partial Differential Equations*, SIAM (Philadelphia, PA).

R. Glowinski, W. Kinton and M.F. Wheeler (1990a), 'Acceleration of domain decomposition algorithms for mixed finite elements by multi-level methods', in *Third Int. Symp. on Domain Decomposition Methods for Partial Differential Equations* (T. Chan, R. Glowinski, J. Périaux and O. Widlund, eds), SIAM (Philadelphia, PA).

R. Glowinski, Y.A. Kuznetsov, G.A. Meurant, J. Périaux and O. Widlund, eds (1991), *Fourth Int. Symp. on Domain Decomposition Methods for Partial Differential Equations*, SIAM (Philadelphia, PA).

R. Glowinski, J. Périaux and G. Terrasson (1990b), 'On the coupling of viscous and inviscid models for compressible fluid flows via domain decomposition', in *Domain Decomposition Methods for Partial Differential Equations* (T. Chan, R. Glowinski, J. Périaux and O. Widlund, eds), SIAM (Philadelphia, PA).

G. Golub and D. Mayers (1984), 'The use of preconditioning over irregular regions', in *Computing Methods in Applied Sciences and Engineering, VI* (R. Glowinski and J.L. Lions, eds), North-Holland (Amsterdam, New York, Oxford) 3–14.

D. Goovaerts (1989), 'Domain decomposition methods for elliptic partial differential equations', PhD thesis, Department of Computer Science, Catholic University of Leuven.

D. Goovaerts, T. Chan and R. Piessens (1991), 'The eigenvalue spectrum of domain decomposed preconditioners', *Appl. Numer. Math.* **8**, 389–410.

M. Griebel (1991), 'Multilevel algorithms considered as iterative methods on indefinite systems', Inst. für Informatik, Tech. Univ. Munchen, SFB 342/29/91A.

M. Griebel and P. Oswald (1993), 'Remarks on the abstract theory of additive and multiplicative Schwarz algorithms', Inst. für Informatik, Tech. Univ. Munchen, SFB 342/6/93A.

W.D. Gropp (1992), 'Parallel computing and domain decomposition', in *Fifth Conf. 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), SIAM (Philadelphia, PA).

W. Gropp and D. Keyes (1993), 'Domain decomposition as a mechanism for using asymptotic methods', in *Asymptotic and Numerical Methods for Partial Differential Equations with Critical Parameters* (H. Kaper and M. Garbey, eds), Vol. 384, NATO ASI Series C, 93–106.

W. Gropp and B. F. Smith (1992), 'Experiences with domain decomposition in three dimensions: overlapping Schwarz methods', Mathematics and Computer Science Division, Argonne National Laboratory, to appear in *Proc. Sixth Int. Symp. on Domain Decomposition Methods.*

W. Hackbusch (1993), *Iterative Methods for Large Sparse Linear Systems*, Springer (Heidelberg).

G.W. Hedstrom and F.A. Howes (1990), 'Domain decomposition for a boundary value problem with a shock layer', in *Third Int. Symp. on Domain Decomposition Methods for Partial Differential Equations* (T.F. Chan, R.Glowinski, J. Périaux and O. Widlund, eds), SIAM (Philadelphia, PA).

F. D'Hennezel (1992), 'Domain decomposition method with non-symmetric interface operator', *Fifth Conf. on Domain Decomposition Methods for Partial Differential Equations* (Tony F. Chan, David E. Keyes, Gérard A. Meurant, Jeffrey S. Scroggs and Robert G. Voigt, eds), SIAM (Philadelphia, PA).

K.-H. Hoffmann and J. Zou (1992), 'Solution of biharmonic problems by the domain decomposition method', Technical Report No. 387, DFG-SPP, Technical University of Munich.

Kang, L. S. (1987), *Parallel Algorithms and Domain Decomposition*, Wuhan University Press (China). In Chinese.

B. Kernighan and S. Lin (1970), 'An efficient heuristic procedure for partitioning graphs', *Bell Systems Tech. J.* **29**, 291–307.

D.E. Keyes and W.D. Gropp (1987), 'A comparison of domain decomposition techniques for elliptic partial differential equations and their parallel implementation', *SIAM J. Sci. Comput.* **8**(2), s166–s202.

D.E. Keyes and W.D. Gropp (1989), 'Domain decomposition with local mesh refinement', Technical Report YALEU/DCS/RR-726, Yale University.

G. Kron (1953), 'A set of principles to interconnect the solutions of physical systems', *J. Appl. Phys.* **24**(8), 965.

Y.A. Kuznetsov (1988), 'New algorithms for approximate realization of implicit difference schemes', *Sov. J. Numer. Anal. Math. Modell.* **3**, 99–114.

Y.A. Kuznetsov (1990), 'Domain decomposition methods for unsteady convection–diffusion problems', in *IXth Int. Conf. on Computing Methods in Applied Science and Engineering* (R. Glowinski and J.L. Lions, eds), INRIA (Paris) 327–344.

Y.A. Kuznetsov (1991), 'Overlapping domain decomposition methods for fe-problems with elliptic singular perturbed operators', in *Fourth Int. Symp. on Domain Decomposition Methods for Partial Differential Equations* (R. Glowinski, Y.A. Kuznetsov, G.A. Meurant, J. Périaux and O. Widlund, eds), SIAM (Philadelphia, PA).

Yu. M. Laevsky (1992), 'Direct domain decomposition method for solving parabolic equations', Preprint no. 940, Novosibirsk, Computing Center Siberian Branch Academy of Sciences. In Russian.

Yu. M. Laevsky (1993), 'On the domain decomposition method for parabolic problems', *Bull. Novosibirsk Comput. Center* **1**, 41–62.

P. Le Tallec (1994), 'Domain decomposition methods in computational mechanics', *J. Comput. Mech. Adv.*, to appear.

P. Le Tallec, Y.-H. De Roeck and M. Vidrascu (1991), 'Domain-decomposition methods for large linearly elliptic three-dimensional problems', *J. Comput. Appl. Math.* **34**.

V.I. Lebedev (1986), *Composition Methods*, USSR Academy of Sciences, Moscow. In Russian.

P.L. Lions (1988), 'On the Schwarz alternating method. I.', in *First Int. Symp. on Domain Decomposition Methods for Partial Differential Equations* (R. Glowinski, G.H. Golub, G.A. Meurant and J. Périaux, eds), SIAM (Philadelphia, PA).

P.L. Lions (1989), 'On the Schwarz alternating method. II', in *Second Int. Conf. on Domain Decomposition Methods* (T. Chan, R. Glowinski, J. Périaux and O. Widlund, eds), SIAM (Philadelphia, PA).

J.-L. Lions and E. Magenes (1972), *Nonhomogeneous Boundary Value Problems and Applications*, Vol. I, Springer (New York, Heidelberg, Berlin).

Lu Tao, T. Shih and C. Liem (1992), *Domain Decomposition Methods: New Numerical Techniques for Solving PDE*, Science Publishers (Beijing, China).

Y. Maday and A.T. Patera (1989), 'Spectral element methods for the Navier–Stokes equations', in *State of the Art Surveys in Computational Mechanics* (A.K. Noor and J.T. Oden, eds), ASME (New York).

J. Mandel (1989a), 'Efficient domain decomposition preconditioning for the p-version finite element method in three dimensions', Technical Report, Computational Mathematics Group, University of Colorado at Denver.

J. Mandel (1989b), 'Two-level domain decomposition preconditioning for the p-version finite element version in three dimensions', *Int. J. Numer. Meth. Engrg*, **29**, 1095–1108.

J. Mandel (1990), 'Hierarchical preconditioning and partial orthogonalization for the p-version finite element method', in *Third Int. Symp. on Domain Decomposition Methods for Partial Differential Equations* (T.F. Chan, R. Glowinski, J. Périaux and O. Widlund, eds), SIAM (Philadelphia, PA).

J. Mandel (1992), 'Balancing domain decomposition', *Commun. Numer. Meth. Engrg* **9**, 233–241.

J. Mandel and M. Brezina (1992), 'Balancing domain decomposition: Theory and computations in two and three dimensions', Technical Report, Computational Mathematics Group, University of Colorado at Denver.

J. Mandel and S. McCormick (1989), 'Iterative solution of elliptic equations with refinement: The model multi-level case', in *Second Int. Conf. on Domain Decomposition Methods* (T. Chan, R. Glowinski, J. Périaux and O. Widlund, eds), SIAM (Philadelphia, PA).

G.I. Marchuk, Y.A. Kuznetsov and A.M. Matsokin (1986), 'Fictitious domain and domain decomposition methods', *Sov. J. Numer. Anal. Math. Modell.* **1**, 3–61.

L.D. Marini and A. Quarteroni (1989), 'A relaxation procedure for domain decomposition methods using finite elements', *Numer. Math.* **56**, 575–598.

T.P. Mathew (1989), 'Domain decomposition and iterative refinement methods for mixed finite element discretisations of elliptic problems', PhD thesis, Courant Institute of Mathematical Sciences. Technical Report 463, Department of Computer Science, Courant Institute.

T.P. Mathew (1993a), 'Schwarz alternating and iterative refinement methods for mixed formulations of elliptic problems, part I: Algorithms and numerical results', *Num. Math.* **65**(4), 445–468.

T.P. Mathew (1993b), 'Schwarz alternating and iterative refinement methods for mixed formulations of elliptic problems, part II: Theory', *Num. Math.* **65**(4), 469–492.

A.M. Matsokin and S.V. Nepomnyaschikh (1985), 'A Schwarz alternating method in a subspace', *Sov. Math.* **29(10)**, 78–84.

S. McCormick (1984), 'Fast adaptive composite grid (FAC) methods', in *Defect Correction Methods: Theory and Applications* (K. Böhmer and H.J. Stetter, eds), Computing Supplement 5, Springer (Wien), 115–121.

S.F. McCormick (1989), *Multilevel Adaptive Methods for Partial Differential Equations*, SIAM (Philadelphia, PA).

S. Meddahi (1993), 'The Schwarz algorithm for a Raviart–Thomas mixed method', preprint, to appear.

G.A. Meurant (1991), 'Numerical experiments with a domain decomposition method for parabolic problems on parallel computers', in *Fourth Int. Symp. on Domain Decomposition Methods for Partial Differential Equations* (R. Glowinski, Y.A. Kuznetsov, G.A. Meurant, J. Périaux and O. Widlund, eds), SIAM (Philadelphia, PA).

K. Miller (1965), 'Numerical analogs of the Schwarz alternating procedure', *Numer. Math.* **7**, 91–103.

Y. Morchoisne (1984), 'Inhomogeneous flow calculations by spectral methods: Mono-domain and multi-domain techniques', in *Spectral Methods for Partial Differential Equations* (R.G. Voigt, D. Gottlieb and M.Y. Hussaini, eds), SIAM-CBMS, 181–208.

D. Morgenstern (1956), 'Begründung des alternierenden Verfahrens durch Orthogonalprojektion', *ZAMM* **36**, 7–8.

F. Nataf and F. Rogier (1993), 'Factorization of the advection-diffusion operator and domain decomposition method', in *Asymptotic and Numerical Methods for Partial Differential Equations with Critical Parameters* (H. Kaper and M. Garbey, eds), Vol. 384, NATO ASI Series C, 123–133.

J. Nečas (1967), *Les Méthodes Directes en Théorie des Équations Elliptiques*, Academia (Prague).

S.V. Nepomnyaschikh (1984), 'On the application of the method of bordering for elliptic mixed boundary value problems and on the difference norms of $W_2^{1/2}(S)$'. In Russian.

S.V. Nepomnyaschikh (1986), 'Domain decomposition and Schwarz methods in a subspace for the approximate solution of elliptic boundary value problems', Computing Center of the Siberian Branch of the USSR Academy of Sciences, Novosibirsk, USSR.

D.P. O'Leary and O.B. Widlund (1979), 'Capacitance matrix methods for the helmholtz equation on general three-dimensional regions', *Math. Comput.* **33**, 849–879.

M. Ong (1989), 'Hierarchical basis preconditioners for second-order elliptic problems in three dimensions', Technical Report 89–3, Department of Applied Maths, University of Washington, Seattle.

P. Oswald (1991), 'On discrete norm estimates related to multilevel preconditioners in the finite element method', in *Proc. Int. Conf. Theory of Functions, Varna 91*, to appear.

S. Pahl (1993), 'Domain decomposition for the Stokes problem', Master's thesis, University of Witwatersrand, Johannesburg.

L.F. Pavarino (1992), Domain decomposition algorithms for the p-version finite element method for elliptic problems, PhD thesis, Courant Institute of Mathematical Sciences, Department of Mathematics.

L.F. Pavarino (1993a), 'Schwarz methods with local refinement for the p-version finite element method', Technical Report TR93-01, Rice University, Department of Computational and Applied Mathematics, submitted to *Numer. Math.*

L.F. Pavarino (1993b), 'Some Schwarz algorithms for the spectral element method', in *Sixth Conf. on Domain Decomposition Methods for Partial Differential Equations* (A. Quarteroni, ed.), American Mathematical Society (Providence, RI), to appear. Technical Report 614, Department of Computer Science, Courant Institute.

L.F. Pavarino and O.B. Widlund (1993), 'Iterative substructuring methods for p-version finite elements in three dimensions', Technical Report, Courant Institute of Mathematical Sciences, Department of Computer Science, to appear.

T.N. Phillips (1992), 'Pseudospectral domain decomposition techniques for the Navier–Stokes equations', in *Fifth Int. Symp. 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), SIAM (Philadelphia, PA).

A. Pothen, H. Simon and K. Liou (1990), 'Partitioning sparse matrices with eigenvector of graphs', *SIAM J. Math. Anal. Appl.* **11**(3), 430–452.

W. Proskurowski and P. Vassilevski (1992), 'Preconditioning nonsymmetric and indefinite capacitance matrix problems in domain imbedding', Technical Report, UCLA, CAM Report 92-48, University of California, Los Angeles. *SIAM J. Sci. Comput.*, to appear.

W. Proskurowski and P. Vassilevski (1994), 'Preconditioning capacitance matrix problems in domain imbedding', *SIAM J. Sci. Comput.*, to appear.

W. Proskurowski and O.B. Widlund (1976), 'On the numerical solution of helmholtz's equation by the capacitance matrix method', *Math. Comput.* **30**, 433–468.

J.S. Przemieniecki (1963), 'Matrix structural analysis of substructures', *Amer. Inst. Aero. Astro. J.* **1**(1), 138–147.

A. Quarteroni (1989), 'Domain decomposition algorithms for the Stokes equations', in *Second Int. Conf. on Domain Decomposition Methods* (T. Chan, R. Glowinski, J. Périaux and O. Widlund, eds), SIAM (Philadelphia, PA).

A. Quarteroni, ed. (1993), *Domain Decomposition Methods in Science and Engineering*, American Mathematical Society (Providence, RI).

A. Quarteroni and A. Valli (1990), 'Theory and applications of Steklov–Poincaré operators for boundary-value problems: the heterogeneous operator case', in *Proc. 4th Int. Conf. on Domain Decomposition Methods, Moscow* (T. Chan, R. Glowinski, J. Périaux and O. Widlund, eds), SIAM (Philadelphia, PA).

D.C. Resasco (1990), 'Domain decomposition algorithms for elliptic partial differential equations', PhD thesis, Department of Computer Science, Yale University.

T. Rusten (1991), 'Iterative methods for mixed finite element systems', University of Oslo.

T. Rusten and R. Winther (1992), 'A preconditioned iterative method for saddle point problems', *SIAM J. Matrix Anal.* **13**(3), 887.

M. Sarkis (1993), 'Two-level Schwarz methods for nonconforming finite elements and discontinuous coefficients', Technical Report 629, Department of Computer Science, Courant Institute of Mathematical Sciences, New York University.

F. Scapini (1990), 'The alternating Schwarz method applied to some biharmonic variational inequalities', *Calcolo* **27**, 57–72.

F. Scapini (1991), 'A decomposition method for some biharmonic problems', *J. Comput. Math.* **9**, 291–300.

H.A. Schwarz (1890), *Gesammelte Mathematische Abhandlungen*, Vol. 2, Springer (Berlin) 133–143. First published in *Vierteljahrsschrift der Naturforschenden Gesellschaft* in Zürich, Vol. 15, 1870, 272–286.

J.S. Scroggs (1989), 'A parallel algorithm for nonlinear convection diffusion equations', in *Third Int. Symp. on Domain Decomposition Methods for Partial Differential Equations*, SIAM (Philadelphia, PA).

H. Simon (1991), 'Partitioning of unstructured problems for parallel processing', *Comput. Sys. Engrg* **2**(2/3), 135–148.

B.F. Smith (1990), 'Domain decomposition algorithms for the partial differential equations of linear elasticity', PhD thesis, Courant Institute of Mathematical Sciences. Technical Report 517, Department of Computer Science, Courant Institute.

B.F. Smith (1991), 'A domain decomposition algorithm for elliptic problems in three dimensions', *Numer. Math.* **60**(2), 219–234.

B.F. Smith (1992), 'An optimal domain decomposition preconditioner for the finite element solution of linear elasticity problems', *SIAM J. Sci. Comput.* **13**(1), 364–378.

B.F. Smith (1993), 'A parallel implementation of an iterative substructuring algorithm for problems in three dimensions', *SIAM J. Sci. Comput.* **14**(2), 406–423.

B.F. Smith and O.B. Widlund (1990), 'A domain decomposition algorithm using a hierarchical basis', *SIAM J. Sci. Comput.* **11**(6), 1212–1220.

B.F. Smith, P. Bjørstad and W.D. Gropp (1994), 'Domain decomposition: algorithms, implementations and a little theory', to appear.

S.L. Sobolev (1936), 'L'algorithme de Schwarz dans la théorie de l'elasticité', *C. R. (Dokl.) Acad. Sci. URSS* **IV**((XIII) 6), 243–246.

J. Sun and J. Zou (1991), 'Ddm preconditioner for 4th order problems by using B-spline finite element method', in *Fourth Int. Symp. on Domain Decomposition Methods for Partial Differential Equations* (R. Glowinski, Y.A. Kuznetsov, G.A. Meurant, J. Périaux and O. Widlund, eds), SIAM (Philadelphia, PA).

W. P. Tang (1988), 'Schwarz splitting and template operators', Department of Computer Science, Stanford University.

C.H. Tong, T.F. Chan and C.C.J. Kuo (1991), 'A domain decomposition preconditioner based on a change to a multilevel nodal basis', *SIAM J. Sci. Comput.* **12**, 1486–1495.

W. Tsui (1991), 'Domain decomposition of biharmonic and Navier–Stokes equations', PhD thesis, Department of Mathematics, University of California, Los Angeles.

J. Wang (1991), 'Convergence analysis of Schwarz algorithms and multilevel decomposition iterative methods: Part I', *Proc. IMACS Int. Symp. on Iterative Methods in Linear Algebra* (R. Beauwens and P. De Groen, eds), (Belgium).

J. Wang (1993), 'Convergence analysis of schwarz algorithms and multilevel decomposition iterative methods: Part II', *SIAM J. Numer. Anal.* **30**(4), 953–970.

M.F. Wheeler and R. Gonzalez (1984), 'Mixed finite element methods for petroleum reservoir engineering problems', in *Computing Methods in Applied Sciences and Engineering, VI* (R. Glowinski and J.L. Lions, eds), North-Holland (New York) 639–658.

O.B. Widlund (1987), 'An extension theorem for finite element spaces with three applications', in *Numerical Techniques in Continuum Mechanics: Notes on Numerical Fluid Mechanics*, Vol. 16, (W. Hackbusch and K. Witsch, eds), Friedr. Vieweg und Sohn (Braunschweig/Wiesbaden) 110–122. Proc. Second GAMM Seminar, Kiel, January, 1986.

O.B. Widlund (1988), 'Iterative substructuring methods: Algorithms and theory for elliptic problems in the plane', in *First Int. Symp. on Domain Decomposition*

Methods for Partial Differential Equations (R. Glowinski, G.H. Golub, G.A. Meurant and J. Périaux, eds), SIAM (Philadelphia, PA).

O.B. Widlund (1989a), 'Iterative solution of elliptic finite element problems on locally refined meshes', in *Finite Element Analysis in Fluids* (T.J. Chung and G.R. Karr, eds), University of Alabama in Huntsville Press (Huntsville, Alabama), 462–467.

O.B. Widlund (1989b), 'Optimal iterative refinement methods', in *Second Int. Conf. on Domain Decomposition Methods* (T. Chan, R. Glowinski, J. Périaux and O. Widlund, eds), SIAM (Philadelphia, PA).

R. Williams (1991), 'Performance of dynamic load balancing algorithms for unstructured mesh calculations', *Concurrency* **3**, 457–481.

J. Xu (1989), 'Theory of multilevel methods', PhD thesis, Cornell University.

J. Xu (1992a), 'Iterative methods by space decomposition and subspace correction', *SIAM Rev.* **34**, 581–613.

J. Xu (1992b), 'A new class of iterative methods for nonselfadjoint or indefinite problems', *SIAM J. Numer. Anal.* **29**(2), 303–319.

J. Xu (1992c), 'Iterative methods by SPD and small subspace solvers for nonsymmetric or indefinite problems', *Fifth Conf. 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) SIAM (Philadelphia, PA).

J. Xu and X.-C. Cai (1992), 'A preconditioned GMRES method for nonsymmetric or indefinite problems', *Math. Comput.* **59**, 311–319.

H. Yserentant (1986), 'On the multi-level splitting of finite element spaces', *Numer. Math.* **49**, 379–412.

X. Zhang (1991), 'Studies in domain decomposition: Multilevel methods and the biharmonic Dirichlet problem', PhD thesis, Courant Institute, New York University.

X. Zhang (1992a), 'Domain decomposition algorithms for the biharmonic Dirichlet problem', in *Fifth Conf. 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), SIAM (Philadelphia, PA).

X. Zhang (1992b), 'Multilevel Schwarz methods', *Num. Math.* **63**(4), 521–539.

X. Zhang (1992c), 'Multilevel Schwarz methods for the biharmonic dirichlet problem', Technical Report CS-TR2907 (UMIACS-TR-92-60), University of Maryland, Department of Computer Science, submitted to *SIAM J. Sci. Comput.*