OVERLAPPING SCHWARZ ALGORITHMS FOR SOLVING HELMHOLTZ'S EQUATION

XIAO-CHUAN CAI*, MARIO A. CASARIN, JR. †, FRANK W. ELLIOTT, JR. ‡, AND OLOF B. WIDLUND §

1. Introduction. In this paper, we give a progress report on the development of a new family of domain decomposition methods for the solution of Helmholtz's equation. We present three algorithms based on overlapping Schwarz methods; in our favorite method we proceed to the continuous finite element approximation of the Helmholtz's equation through a sequence of discontinuous iterates. While this is, quite possibly, a new type of overlapping Schwarz methods, we have been inspired to develop this idea by the thesis of Després [4].

The basic domain decomposition algorithm considered by Després is defined as follows: The given region Ω is divided into two nonoverlapping subregions Ω_1 and Ω_2, and the iteration is advanced by simultaneously solving

$$
\begin{aligned}
-\Delta u_j^{n+1} - k^2 u_j^{n+1} &= f \quad x \in \Omega_j, \\
\partial u_j^{n+1}/\partial n_{int} - iku_j^{n+1} &= -\partial u_{out}^{n}/\partial n_{out} - iku_{out}^{n} \quad x \in \Gamma, \qquad (1)\\
\partial u_j^{n+1}/\partial n_{int} - iku_j^{n+1} &= g \quad x \in \partial\Omega,
\end{aligned}
$$

in the two subregions. Here f and g are data given for the original problem, k is a real parameter, and Γ the interface, i.e. the parts common to $\partial\Omega_1$ and $\partial\Omega_2$, the boundaries of subregions. We note that Sommerfeld-type boundary conditions are used and that the subregions themselves can be the union of a number of disjoint regions as in the case when Ω is cut into strips and the strips colored using two colors. The iterates generally have jumps across the interface; the jump will go to zero as the iteration converges.

In his thesis, Després proves convergence in a relatively weak sense for a quite general decomposition into nonoverlapping subregions and also conducts a detailed theoretical and numerical study for a rectangular Ω cut into two. The convergence is slow, but it is shown that under-relaxation can lead to an improvement. Després also briefly

* University of Colorado, Boulder, CO 80309. Electronic mail address: cai@cs.colorado.edu URL: http://www.cs.colorado.edu/~cai. This work was supported in part by the National Science Foundation under Grants ASC-9457534, ECS-9527169, and ECS-9725004, and in part by AFOSR Grant F49620-97-1-0059.

† IMECC-UNICAMP, Caixa Postal 6065, 13081 - 970 - Campinas - SP, Brazil. Electronic mail address: casarin@ime.unicamp.br URL: http://www.ime.unicamp.br/~casarin. This work was supported in part by the National Science Foundation under Grant NSF-ECS-9527169, and in part by CNPq Brazil.

‡ Courant Institute of Mathematical Sciences, 251 Mercer Street, New York, N.Y. 10012. Electronic mail address: elliott@cims.nyu.edu This work was supported in part by the National Science Foundation under Grant NSF-ECS-9527169, and in part by the U.S. Department of Energy under Contract DE-FG02-92ER25127.

§ Courant Institute of Mathematical Sciences, 251 Mercer Street, New York, N.Y. 10012. Electronic mail address: widlund@cs.nyu.edu URL: http://cs.nyu.edu/cs/faculty/widlund/index.html. This work was supported in part by the National Science Foundation under Grants NSF-CCR-9503408 and NSF-ODURF-354151, and in part by the U.S. Department of Energy under Contract DE-FG02-92ER25127.

considers the use of overlap; it is shown that this leads to a considerable improvement in the rate of convergence of the iteration for the two subregion case.

In the limit of increasing domain diameter, the Sommerfeld boundary condition provides the correct far-field condition for propagation of waves in the frequency domain, but for a bounded region it does not provide perfect transparency. It can be argued that an alternative boundary condition, which more closely approximates the correct nonlocal non-reflecting boundary condition, would lead to more rapid convergence. These ideas have indeed been tested with some success by Ghanemi [5] and others. This essentially amounts to replacing the ik terms in the interface condition (1) with ikT, where T is an appropriate nonlocal operator. See also [9] for work more closely related to ours.

In our own work, we have instead attempted to use three ideas that have proven successful in studies of other types of problems: We have focused almost exclusively on methods based on overlapping decompositions of the region Ω. In addition, we are exploring the possible benefits of a coarse solver as a part of our preconditioner; we note that the use of a coarse space correction is required to establish convergence of domain decomposition algorithms for a class of nonsymmetric and indefinite elliptic problems previously considered by Cai and Widlund [2, 3]. We also take advantage of well-known accelerators of the basic iteration schemes, in particular the GMRES algorithm.

We refer to Smith, Bjørstad, and Gropp [11] for an introduction to domain decomposition methods in general, and these ideas in particular. We note that there are a number of variants of the Schwarz algorithms: additive, hybrid, restricted, etc. In our work, we are now focusing on the classical, multiplicative algorithm.

2. Differential and Discrete Model Problems. We consider a Helmholtz model problem given by

$$-\Delta u - k^2 u = f \quad x \in \Omega, \quad \partial u / \partial n - iku = g \quad x \in \partial\Omega, \tag{2}$$

where Ω is a bounded two or three-dimensional region. This equation is uniquely solvable, and we note that the boundary condition, said to be of Sommerfeld type, is essential in the proof of this fact.

We use Green's formula, and complex conjugation of the test functions, to convert (2) into variational form: Find $u \in H^1(\Omega)$ such that,

$$\begin{aligned} b(u,v) &= \int_\Omega (\nabla u \cdot \nabla \bar{v} - k^2 u \bar{v}) dx - ik \int_{\partial\Omega} u \bar{v} ds \\ &= \int_\Omega f \bar{v} dx + \int_{\partial\Omega} g \bar{v} ds \;=\; F(v) \quad \forall v \in H^1(\Omega). \end{aligned}$$

Finite element problems can now be defined straightforwardly by replacing $H^1(\Omega)$ by a suitable conforming finite element space. So far, we have worked mainly with lower order elements but have made progress towards extending our studies and numerical experiments to spectral elements.

Our interest in the spectral element case has been inspired by the work of Ihlenburg and Babuška [7, 8, 6] and the thesis by Melenk [10]. They have considered the well-posedness of the original problem and different finite element discretizations and proven, for a model problem in one dimension, that the basic estimate

$$|u|_{H^1} \leq Ck|F|_{H^{-1}}$$

holds. In the finite element case, an assumption of $hk < 1$ is used. The constant C is independent of p, the degree of the finite elements. Ihlenburg has also conducted extensive numerical experiments which suggest that this bound also holds for problems in two or three dimensions. Error bounds of the following form are also given for $p = 1$ and kh small enough:

$$|error|_{H^1} \leq C_1\theta + C_2 k\theta^2 \text{ where } \theta = \text{best } H^1\text{−error}.$$

With oscillatory solutions typical, we can expect θ to be on the order of kh. In that case, the second term, which is due to the phase error, will dominate unless k^2h is on the order of 1. Larger values of p appear attractive since Ihlenburg and Babuška have also shown that

$$|error|_{H^1} \leq \theta^p(C_1 + C_2\theta^2) + C_3 k\theta^{2p}.$$

Here $\theta = hk/2p$, and the phase error is now relatively less important.

3. Overlapping Schwarz Algorithms. The basic multiplicative, one-level overlapping Schwarz method can be described as follows: Let $\{\Omega_j\}$ be a set of open subregions that covers the given region Ω. Just as in the strip case of Section 1, each subregion Ω_j can have many disconnected components; it is often profitable to color the subregions of an original overlapping decomposition of Ω using different colors for any pair of subregions that intersect. The original set of subregions can then be partitioned into sets of subregions, one for each color, effectively reducing the number of subregions. This decreases the number of fractional steps of our Schwarz methods and helps make the algorithms parallel. The number of colors is denoted by J.

In many cases, it is appropriate to view a multiplicative Schwarz method as follows: A full iteration step proceeds through J fractional steps,

$$u^{n+j/J} - u^{n+(j-1)/J} \; = \; P_j(u - u^{n+(j-1)/J}),$$

where $P_j, j = 1, \ldots, J$, is a projection onto a subspace V_j related to Ω_j and u is the exact finite element solution. Such a fractional step can be more easily understood by rewriting it in the form

$$b_j(u^{n+j/J} - u^{n+(j-1)/J}, v) \; = \; F(v) - b(u^{n+j-1)/J}, v) \;\; \forall v \in V_j. \tag{3}$$

The choice of the local sesquilinear form $b_j(\cdot,\cdot)$ and the space V_j determines the projection P_j. We will examine several choices one of which has discontinuous iterates, and for it we will need an alternative to formula (3). Introducing a splitting of the Helmholtz form with respect to each Ω_j and its complement Ω_j^c,

$$b(u, v) = b_j(u, v) + b_j^c(u, v), \tag{4}$$

which we will further describe below, we can replace (3) by

$$b_j(u^{n+j/J}, v) = F(v) - b_j^c(u^{n+(j-1)/J}, v). \tag{5}$$

It is also easy to introduce a coarse space correction and a second level into the algorithm. An additional fractional step is then used; we choose to make this correction prior to the other, local steps. In our experiments, we have so far only used the same low order finite element method on a coarser mesh. The space related to this mesh and fractional step is denoted by V_0, and we use formula (3) to define the related, special update. We note that all the solvers used in the fractional steps are smaller, often much smaller, instances of the original problem.

One difficulty with faithfully implementing a generalization of Després' algorithm with overlap is the appearance and disappearance of multiple values, i.e. jumps, across different parts of the interface Γ, which is now defined by

$$\cup \partial\Omega_i \setminus \partial\Omega.$$

In our first two algorithms, we avoid jumps and use traditional domain decomposition techniques, but in the third and most successful algorithm jumps in the solution are fully accommodated.

The three algorithms can now be defined in terms of the sesquilinear forms $b_j(\cdot,\cdot)$ and the subspaces V_j.

ALG1 An update with zero Dirichlet condition on $\partial\Omega_j \setminus \Omega$ is used in the *jth* fractional step; this preserves the continuity of the iterates. The test functions of V_j then vanish at all mesh points in the closure of Ω_j^c. The sesquilinear form is defined by

$$b_j(u,v) = \int_{\Omega_j} \nabla u \cdot \nabla \overline{v} dx - ik \int_{\partial\Omega \cap \partial\Omega_j} u\overline{v} ds.$$

We note that for an interior subregion we cannot guarantee solvability of the subproblem except by making the diameters of the components of the subregion Ω_j small enough. The same preconditioner can also be obtained by a matrix splitting based on the diagonal blocks of variables associated with the nodes in Ω_j and on $\partial\Omega \cap \partial\Omega_j$.

ALG2 The sesquilinear form is chosen as

$$b_j(u,v) = \int_{\Omega_j} \nabla u \cdot \nabla \overline{v} dx - ik \int_{\partial\Omega_j} u\overline{v} ds,$$

and the elements of the space V_j are now required to vanish at all the nodes in the open set Ω_j^c. We require that the solution of equation (5) belong to the same space. Continuity of the iterates is maintained by overwriting all the old values at all the nodes of the closure of Ω_j.

ALG3 The same V_j and $b_j(\cdot,\cdot)$ are used as in ALG2, but both the old and the new values on $\partial\Omega_j$ are saved. This will typically produce a jump across this part of the interface. At the same time the jump across the interface interior to Ω_j is eliminated. Further details will be given; we note that the new features of this algorithm have required a redesign of our data structures, and that there are consequences of the jumps that need careful scrutiny in order to understand ALG3 correctly.

Since we use completely standard techniques for the coarse grid correction, we describe only the fine grid fractional steps of ALG3 in some detail. We must first realize that the lack of continuity across the interface Γ forces us to use broken norms, i.e. to replace integrals over Ω and the Ω_j by sums of integrals over *atomic* subregions defined by Γ. We proceed by finding the common refinement of all splittings like (4). Let $\{\mathrm{A}_q | q = 1, \ldots, Q\}$ be the *open atoms* generated by $\{\Omega_j\}$, i.e. the collection of the largest open sets satisfying $\mathrm{A}_q \subseteq \Omega_j$ or $\mathrm{A}_q \subseteq \Omega_j^c$ for all j and q. We refine (4) by expressing each term as a sum of Helmholtz forms defined on the collection of open atoms contained in that region. Thus,

$$\begin{aligned} b_j(u,v) &= \sum_{\mathrm{A}_q \subseteq \Omega_j} \underline{b}_q(u,v) \\ b_j^c(u,v) &= \sum_{\mathrm{A}_q \subseteq \Omega_j^c} \underline{b}_q(u,v). \end{aligned}$$

These are the splittings needed to solve equation (5) in the presence of jumps and to represent the solution in atomic form for further steps. The sesquilinear forms corresponding to the individual atoms are defined by

$$\underline{b}_q(u,v) = a_{\mathrm{A}_q}(u,v) - ik(u,v)_{\tilde{\Xi}_q} - ik(u,v)_{\tilde{\Gamma}_q^-} + ik(u,v)_{\tilde{\Gamma}_q^+}$$

where,

$$\begin{aligned} \tilde{\Xi}_q &= \partial\mathrm{A}_q \cap \partial\Omega \\ \tilde{\Gamma}_q^+ &= \bigcup \left\{ \partial\mathrm{A}_q \cap \partial\Omega_j \,\middle|\, \Omega_j^c \supseteq \mathrm{A}_q, j = 1,\ldots,J \right\} \\ \tilde{\Gamma}_q^- &= \bigcup \left\{ \partial\mathrm{A}_q \cap \partial\Omega_j^c \,\middle|\, \Omega_j \supseteq \mathrm{A}_q, j = 1,\ldots,J \right\}. \end{aligned}$$

For a valid splitting, we have to assume that the boundaries of any two intersecting subdomains, Ω_i and Ω_j, must have the same unit normal where they intersect, except on sets of measure zero.

The principal difficulty in implementing a multiplicative Schwarz cycle based on (5) is that it requires that multiple values be kept at the atom interfaces $\tilde{\Gamma}_q^-$ and $\tilde{\Gamma}_q^+$ because continuity is not enforced. Therefore, we represent the solution function u as an element in the direct product of finite element spaces, one for each atom. At iteration $n + j/J$ of the algorithm, see (5), the right hand side is computed atomic subregion by atomic subregion. It is therefore practical to store the nodal values of each atom separately. We also note that the test functions v are continuous functions in the closure of Ω and that the solution $u^{n+j/J}$ of (5) is continuous in the closure of Ω_j. Once it is found, it is scattered to the individual atoms of Ω_j. The set of nodal values of the iterate is exactly what is required in the computation of the contribution from that atom to the next set of right hand sides.

After a full sweep through all the subregions, the residuals interior to each atomic subregion are zero, and across any segment of Γ, the solution is either continuous or satisfies the flux condition. Then, by Green's formula, the approximate solution u^n satisfies

$$b(u^n, v) = F(v) + ik \int_\Gamma [u^n]\bar{v}ds \quad \forall v \in V.$$

In the limit, the jump goes to zero; this conveniently signals convergence. At this point of the iteration, the residuals can be computed from the jump directly, but also conventionally through its contributions from the atomic subregions.

4. Theoretical Results. Optimal convergence has been established for ALG1 and ALG2, with a coarse space V^H and GMRES acceleration, using essentially only our older theory; see [2, 3] or [11, Chapter 5.4]. Our result is also valid for ALG3 in the case when there are no cross points. Our current proofs require H^2-regularity and that k^3H^2 is sufficiently small, i.e. the phase error of the coarse space solution is small enough. (We believe that the H^2-regularity can be weakened at the expense of a more severe restriction on k and H.) We note that while these types of conditions are meaningful asymptotically, since our results show that the number of iterations will be independent of h, only experiments can tell if the restriction imposed on H makes our results irrelevant for a choice of mesh points that corresponds to a realistic number of mesh points per wave length. We also note that our current theory fails to explain the quite satisfactory performance that we have observed in many of our experiments even without a coarse correction.

The bound for ALG1 is independent of k and h while that of ALG2 deteriorates linearly with the number of points per wave length.

In view of our results and the formulas for the phase error, we have made series of experiments with a fixed k^3H^2 as well as k^3h^2.

5. Numerical Results. The software used was developed with the numerical linear algebra library PETSc [1] supplied by Argonne National laboratories and Matlab (TM), a product of Mathworks, Inc. We would like to acknowledge the generous help of the PETSc implementors in developing and debugging our code. The platforms for our computations are a Silicon Graphics Reality Monster (TM) parallel computer at Argonne National Laboratories and local workstations.

We now describe the geometry and discretization used in our numerical experiments. In all cases Ω is a unit square discretized with Q_1 elements, and the subregions Ω_j are built from a decomposition of Ω into nonoverlapping square subregions with a layer of δ elements added in all directions. The relevant parameters for describing an experiment are:

- n the number grid points per side of the square fine mesh, n_c the analog for the coarse mesh; $h = 1/(n-1)$ with Ω a unit square;
- k the spatial frequency;
- the number of subregions;
- δ the number of elements across half of the overlap;
- ppw the number of points per wave length; $ppw = \frac{2\pi}{kh}$.

The number of points per wavelength is a non-dimensional measure of resolution.

We first compare the performance of the three algorithms on a 2×2 set of square subregions without using a coarse grid. Here and below, we say that an algorithm has converged at a given iteration if the ℓ^2 norm of the residual is less than 10^{-6} of that of the original residual. We also discuss the relative error at termination of the iteration measured by the ℓ_2 norm of the difference between the final iterate and the exact solution of the discrete system divided by the ℓ_2 norm of the same exact solution. An iteration number given in parentheses indicates divergence and the iteration number at which the iteration was stopped.

For ALG1 GMRES acceleration had to be used at all times to obtain convergence; moreover, given $n = 97$ or 129, $\delta = 1, 2, 3, 4,$ or 5, and $ppw = 10$ or 20, ALG1 never converges in fewer than 40 iterations. Moreover, the relative error in the solution always exceeds 10^{-4}, possibly reflecting ill-conditioning.

For ALG2, we obtain convergence provided we use either overlap or GMRES acceleration. Given the resolution $n = 97$ or 129 and $ppw = 10$, ALG2 diverges when

TABLE 1
Table for **ALG3**

δ	3		2		1	
nc=0;	(101)	3.87e-01	31	9.90e-07	20	6.91e-07
k= 20.1	15	8.97e-07	16	6.84e-07	17	5.75e-07
nc=9;	54	9.47e-07	(101)	1.15e+01	(34)	5.85e+01
k= 20.1	14	5.99e-07	14	6.50e-07	15	7.05e-07
nc=17;	45	9.58e-07	20	9.28e-07	19	9.53e-07
k= 20.1	13	9.85e-07	12	7.44e-07	11	4.44e-07
nc=33;	42	9.44e-07	21	7.33e-07	17	7.04e-07
k= 20.1	13	4.17e-07	12	9.41e-07	11	4.28e-07

δ	3		2		1	
nc=0;	18	9.29e-07	21	7.32e-07	24	7.15e-07
k= 31.9	15	6.19e-07	16	6.38e-07	18	7.98e-07
nc=17;	19	8.09e-07	(101)	5.49e-03	(30)	2.59e+01
k= 31.9	13	6.60e-07	14	9.70e-07	17	5.72e-07
nc=33;	36	9.12e-07	17	7.98e-07	22	8.42e-07
k= 31.9	13	5.41e-07	13	4.86e-07	13	7.81e-07
nc=65;	18	9.09e-07	17	6.87e-07	23	8.72e-07
k= 31.9	12	6.20e-07	12	6.58e-07	13	9.02e-07

$\delta = 0$ unless GMRES acceleration is applied. Even with acceleration but without overlap, ALG2 appears to be worse than ALG1. With the smallest overlap, $\delta = 1$, ALG2 converges without acceleration. The convergence rate, about 42 iterations, is similar to that of ALG1, but the error is one tenth of that encountered in ALG1, apparently reflecting better conditioning. With acceleration and $\delta = 1$, ALG2 converges in 13 iterations for both resolutions.

ALG3 does not converge at all without overlap, but with overlap it outperforms both ALG1 and ALG2. Given $ppw = 10$ the resolution $n = 97$ or 129 and $\delta = 1, 2, 3$ or 4, ALG3 consistently converges at least twice as fast as ALG2. For these parameter values ALG3 always converges in fewer than 10 iterations, and the relative error is always less than 10^{-6}. Given $ppw = 20$, whether $n = 97$ or $n = 129$, the results are the *same* provided that the ratio of δ to ppw is maintained; this indicates that ALG3 has converged with respect to resolution. However, when a coarse grid is used, the ratio of δ to ppw ceases to be accurate determinant of performance.

In the remaining numerical results we investigate the behavior of ALG3 in the case of many subregions. Generally speaking, in the many subregion case ALG3 needs either GMRES acceleration or a sufficiently fine coarse grid to converge. Table 1 shows the results of several runs with ALG3. For the two sub-tables, ppw is 20.00 (above) and 25.20 (below), and the mesh size is 65×65 (above) and 129×129 (below). The two choices of parameters are related by a constant value of k^3h^2. In all cases there are 8×8 subregions, and coarse mesh size varies as indicated in the left column. Within each cell are presented data for the unaccelerated algorithm (upper row), the accelerated algorithm (lower row), the iteration count (left column), the normalized residual (right column).

In all cases without a coarse grid, GMRES forces convergence in 18 iterations or less. By contrast, the algorithm sometimes fails to converge when neither a coarse grid nor GMRES is used. Indeed, other experimental results suggest that with a larger number of subregions ($> 1,000$) convergence without a coarse grid *always requires* GMRES acceleration.

In particular for a run not shown in our tables with ppw=40, a 257×257 fine grid, no coarse grid, $\delta = 2$, and a 32×32 array of overlapping subregions, the accelerated algorithm converges in 74 steps while the unaccelerated algorithm diverges. A coarse grid correction with a 65×65 grid causes convergence in 17 steps for the accelerated algorithm and 86 steps for the unaccelerated algorithm.

Our highest resolution computations, so far, uses a 385×385 fine grid with a 24×24 array of subregions, $\delta = 2$, and $ppw = 53.33$. Even without a coarse grid, the unaccelerated algorithm converges in 46 iterations; with GMRES acceleration the algorithm converges in 42 iterations. With a 97×97 coarse grid, the unaccelerated algorithm requires 22 iterations, but the accelerated algorithm only 9 iterations.

For some problems that we consider the minimum number of coarse grid nodes per wavelength required for the coarse grid correction to be helpful has been as small as three, but this number appears to grow somewhat with larger values of k and larger values of subregions.

REFERENCES

[1] Satish Balay, William Gropp, Lois Curfman McInnes, and Barry F. Smith. *PETSc, the Portable, Extensible Toolkit for Scientific Computation*. Argonne National Laboratory, 2.0.17 April 5, 1997 edition.

[2] Xiao-Chuan Cai and Olof Widlund. Domain decomposition algorithms for indefinite elliptic problems. *SIAM J. Sci. Statist. Comput.*, 13(1):243–258, January 1992.

[3] Xiao-Chuan Cai and Olof Widlund. Multiplicative Schwarz algorithms for some nonsymmetric and indefinite problems. *SIAM J. Numer. Anal.*, 30(4):936–952, August 1993.

[4] Bruno Després. *Méthodes de Décomposition de Domaine pour les Problémes de Propagation d'Ondes en Régime Harmonique*. PhD thesis, Paris IX Dauphine, October 1991.

[5] Souad Ghanemi. *Méthode de Décomposition de Domaine avec Conditions de Transmissions Non Locales pour des Problémes de Propagation d'Ondes*. PhD thesis, Paris IX Dauphine, January 1996.

[6] Frank Ihlenburg and Ivo Babuška. Dispersion analysis and error estimation of Galerkin finite element methods for the Helmholtz equation. *Inter. J. Numer. Meth. Eng*, 38:3745–3774, 1995.

[7] Frank Ihlenburg and Ivo Babuška. Finite element solution of the Helmholtz equation with high wave number, part i: The h-version of the FEM. *Computers Math. Applic.*, 30(9):9–37, 1995.

[8] Frank Ihlenburg and Ivo Babuška. Finite element solution of the Helmholtz equation with high wave number, part ii: The h-p-version of the FEM. *SIAM J. Numer. Anal.*, 34(1):315–358, 1997.

[9] L. C. McInnes, R. F. Susan-Resiga, D. E. Keyes, and H. M. Atassi. Additive Schwarz methods wih nonreflecting boundary conditions for the parallel computation of Helmholtz problems. In Xiao-Chuan Cai, Charbel Farhat, and Jan Mandel, editors, *Tenth International Symposium on Domain Decomposition Methods for Partial Differential Equations*. AMS, 1997. Submitted.

[10] Jens M. Melenk. *On Generalized Fininte Element Methods*. PhD thesis, University of Maryland, 1995.

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