SIAM J. MATRIX ANAL. APPL.
Vol. 33, No. 2, pp. 653–680

AN OPTIMAL BLOCK ITERATIVE METHOD AND PRECONDITIONER FOR BANDED MATRICES WITH APPLICATIONS TO PDES ON IRREGULAR DOMAINS*

MARTIN J. GANDER†, SÉBASTIEN LOISEL‡, AND DANIEL B. SZYLD§

Abstract. Classical Schwarz methods and preconditioners subdivide the domain of a PDE into subdomains and use Dirichlet transmission conditions at the artificial interfaces. Optimized Schwarz methods use Robin (or higher order) transmission conditions instead, and the Robin parameter can be optimized so that the resulting iterative method has an optimized convergence factor. The usual technique used to find the optimal parameter is Fourier analysis; but this is applicable only to certain regular domains, for example, a rectangle, and with constant coefficients. In this paper, we present a completely algebraic version of the optimized Schwarz method, including an algebraic approach to finding the optimal operator or a sparse approximation thereof. This approach allows us to apply this method to any banded or block banded linear system of equations, and in particular to discretizations of PDEs in two and three dimensions on irregular domains. With the computable optimal operator, we prove that the optimized Schwarz method converges in no more than two iterations, even for the case of many subdomains (which means that this optimal operator communicates globally). Similarly, we prove that when we use an optimized Schwarz preconditioner with this optimal operator, the underlying minimal residual Krylov subspace method (e.g., GMRES) converges in no more than two iterations. Very fast convergence is attained even when the optimal transmission operator is approximated by a sparse matrix. Numerical examples illustrating these results are presented.

Key words. linear systems, banded matrices, block matrices, Schwarz methods, optimized Schwarz methods, iterative methods, preconditioners

AMS subject classifications. 65F08, 65F10, 65N22, 65N55

DOI. 10.1137/100796194

1. Introduction.

Finite difference or finite element discretizations of PDEs usually produce matrices which are banded or block banded (e.g., block tridiagonal or block pentadiagonal). In this paper, we present a novel iterative method for such block and banded matrices guaranteed to converge in at most two steps, even in the case of many subdomains. Similarly, its use as a preconditioner for minimal residual methods also achieves convergence in two steps. The formulation of this method proceeds by appropriately replacing a small block of the matrix in the iteration operator. As we will show, approximations of this replacement also produce very fast convergence. The method is based on an algebraic rendition of optimized Schwarz methods.

*Received by the editors May 21, 2010; accepted for publication (in revised form) May 1, 2012; published electronically June 28, 2012.

http://www.siam.org/journals/simax/33-2/79619.html

†Section de Mathématiques, Université de Genève, CP 64, CH-1211 Geneva, Switzerland (gander@math.unige.ch).

‡Department of Mathematics, Heriot-Watt University, Riccarton EH14 4AS, UK (S.Loisel@hw.ac.uk). This author's work was supported in part by The Numerical Algorithms and Intelligent Software Centre funded by UK EPSRC grant EP/G036136 and the Scottish Funding Council, as well as by the U.S. Department of Energy under grant DE-FG02-05ER25672, while he was affiliated with Temple University.

§Department of Mathematics, Temple University (038-16), 1805 N. Broad Street, Philadelphia, PA 19122-6094 (szyld@temple.edu). Part of this research was performed during this author's visit to the Université de Genève, which was supported by the Fond National Suisse under grant FNS 200020-121561/1. This author's work was also supported in part by the U.S. Department of Energy under grant DE-FG02-05ER25672 and by the U.S. National Science Foundation under grant DMS-1115520.

Schwarz methods are important tools for the numerical solution of PDEs. They are based on a decomposition of the domain into subdomains, and on the (approximate) solution of the (local) problems in each subdomain. In the classical formulation, Dirichlet boundary conditions at the artificial interfaces are used; see, e.g., [28], [34], [37], [42]. In optimized Schwarz methods, Robin and higher order boundary conditions are used in the artificial interfaces, e.g., of the form $\partial_n u(x) + pu(x)$. By optimizing the parameter p, one can obtain optimized convergence of the Schwarz methods; see, e.g., [4], [6], [7], [8], [13], [14], [15], [17], [22], and also [19]. The tools usually employed for the study of optimized Schwarz methods and its parameter estimation are based on Fourier analysis. This limits the applicability of the technique to certain classes of differential equations and to simple domains, e.g., rectangles or spheres.

Algebraic analyses of classical Schwarz methods were shown to be useful in their understanding and extensions; see, e.g., [2], [10], [20], [30], [38], [41]. In particular, it follows that the classical additive and multiplicative Schwarz iterative methods and preconditioners can be regarded as the classical block Jacobi or block Gauss–Seidel methods, respectively, with the addition of overlap; see section 2. Inspired in part by the earlier work on algebraic Schwarz methods, in this paper, we mimic the philosophy of optimized Schwarz methods when solving block banded linear systems; see also [24], [25], [26], [27]. Our approach consists of optimizing the block which would correspond to the artificial interface (called transmission matrix) so that the spectral radius of the iteration operator is reduced; see section 3. With the optimal transmission operator, we show that the new method is guaranteed to converge in no more than two steps; see section 4. Such optimal iterations are sometimes called nilpotent [33]; see also [31], [32]. When we use our optimal approach to precondition a minimal residual Krylov subspace method, such as GMRES, the preconditioned iterations are also guaranteed to converge in no more than two steps.

Because the calculation of the optimal transmission matrices is expensive, we propose two general ways of approximating them. We can approximate some inverses appearing in the expression of the optimal matrices, e.g., by using an incomplete LU (ILU) factorization (see also [39, 40] on parallel block ILU preconditioners). We also show how to approximate the optimal transmission matrices using scalar (O0s), diagonal (O0), and tridiagonal (O2) transmission matrices or using some prescribed sparsity pattern.

For a model problem, we compare our algebraic results to those that can be obtained with Fourier analysis on the discretized differential equation; see section 5. Since the new method is applicable to any (block) banded matrix, we can use it to solve systems arising from the discretization of PDEs on unstructured meshes, and/or on irregular domains, and we show in section 6 how our approach applies to many subdomains, while still maintaining convergence in two iterations when the optimal transmission matrices are used.

In section 7, we present several numerical experiments. These experiments show that our methods can be used either iteratively or as preconditioners for Krylov subspace methods. We show that the optimal transmission matrices produce convergence in two iterations, even when these optimal transmission matrices are approximated using an ILU factorization. We also show that our new O0s, O0, and O2 algorithms generally perform better than classical methods such as block Jacobi and restricted additive Schwarz preconditioners. We end with some concluding remarks in section 8.

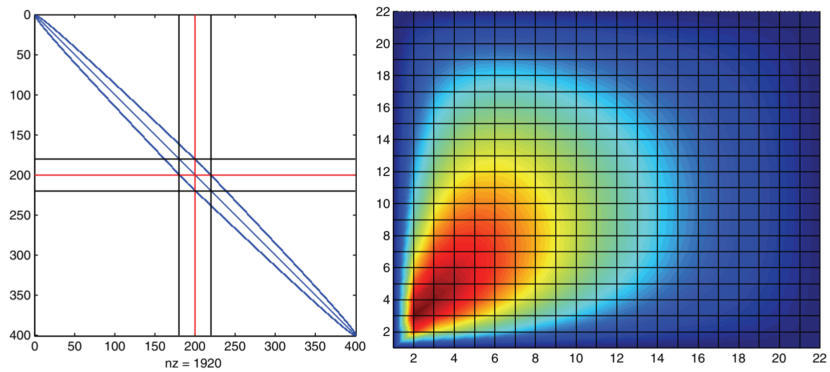

FIG. 2.1. *Left: A 400×400 band matrix partitioned into 4×4 blocks. Right: The corresponding solution to a boundary value problem.*

2. Classical block iterative methods. Our aim is to solve a linear system of equations of the form $Au = f$, where the $n \times n$ matrix A is banded, or block-banded, or, more generally, of the form

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \\ A_{21} & A_{22} & A_{23} & \\ & A_{32} & A_{33} & A_{34} \\ & A_{42} & A_{43} & A_{44} \end{bmatrix}. \tag{2.1}$$

In most practical cases, where A corresponds to a discretization of a differential equation, one has that $A_{13} = A_{42} = O$; i.e., they are zero blocks. Each block A_{ij} is of order $n_i \times n_j$, $i, j = 1, \ldots, 4$, and $\sum_i n_i = n$. We have in mind the situation where $n_1 \gg n_2$ and $n_4 \gg n_3$, as illustrated, e.g., in Figure 2.1.

2.1. Block Jacobi and block Gauss–Seidel methods. Consider first the two diagonal blocks (without overlap)

$$A_1 = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \qquad A_2 = \begin{bmatrix} A_{33} & A_{34} \\ A_{43} & A_{44} \end{bmatrix}, \tag{2.2}$$

which are square but not necessarily of the same size; cf. the example in Figure 2.1 (left). The block Jacobi preconditioner (or block diagonal preconditioner is)

$$M^{-1} = M_{BJ}^{-1} = \begin{bmatrix} A_1^{-1} & O \\ O & A_2^{-1} \end{bmatrix} = \sum_{i=1}^{2} R_i^T A_i^{-1} R_i, \tag{2.3}$$

where the restriction operators are

$$R_1 = [\ I \quad O\ ] \text{ and } R_2 = [\ O \quad I\ ],$$

which have order $(n_1 + n_2) \times n$ and $(n_3 + n_4) \times n$, respectively. The transpose of these operators, R_i^T, are prolongation operators. The standard block Jacobi method using these two blocks has an iteration operator of the form

$$T = T_{BJ} = I - M_{BJ}^{-1} A = I - \sum R_i^T A_i^{-1} R_i A.$$

The iterative method is then, for a given initial vector u^0, $u^{k+1} = Tu^k + M^{-1}f$, $k = 0, 1, \ldots$, and its convergence is linear with an asymptotic convergence factor $\rho(T)$, the spectral radius of the iteration operator; see, e.g., the classical reference [43].

Similarly, the block Gauss–Seidel iterative method for a system with a coefficient matrix (2.1) is defined by an iteration matrix of the form

$$T = T_{GS} = (I - R_2^T A_2^{-1} R_2 A)(I - R_1^T A_1^{-1} R_1 A) = \prod_{i=2}^{1} (I - R_i^T A_i^{-1} R_i A),$$

where the corresponding preconditioner can thus be written as

$$M_{GS}^{-1} = [I - (I - R_2^T A_2^{-1} R_2 A)(I - R_1^T A_1^{-1} R_1 A)]A^{-1}. \tag{2.4}$$

2.2. Additive and multiplicative Schwarz methods. We consider now the same blocks (2.2) with *overlap*, and using the same notation we write the new blocks with overlap as

$$A_1 = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ & A_{32} & A_{33} \end{bmatrix}, \quad A_2 = \begin{bmatrix} A_{22} & A_{23} & \\ A_{32} & A_{33} & A_{34} \\ A_{42} & A_{43} & A_{44} \end{bmatrix}. \tag{2.5}$$

The corresponding restriction operators are again

$$R_1 = [\begin{array}{cc} I & O \end{array}] \quad \text{and} \quad R_2 = [\begin{array}{cc} O & I \end{array}], \tag{2.6}$$

which now have order $(n_1+n_2+n_3)\times n$ and $(n_2+n_3+n_4)\times n$, respectively. With this notation, the additive and multiplicative Schwarz preconditioners (with or without overlap) are

$$M_{AS}^{-1} = \sum_{i=1}^{2} R_i^T A_i^{-1} R_i \quad \text{and} \quad M_{MS}^{-1} = [I - (I - R_2^T A_2^{-1} R_2 A)(I - R_1^T A_1^{-1} R_1 A)]A^{-1}, \tag{2.7}$$

respectively; see, e.g., [37], [42]. By comparing (2.3) and (2.4) with (2.7), one concludes that the classical Schwarz preconditioners can be regarded as block Jacobi or block Gauss–Seidel with the addition of overlap; for more details, see [14].

2.3. Restricted additive and multiplicative Schwarz methods. From the preconditioners (2.7), one can write explicitly the iteration operators for the additive and multiplicative Schwarz iterations as

$$T_{AS} = I - \sum_{i=1}^{2} R_i^T A_i^{-1} R_i A \tag{2.8}$$

$$\text{and} \quad T_{MS} = \prod_{i=2}^{1} (I - R_i^T A_i^{-1} R_i A),$$

respectively. The additive Schwarz iteration (with overlap) associated with the iteration operator in (2.8) is usually not convergent; this is because it holds that with overlap $\sum R_i^T R_i > I$. The standard approach is to use a damping parameter $0 < \gamma < 1$ so that the iteration operator $T_R(\gamma) = I - \gamma \sum_{i=1}^{2} R_i^T A_i^{-1} R_i A$ is such that $\rho(T_R(\gamma)) < 1$;

see, e.g., [37], [42]. We will not pursue this strategy here. Instead we consider the restricted additive Schwarz (RAS) iterations [3], [9], [11].

The RAS method consists of using the local solvers with the overlap (2.5), with the corresponding restriction operators R_i, but using the prolongations $\tilde{R}_i^T$ without the overlap, which are defined as

$$\tilde{R}_1 = \begin{bmatrix} I & O \\ O & O \end{bmatrix} \quad \text{and} \quad \tilde{R}_2 = \begin{bmatrix} O & O \\ O & I \end{bmatrix}, \tag{2.9}$$

having the same order as the matrices R_i in (2.6), and where the identity in $\tilde{R}_1$ is of order $n_1 + n_2$ and that in $\tilde{R}_2$ is of order $n_3 + n_4$. These restriction operators select the variables without the overlap. Note that we now have $\sum \tilde{R}_i^T R_i = I$. In this way, there is no "double counting" of the variables on the overlap, and, under certain hypotheses, there is no need to use a relaxation parameter to obtain convergence; see [11], [14] for details. Thus, the RAS iteration operator is

$$T_{RAS} = I - \sum \tilde{R}_i^T A_i^{-1} R_i A. \tag{2.10}$$

Similarly, one can have restricted multiplicative Schwarz (RMS) [3], [29], and the iteration operator is

$$T = T_{RMS} = \prod_{i=2}^{1} (I - \tilde{R}_i^T A_i^{-1} R_i A) = (I - \tilde{R}_2^T A_2^{-1} R_2 A)(I - \tilde{R}_1^T A_1^{-1} R_1 A), \tag{2.11}$$

although in this case the $\tilde{R}_i^T$ are not necessary to avoid double counting. We include this method just for completeness.

3. Convergence factor for modified restricted Schwarz methods.

Our proposed new method consists of replacing the transmission matrices A_{33} (lowest right corner) in A_1 and A_{22} (upper left corner) in A_2 so that the modified operators of the form (2.10) and (2.11) have small spectral radii, and thus, the corresponding iterative methods have fast convergence. Let the replaced blocks in A_1 and in A_2 be

$$S_1 = A_{33} + D_1 \text{ and } S_2 = A_{22} + D_2, \tag{3.1}$$

respectively, and let us call the modified matrices $\tilde{A}_i$; i.e., we have

$$\tilde{A}_1 = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ & A_{32} & S_1 \end{bmatrix}, \quad \tilde{A}_2 = \begin{bmatrix} S_2 & A_{23} & \\ A_{32} & A_{33} & A_{34} \\ A_{42} & A_{43} & A_{44} \end{bmatrix}; \tag{3.2}$$

cf. (2.5). We consider additive and multiplicative methods in the following two subsections.

3.1. Modified RAS methods.

With the above notation, our proposed modified RAS iteration operator is

$$T_{MRAS} = I - \sum \tilde{R}_i^T \tilde{A}_i^{-1} R_i A, \tag{3.3}$$

and we want to study modifications D_i so that $\|T_{MRAS}\| \ll 1$ for some suitable norm. This would imply of course that $\rho(T_{MRAS}) \ll 1$. Finding the appropriate modifications D_i is analogous to finding the appropriate parameter p in optimized Schwarz methods; see our discussion in section 1 and references therein.

To that end, we first introduce some notation. Let E_3 be the $(n_1+n_2+n_3)\times n_3$ matrix given by $E_3^T = [\ O \quad O \quad I\ ]$, and let E_1 be the $(n_2+n_3+n_4)\times n_2$ matrix given by $E_1^T = [\ I \quad O \quad O\ ]$. Let

$$A_1^{-1}E_3 =: B_3^{(1)} = \begin{bmatrix} B_{31} \\ B_{32} \\ B_{33} \end{bmatrix}, \quad A_2^{-1}E_1 =: B_1^{(2)} = \begin{bmatrix} B_{11} \\ B_{12} \\ B_{13} \end{bmatrix}, \tag{3.4}$$

i.e., the last block column of A_1^{-1}, and the first block column of A_2^{-1}, respectively. Furthermore, we denote

$$\tilde{B}_3^{(1)} = \begin{bmatrix} B_{31} \\ B_{32} \end{bmatrix}, \qquad \tilde{B}_1^{(2)} = \begin{bmatrix} B_{12} \\ B_{13} \end{bmatrix}, \tag{3.5}$$

i.e., pick the first two blocks of $B_3^{(1)}$ of order $(n_1+n_2)\times n_3$ and the last two blocks of $B_1^{(2)}$ of order $(n_3+n_4)\times n_2$. Finally, let

$$\bar{E}_1^T = [\ I \quad O\ ] \text{ and } \bar{E}_2^T = [\ O \quad I\ ], \tag{3.6}$$

which have order $(n_1+n_2)\times n$ and $(n_3+n_4)\times n$, respectively.

LEMMA 3.1. *The new iteration matrix* (3.3) *has the form*

$$T = T_{MRAS} = \left[\begin{array}{c|c} O & K \\ \hline L & O \end{array}\right], \tag{3.7}$$

where

$$\begin{aligned} K &= \tilde{B}_3^{(1)}(I+D_1B_{33})^{-1}\left[\bar{D}_1\bar{E}_1^T - A_{34}\bar{E}_2^T\right], \\ L &= \tilde{B}_1^{(2)}(I+D_2B_{11})^{-1}\left[-A_{21}\bar{E}_1^T + D_2\bar{E}_2^T\right]. \end{aligned} \tag{3.8}$$

Proof. We can write

$$\tilde{A}_1 = A_1 + E_3D_1E_3^T, \quad \tilde{A}_2 = A_2 + E_1D_2E_1^T.$$

Using the Sherman–Morrison–Woodbury formula (see, e.g., [21]) we can explicitly write $\tilde{A}_i^{-1}$ in terms of A_i^{-1} as

$$\tilde{A}_1^{-1} = A_1^{-1} - A_1^{-1}E_3(I + D_1E_3^TA_1^{-1}E_3)^{-1}D_1E_3^TA_1^{-1} =: A_1^{-1} - C_1, \tag{3.9}$$

$$\tilde{A}_2^{-1} = A_2^{-1} - A_2^{-1}E_1(I + D_2E_1^TA_2^{-1}E_1)^{-1}D_2E_1^TA_2^{-1} =: A_2^{-1} - C_2 \tag{3.10}$$

and observe that $E_3^TA_1^{-1}E_3 = B_{33}$, $E_1^TA_2^{-1}E_1 = B_{11}$.

Let us first consider the term with $i = 1$ in (3.3). We begin by noting that from (2.1) it follows that $R_1A = \left[\ A_1 \ \middle|\ E_3A_{34}\ \right]$. Thus,

$$\tilde{A}_1^{-1}R_1A = (A_1^{-1} - C_1)\left[\ A_1 \ \middle|\ E_3A_{34}\ \right] = \left[\ I - C_1A_1 \ \middle|\ A_1^{-1}E_3A_{34} - C_1E_3A_{34}\ \right]. \tag{3.11}$$

We look now at each part of (3.11). First from (3.9), we have that $C_1A_1 = B_3^{(1)}(I + D_1B_{33})^{-1}D_1E_3^T$. Then we see that $C_1E_3A_{34} = B_3^{(1)}(I+D_1B_{33})^{-1}D_1E_3^TA_1^{-1}E_3A_{34} = B_3^{(1)}(I + D_1B_{33})^{-1}D_1B_{33}A_{34}$, and therefore

$$\begin{aligned} A_1^{-1}E_3A_{34} - C_1E_3A_{34} &= B_3^{(1)}[I - (I + D_1B_{33})^{-1}D_1B_{33}]A_{34} \\ &= B_3^{(1)}(I + D_1B_{33})^{-1}(I + D_1B_{33} - D_1B_{33})A_{34} = B_3^{(1)}(I + D_1B_{33})^{-1}A_{34}. \end{aligned}$$

Putting this together, we have

$$\tilde{A}_1^{-1} R_1 A = \left[\begin{array}{c|c} I - B_3^{(1)}(I + D_1 B_{33})^{-1} D_1 E_3^T & B_3^{(1)}(I + D_1 B_{33})^{-1} A_{34} \end{array}\right]. \tag{3.12}$$

It is important to note that the lower blocks in this expression, corresponding to the overlap, will not be considered once it is multiplied by $\tilde{R}_1^T$. An analogous calculation produces

$$\tilde{A}_2^{-1} R_2 A = \left[\begin{array}{c|c} B_1^{(2)}(I + D_2 B_{11})^{-1} A_{21} & I - B_1^{(2)}(I + D_2 B_{11})^{-1} D_2 E_1^T \end{array}\right], \tag{3.13}$$

and again, one should note that the upper blocks would be eliminated with the multiplication by $\tilde{R}_2^T$. We remark that the numbers of columns of the blocks in (3.12) and (3.13) are not the same. Indeed, the first block in (3.12) is of order $(n_1+n_2)\times(n_1+n_2+n_3)$, and the second is of order $(n_1+n_2)\times n_4$, while the first block in (3.13) is of order $(n_3+n_4)\times n_1$, and the second is of order $(n_3+n_4)\times(n_2+n_3+n_4)$.

We apply the prolongations $\tilde{R}_i^T$ from (2.9) to (3.12) and (3.13) and collect terms to form (3.3). First notice that the identity matrix in (3.3) and the identity matrices in (3.12) and (3.13) cancel each other. We thus have

$$\begin{aligned} T &= \left[\begin{array}{c|c} \tilde{B}_3^{(1)}(I + D_1 B_{33})^{-1} D_1 E_3^T & -\tilde{B}_3^{(1)}(I + D_1 B_{33})^{-1} A_{34} \\ -\tilde{B}_1^{(2)}(I + D_2 B_{11})^{-1} A_{21} & \tilde{B}_1^{(2)}(I + D_2 B_{11})^{-1} D_2 E_1^T \end{array}\right] \\ &= \begin{bmatrix} \tilde{B}_3^{(1)}(I + D_1 B_{33})^{-1} \left[\begin{array}{c|c} D_1 E_3^T & -A_{34} \end{array}\right] \\ \tilde{B}_1^{(2)}(I + D_2 B_{11})^{-1} \left[\begin{array}{c|c} -A_{21} & D_2 E_1^T \end{array}\right] \end{bmatrix} \\ &= \begin{bmatrix} \tilde{B}_3^{(1)}(I + D_1 B_{33})^{-1} \left[D_1 \tilde{E}_3^T - A_{34} \tilde{E}_4^T\right] \\ \tilde{B}_1^{(2)}(I + D_2 B_{11})^{-1} \left[-A_{21} \tilde{E}_1^T + D_2 \tilde{E}_2^T\right] \end{bmatrix}, \end{aligned} \tag{3.14}$$

where the last equality follows from enlarging $E_3 = [\,O \;\; O \;\; I\,]^T$ to $\tilde{E}_3 = [\,O \;\; O \;\; I \;\; O\,]^T$ and $E_1 = [\,I \;\; O \;\; O\,]^T$ to $\tilde{E}_2 = [\,O \;\; I \;\; O \;\; O\,]^T$ and introducing $\tilde{E}_4 = [\,O \;\; O \;\; O \;\; I\,]^T$ and $\tilde{E}_1 = [\,I \;\; O \;\; O \;\; O\,]^T$. A careful look at the form of the matrix (3.15) reveals the block structure (3.7) with (3.8). $\square$

Recall that our goal is to find appropriate matrices D_1, D_2 in (3.1) to obtain a small $\rho(T_{MRAS})$. Given the form (3.7) we obtained, it would suffice to minimize $\|K\|$ and $\|L\|$. As it turns out, even in simple cases, the best possible choices of the matrices D_1, D_2 produce a matrix $T = T_{MRAS}$ with $\|T\| > 1$ (although $\rho(T) < 1$); see, for example, the case reported in Figure 3.1. In this case, we show the Laplacian with $D_i = \beta I$. We computed the value of $\|T\|$, $\|T^2\|$ (the 2-norm), and $\rho(T)$ for varying values of the parameter β. We also show an optimized choice of β given by solving an approximate minimization problem which we discuss shortly. It can be appreciated that while $\rho(T) < 1$ for all values of $\beta \in [0, 1]$, $\|T\| > 1$ for most of those values. Furthermore the curve for $\|T^2\|$ is pretty close to that of $\rho(T)$ for a wide range of values of β.

Thus, another strategy is needed. We proceed by considering T^2, which can easily be computed from (3.7) to obtain

$$T^2 = \left[\begin{array}{c|c} KL & O \\ \hline O & LK \end{array}\right]. \tag{3.15}$$

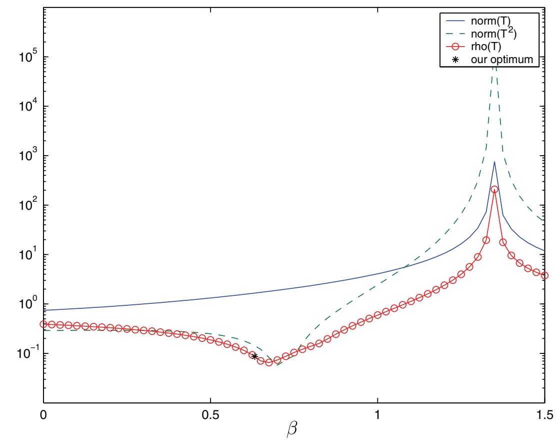

FIG. 3.1. $\|T\|$, $\|T^2\|$, *and* $\rho(T)$ *for* $D_i = \beta I$ *for varying* β. *Laplacian.*

THEOREM 3.2. *The asymptotic convergence factor of the modified RAS method given by* (3.3) *is bounded by the product of the following two norms:*

$$\begin{aligned} &\|(I + D_1 B_{33})^{-1}\,[D_1 B_{12} - A_{34} B_{13}]\,\|, \\ &\|(I + D_2 B_{11})^{-1}\,[D_2 B_{32} - A_{21} B_{31}]\,\|. \end{aligned} \tag{3.16}$$

Proof. We consider T^2 as in (3.15). Using (3.8), (3.6), and (3.5), we can write

(3.17)

$$\begin{aligned} KL &= \tilde{B}_3^{(1)}(I + D_1 B_{33})^{-1}\left[D_1 \bar{E}_1^T - A_{34}\bar{E}_2^T\right]\tilde{B}_1^{(2)}(I + D_2 B_{11})^{-1}\left[-A_{21}\bar{E}_1^T + D_2 \bar{E}_2^T\right] \\ &= \tilde{B}_3^{(1)}(I + D_1 B_{33})^{-1}\,[D_1 B_{12} - A_{34} B_{13}]\,(I + D_2 B_{11})^{-1}\left[-A_{21}\bar{E}_1^T + D_2 \bar{E}_2^T\right] \end{aligned}$$

and similarly

$$LK = \tilde{B}_1^{(2)}(I + D_2 B_{11})^{-1}\,[D_2 B_{32} - A_{21} B_{31}]\,(I + D_1 B_{33})^{-1}\left[D_1 \bar{E}_1^T - A_{34}\bar{E}_2^T\right].$$

Furthermore, let us consider the following products, which are present in $KLKL$ and in $LKLK$:

(3.18)

$$KL\tilde{B}_3^{(1)} = \tilde{B}_3^{(1)}(I + D_1 B_{33})^{-1}\,[D_1 B_{12} - A_{34} B_{13}]\,(I + D_2 B_{11})^{-1}\,[D_2 B_{32} - A_{21} B_{31}]\,,$$

(3.19)

$$LK\tilde{B}_1^{(2)} = B_1^{(2)}(I + D_2 B_{11})^{-1}\,[D_2 B_{32} - A_{21} B_{31}]\,(I + D_1 B_{33})^{-1}\,[D_1 B_{12} - A_{34} B_{13}]\,.$$

These factors are present when considering the powers T^{2k}, and, therefore, asymptotically their norm provides the convergence factor in which T^2 goes to zero. Thus, the asymptotic convergence factor is bounded by the product of the two norms (3.16). ∎

3.2. Modified RMS methods. We now study the idea of using the modified matrices (3.2) for the RMS iterations, obtained by modifying the iteration operator (2.11); i.e., we have

$$T = T_{MRMS} = \prod_{i=2}^{1}(I - \tilde{R}_i^T \tilde{A}_i^{-1} R_i A) = (I - \tilde{R}_2^T \tilde{A}_2^{-1} R_2 A)(I - \tilde{R}_1^T \tilde{A}_1^{-1} R_1 A) \tag{3.20}$$

and its associated preconditioner.

From (3.12), (3.13), and (3.8), we see that

$$(I - \tilde{R}_1^T \tilde{A}_1^{-1} R_1 A) = \left[\begin{array}{c|c} O & K \\ \hline O & I \end{array}\right], \quad (I - \tilde{R}_2^T \tilde{A}_2^{-1} R_2 A) = \left[\begin{array}{c|c} I & O \\ \hline L & O \end{array}\right].$$

As a consequence, putting together (3.20), we have the structure

$$T = T_{MRMS} = \left[\begin{array}{c|c} O & K \\ \hline O & LK \end{array}\right],$$

and from it we can obtain the following result on its eigenvalues.

PROPOSITION 3.3. *Let*

$$T_{MRAS} = \left[\begin{array}{c|c} O & K \\ \hline L & O \end{array}\right], \quad T_{MRMS} = \left[\begin{array}{c|c} O & K \\ \hline O & LK \end{array}\right].$$

If $\lambda \in \sigma(T_{MRAS})$, *then* $\lambda^2 \in \sigma(T_{MRMS})$.

Proof. Let $[x, v]^T$ be the eigenvector of T_{MRAS} corresponding to λ, i.e.,

$$\left[\begin{array}{c|c} O & K \\ \hline L & O \end{array}\right] \left[\begin{array}{c} x \\ v \end{array}\right] = \lambda \left[\begin{array}{c} x \\ v \end{array}\right].$$

Thus, $Kv = \lambda x$, and $Lx = \lambda v$. Then, $LKv = \lambda Lx = \lambda^2 v$, and the eigenvector for T_{MRMS} corresponding to λ^2 is $[x, \lambda v]^T$. ∎

Remark 3.4. We note that the structure of (3.7) is the structure of a standard block Jacobi iteration matrix for a "consistently ordered matrix" (see, e.g., [43], [44]), but our matrix is not of a block Jacobi iteration. We note then that a matrix of this form has the property that if $\mu \in \sigma(T)$, then $-\mu \in \sigma(T)$; see, e.g., [35, p. 120, Prop. 4.12]. This is consistent with our calculations of the spectra of the iteration matrices. Note that for consistently ordered matrices $\rho(T_{GS}) = \rho(T_J)^2$; see, e.g., [43, Corollary 4.26]. Our generic block matrix A is not consistently ordered, but in Proposition 3.3 we proved a similar result.

Observe that in Proposition 3.3 we provide only half of the eigenvalues of T_{MRMS}; the other eigenvalues are zero. Thus we have that $\rho(T_{MRMS}) = \rho(T_{MRAS})^2$, indicating a much faster asymptotic convergence of the multiplicative version.

4. Optimal and optimized transmission matrices. In the present section, we discuss various choices of the transmission matrices D_1 and D_2. We first show that, using Schur complements, the iteration matrix T can be made nilpotent. Since this is an expensive procedure, we then discuss how an ILU approximation can be used to obtain a method which converges very quickly. We finally look at sparse approximations inspired from the optimized Schwarz literature.

4.1. Schur complement transmission conditions. We want to make the convergence factor (3.16) equal to zero, and so we set

$$D_1 B_{12} - A_{34} B_{13} = O \text{ and } D_2 B_{32} - A_{21} B_{31} = O \tag{4.1}$$

and solve for D_1 and D_2. We consider the practical case when $A_{13} = A_{42} = O$. From the definition (3.4), we have that $A_{43}B_{12} + A_{44}B_{13} = O$ or $B_{13} = -A_{44}^{-1}A_{43}B_{12}$ and, similarly, $B_{31} = -A_{11}^{-1}A_{12}B_{32}$. Combining with (4.1), we find the following equations for D_1 and D_2:

$$(D_1 + A_{34}A_{44}^{-1}A_{43})B_{12} = O \text{ and } (D_2 + A_{21}A_{11}^{-1}A_{12})B_{32} = O. \tag{4.2}$$

This can be achieved by using the following Schur complements:

$$D_1 = -A_{34}A_{44}^{-1}A_{43} \text{ and } D_2 = -A_{21}A_{11}^{-1}A_{12}; \tag{4.3}$$

we further note that this is the only choice if B_{12} and B_{32} are invertible. We note that B_{12} and B_{32} are often of low rank and then there may be cheaper choices for D_1 and D_2 that produce nilpotent iterations.

Although the above methods can be used as iterations, it is often beneficial to use them as preconditioners for Krylov subspace solvers such as GMRES or MINRES; see, e.g., [35], [36]. For example, the modified RAS preconditioner is

$$M_{MRAS}^{-1} = \sum \tilde{R}_i^T \tilde{A}_i^{-1} R_i. \tag{4.4}$$

Similarly, we can have a modified RMS preconditioner corresponding to the iteration matrix (3.20). If the Schur complements (4.3) are used for the transmission matrices, then the Krylov space solvers will converge in at most two steps.

PROPOSITION 4.1. *Consider a linear system with coefficient matrix of the form* (2.1) *and a minimal residual method for its solution with either the modified RAS preconditioner* (4.4) *or the modified RMS preconditioner, with* $\tilde{A}_i$ *of the form* (3.2) *and* D_i *(*$i = 1, 2$*) solutions of* (4.2)*. Then, the preconditioned minimal residual method converges in at most two iterations.*

Proof. We can write $T^2 = (I - M^{-1}A)^2 = p_2(M^{-1}A) = 0$, where $p_2(z) = (1-z)^2$ is a particular polynomial of degree 2 with $p_2(0) = 1$. Thus, the minimal residual polynomial $q_2(z)$ of degree 2 also satisfies $q_2(z) = 0$. ∎

4.2. Approximate Schur complements. The factors A_{11}^{-1} and A_{44}^{-1} appearing in (4.3) pose a problem in practice since the matrices A_{11} and A_{44} are large. Hence it is desirable to solve approximately the linear systems

$$A_{44}X = A_{43} \text{ and } A_{11}Y = A_{12}. \tag{4.5}$$

There are of course many computationally attractive ways to do this, including ILU factorizations of the block A_{44} [35] (or of each diagonal block in it in the case of multiple blocks (see section 6), where the factorization can be performed in parallel) or the use of sparse approximate inverse factorizations [1]. In our experiments in this paper we use ILU to approximate the solution of systems like (4.5). An experimental study showing the effectiveness of ILU in this context is presented later in section 7.

4.3. Sparse transmission matrices. The Schur complement transmission matrices (4.3) are dense. We now impose sparsity structures on the transmission matrices D_1 and D_2:

$$D_1 \in \mathcal{Q}_1 \text{ and } D_2 \in \mathcal{Q}_2,$$

where $\mathcal{Q}_1$ and $\mathcal{Q}_2$ denote spaces of matrices with certain sparsity patterns. The optimized choices of D_1 and D_2 are then given by solving the following nonlinear optimization problems:

$$\begin{aligned} &\min_{D_1 \in \mathcal{Q}_1} \|(I + D_1B_{33})^{-1}[D_1B_{12} - A_{34}B_{13}]\|, \\ &\min_{D_2 \in \mathcal{Q}_2} \|(I + D_2B_{11})^{-1}[D_2B_{32} - A_{21}B_{31}]\|. \end{aligned} \tag{4.6}$$

As an approximation, one can also consider the following linear problems:

$$\min_{D_1 \in \mathcal{Q}_1} \|D_1 B_{12} - A_{34} B_{13}\|, \quad \min_{D_2 \in \mathcal{Q}_2} \|D_2 B_{32} - A_{21} B_{31}\|. \tag{4.7}$$

Successful sparsity patterns have been identified in the optimized Schwarz literature. Order 0 methods ("OO0") use diagonal matrices D_1 and D_2, while order 2 methods ("OO2") include off-diagonal components that represent tangential derivatives of order 2; this corresponds to using tridiagonal matrices D_1 and D_2. For details, see [13], [14], and, further, section 5. Inspired by the OO0 and OO2 methods, we propose the following schemes. The O0s scheme uses $D_i = \beta_i I$, where β_i is a scalar parameter to be determined. The O0 scheme uses a general diagonal matrix D_i, and the O2 scheme uses a general tridiagonal matrix D_i.

We choose the Frobenius norm for the linear minimization problem (4.7). For the O0s case, we obtain

$$\begin{aligned} \beta_0 &= \arg\min_{\beta} \|\beta B_{12} - A_{34} B_{13}\|_F = \arg\min_{\beta} \|\beta \, \mathtt{vec}(B_{12}) - \mathtt{vec}(A_{34} B_{13})\|_2 \\ &= \mathtt{vec}(B_{12})^T \mathtt{vec}(A_{34} B_{13}) / \mathtt{vec}(B_{12})^T \mathtt{vec}(B_{12}), \end{aligned} \tag{4.8}$$

where the MATLAB `vec` command produces here an $n_3 \cdot n_2$ vector with the matrix entries. In the O0 case, we look for a diagonal matrix $D_1 = \mathrm{diag}(d_1, \ldots, d_{n_3})$ such that

$$D_1 = \arg\min_{D} \|D B_{12} - A_{34} B_{13}\|_F \tag{4.9}$$

(and similarly for D_2). The problem (4.9) can be decoupled as n_3 problems for each nonzero of D_1 using each column of B_{12} and $A_{34} B_{13}$ to obtain

$$d_i = \arg\min_{d} \|d (B_{12})_i - (A_{34} B_{13})_i\|_F = (B_{12})_i^T (A_{34} B_{13})_i / (B_{12})_i^T (B_{12})_i, \tag{4.10}$$

where we have used the notation X_i to denote the ith column of X. Observe that the cost of obtaining β_0 in (4.8) and that of obtaining the n_3 values of d_i in (4.10) is essentially the same. Similarly, the $O2$ method leads to least squares problems.

Remark 4.2. The methods O0s, O0, and O2 rely on having access to the matrices B_{ij}. Computing the matrices B_{ij} is exactly as difficult as computing the Schur complement, which can then be used to produce nilpotent iterations as per subsection 4.1. Furthermore, any approximation to the B_{ij} can be used to produce approximate Schur complement transmission matrices as per subsection 4.2. In either case, it is not obvious that there is an advantage to approximating these exact or approximate Schur complements sparsely. It remains an open problem to compute sparse matrices D_i without having access to the B_{ij}.

5. Asymptotic convergence factor estimates for a model problem using Fourier analysis. In this section we consider a problem on a simple domain, so we can use Fourier analysis to calculate the optimal parameters as is usually done in optimized Schwarz methods; see, e.g., [13]. We use this analysis to compute the asymptotic convergence factor of the optimized Schwarz iterative method and compare it to what we obtain with our algebraic counterpart.

The model problem we consider is $-\Delta u = f$ in the (horizontal) strip $\Omega = \mathbb{R} \times (0, L)$, with Dirichlet conditions $u = 0$ on the boundary $\partial\Omega$, i.e., at $x = 0, L$. We discretize the continuous operator on a grid whose interval is h in both the x and y

directions, i.e., with vertices at (jh, kh). We assume that $h = L/(m+1)$ so that there are m degrees of freedom along the y axis, given by $y = h, 2h, \ldots, mh$. The stiffness matrix is infinite and block tridiagonal of the form

$$A = \begin{bmatrix} \ddots & \ddots & \ddots & & & \\ & -I & E & -I & & \\ & & -I & E & -I & \\ & & & \ddots & \ddots & \ddots \end{bmatrix},$$

where I is the $m \times m$ identity matrix and E is the $m \times m$ tridiagonal matrix $E = \mathit{tridiag}\,(-1, 4, -1)$. This is the stiffness matrix obtained when we discretize with the finite element method using piecewise linear elements. Since the matrix is infinite, we must specify the space that it acts on. We look for solutions in the space $\ell^2(\mathbb{Z})$ of square-summable sequences. In particular, a solution to $Au = b$ must vanish at infinity. This is similar to solving the Laplace problem in $H_0^1(\Omega)$, where the solution also vanishes at infinity.

We use the subdomains $\Omega_1 = (-\infty, h) \times (0, L)$ and $\Omega_2 = (0, \infty) \times (0, L)$, leading to the decomposition

$$\left[\begin{array}{c|c|c|c} A_{11} & A_{12} & O & O \\ \hline A_{21} & A_{22} & A_{23} & O \\ \hline O & A_{32} & A_{33} & A_{34} \\ \hline O & O & A_{43} & A_{44} \end{array}\right] = \left[\begin{array}{ccc|c|c|ccc} \ddots & \ddots & \ddots & & & & & \\ & -I & E & -I & & & & \\ \hline & & -I & E & -I & & & \\ \hline & & & -I & E & -I & & \\ \hline & & & & -I & E & -I & \\ & & & & & \ddots & \ddots & \ddots \end{array}\right]; \tag{5.1}$$

i.e., we have in this case $n_2 = n_3 = m$.

In optimized Schwarz methods, one uses either Robin conditions (OO0) on the artificial interface or a second order tangential condition (OO2). If we discretize these transmission conditions using the piecewise linear spectral element method (i.e., by replacing the integrals with quadrature rules), we get that $S_i = \frac{1}{2}E + hpI$ for the OO0 iteration, where the scalar p is typically optimized by considering a continuous version of the problem, and using Fourier transforms; see, e.g., [13]. Likewise, for the OO2 iteration, we get that $S_i = \frac{1}{2}E + (hp - \frac{2q}{h})I + \frac{q}{h}J$, where J is the tridiagonal matrix $\mathit{tridiag}\,(1, 0, 1)$, and where p and q are optimized using a continuous version of the problem. In the current paper, we have also proposed the choices $S_i = E - \beta I$ (O0) and $S_i = E - \beta I + \gamma J$ (O2). The O2 and OO2 methods are related via

$$4 - \beta = 2 + hp - \frac{2q}{h} \text{ and } \gamma - 1 = \frac{q}{h} - \frac{1}{2}.$$

However, the O0 method is new and is not directly comparable to the OO0 method, since the off-diagonal entries of $E - \beta I$ cannot match the off-diagonal entries of $E/2 + pI$.[1]

We now obtain an estimate of the convergence factor for the proposed new method.

[1] In OO0, the optimized p is positive because it represents a Robin transmission condition. The best choice of ph is small, and hence the corresponding row sums of $\tilde{A}_i$ are almost zero but positive. We have chosen $D_i = -\beta I$ in order to achieve similar properties for the rows of our $\tilde{A}_i$.

LEMMA 5.1. *Let A be given by (5.1). For $S_1 = A_{33} + D_1$ with $D_1 = -\beta I$ and $\beta \in \mathbb{R}$, the convergence factor estimate (3.16) is*

$$\|(I + D_1 B_{33})^{-1}(D_1 B_{12} - A_{34} B_{13})\| = \max_{k=1,\ldots,m} \left| \frac{-\beta + e^{-w(k)h}}{1 - \beta e^{-w(k)h}} \right| e^{-2w(k)h}, \tag{5.2}$$

where $w(k) = w(k, L, h)$ is the unique positive solution of the relation

$$\cosh(w(k)h) = 2 - \cos\left(\frac{\pi\, kh}{L}\right). \tag{5.3}$$

Note that $w(k)$ is a monotonically increasing function of $k \in [1, m]$.

Proof of Lemma 5.1. Let F be the symmetric orthogonal matrix whose entries are

$$F_{jk} = \sqrt{\frac{m+1}{2}} \sin(\pi jk/(m+1)). \tag{5.4}$$

Consider the auxiliary problem

$$\begin{bmatrix} A_{11} & A_{12} & O \\ A_{21} & A_{22} & A_{23} \\ O & A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \\ C_3 \end{bmatrix} = \begin{bmatrix} O \\ O \\ F \end{bmatrix}. \tag{5.5}$$

Observe that since $F^2 = I$, we have that

$$\begin{bmatrix} B_{31} \\ B_{32} \\ B_{33} \end{bmatrix} = \begin{bmatrix} C_1 \\ C_2 \\ C_3 \end{bmatrix} F. \tag{5.6}$$

We can solve (5.5) for the unknowns C_1, C_2, C_3. By considering (5.1), and since $A_{34} = [-I \ \ O \ \ O \ \ \ldots]$, we see that

$$\begin{bmatrix} A_{11} & A_{12} & O & O \\ A_{21} & A_{22} & A_{23} & O \\ O & A_{32} & A_{33} & -I \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ F \end{bmatrix} = \begin{bmatrix} O \\ O \\ O \end{bmatrix}.$$

Hence, we are solving the discrete problem

$$\begin{cases} (L_h u)(x, y) & = & 0 \text{ for } x = \ldots, -h, 0, h \text{ and } y = h, 2h, \ldots, mh, \\ u(2h, y) & = & \sqrt{\frac{m+1}{2}} \sin(\pi ky/L) \text{ for } y = h, 2h, \ldots, mh, \text{ and} \\ u(x, 0) & = & u(x, L) = 0 \text{ for } x = \ldots, -h, 0, h, \end{cases} \tag{5.7}$$

where the discrete Laplacian L_h is given by

$$(L_h u)(x, y) = 4u(x, y) - u(x - h, y) - u(x + h, y) - u(x, y - h) - u(x, y + h).$$

The two basic solutions to the difference equation are

$$u_\pm(x, y) = e^{\pm w(k)x} \sin(\pi ky/L),$$

where $w(k)$ is the unique positive solution of (5.3).

The subdomain Ω_1 does not contain the $x = \infty$ boundary, but it does contain the $x = -\infty$ boundary. Since we are looking for solutions that vanish at infinity (which, for Ω_1, means $x = -\infty$), the unique solution for the given Dirichlet data at $x = 2h$ is therefore

$$u(x, y) = \left(\sqrt{\frac{m+1}{2}} e^{-2w(k)h} \right) e^{w(k)x} \sin(\pi k y/L).$$

Using (5.4), this gives the formula

$$\left[\begin{array}{c} \\ C_1 \\ \\ \hline C_2 \\ \hline C_3 \end{array} \right] = \left[\begin{array}{c} \vdots \\ FD(3h) \\ \hline FD(2h) \\ \hline FD(h) \end{array} \right],$$

where $D(\xi)$ is the diagonal $m \times m$ matrix whose (k, k)th entry is $e^{-w(k)\xi}$. Hence, from (5.6),

$$\left[\begin{array}{c} \\ B_{31} \\ \\ \hline B_{32} \\ \hline B_{33} \end{array} \right] = \left[\begin{array}{c} \vdots \\ FD(3h)F \\ \hline FD(2h)F \\ \hline FD(h)F \end{array} \right].$$

In other words, the matrix F diagonalizes all the $m \times m$ blocks of $B_3^{(1)}$. Observe that F also diagonalizes J and $E = 4I - J$, and hence all the blocks of A; see the right-hand side of (5.1). Similar reasoning shows that F also diagonalizes the $m \times m$ blocks of $B_1^{(2)}$:

$$\left[\begin{array}{c} B_{11} \\ \hline B_{12} \\ \hline \\ B_{13} \\ \\ \end{array} \right] = \left[\begin{array}{c} FD(h)F \\ \hline FD(2h)F \\ \hline FD(3h)F \\ \vdots \end{array} \right].$$

Hence, the convergence factor estimate (3.16) for our model problem is given by

$$\|(I + D_1 B_{33})^{-1}(D_1 B_{12} - A_{34} B_{13})\| = \|F(I - \beta D(h))^{-1}(-\beta D(2h) + D(3h))F\|,$$

which leads to (5.2). $\square$

LEMMA 5.2. *Consider the cylinder $(-\infty, \infty) \times (0, L)$, where $L > 0$ is the height of the cylinder. Let $h > 0$ be the grid parameter, and consider the domain decomposition* (5.1). *In the limit as the grid parameter h tends to zero, the optimized parameter β behaves asymptotically like*

$$\beta_{opt} = 1 - c\left(\frac{h}{L}\right)^{2/3} + O(h), \tag{5.8}$$

where $c = (\pi/2)^{2/3} \approx 1.35$. The resulting convergence factor is

$$\rho_{opt} = 1 - \left(\frac{32\pi h}{L}\right)^{1/3} + O(h^{2/3}). \tag{5.9}$$

Proof. We begin this proof with some notation and a few observations. Let

$$r(w,h,\beta) = \frac{-\beta + e^{-wh}}{1-\beta e^{-wh}} e^{-2wh}.$$

Then, the convergence factor estimate (5.2) is bounded by and very close to (cf. the argument in [23])

$$\rho(L,h,\beta) = \max_{w\in[w(1),w(m)]} |r(w,h,\beta)|,$$

where $w(k) = w(k,h,L)$ is given by (5.3). Clearly, $r(w,h,0) < r(w,h,\beta)$ whenever $\beta < 0$; hence we will optimize over the range $\beta \geq 0$. Conversely, we must have $1-\beta e^{-wh} > 0$ for every $w \in [w(1), w(m)]$ to avoid an explosion in the denominator of (5.2). By using the value $w = w(1)$, we find that

$$0 \leq \beta < 2 - \cos(\pi h/L) + \sqrt{(3-\cos(\pi h/L))(1-\cos(\pi h/L))} = 1 + \pi\frac{h}{L} + O(h^2). \tag{5.10}$$

We are therefore optimizing β in a closed interval $[0, \beta_{\max}] = [0, 1 + \pi h/L + O(h^2)]$.

We divide the rest of the proof into seven steps.

Step 1. We show that β_{opt} is obtained by solving an equioscillation problem. We define the set $W(\beta) = W(L,h,\beta) = \{w > 0 \text{ such that } |r(w,h,\beta)| = \rho(L,h,\beta)\}$, and we now show that, if $\rho(\beta) = \rho(L,h,\beta)$ is minimized at $\beta = \beta_{opt}$, then $\#W(\beta_{opt}) > 1$. By the envelope theorem [23], if $W(\beta) = \{w^*\}$ is a singleton, then $\rho(\beta) = \rho(L,h,\beta)$ is a differentiable function of β and its derivative is $\frac{\partial}{\partial\beta}|r(w^*,h,\beta)|$. Since

$$\frac{\partial}{\partial\beta} r(w,h,\beta) = \frac{e^{-2wh}-1}{(\beta e^{-wh}-1)^2} e^{-2wh}, \tag{5.11}$$

we obtain

$$0 = \frac{d\rho}{d\beta}(\beta_{opt}) = \frac{e^{-2w^*h}-1}{(\beta_{opt}e^{-w^*h}-1)^2} e^{-2w^*h} sgn(r),$$

which is impossible. Therefore, $\#W(\beta_{opt}) \geq 2$; i.e., β_{opt} is obtained by equioscillating $r(w)$ at at least two distinct points of $W(\beta)$.

Step 2. We find the critical values of $r(w) = r(w,h,\beta)$ as a function of w alone. By differentiating, we find that the critical points are $\pm w_{\min}$, where

$$w_{\min} = w_{\min}(\beta,h) = -\ln\left(\frac{1}{4}\,\frac{3+\beta^2-\sqrt{9-10\,\beta^2+\beta^4}}{\beta}\right) h^{-1}. \tag{5.12}$$

In the situation $\beta > 1$, $w_{\min}$ is complex, and hence there is no critical point. In the situation $\beta = 1$, we have $w_{\min} = 0$, which is outside of the domain $[w(1), w(m)]$ of $r(w)$. Since $r(w)$ is differentiable over its domain $[w(1), w(m)]$, its extrema must be either at critical points or at the endpoints of the interval; i.e.,

$$W(\beta_{opt}) \subset \{w(1), w_{\min}, w(m)\}.$$

We now compute β_{opt} by assuming that[2] $W(\beta_{opt}) = \{w(1), w_{\min}\}$. For this value of β_{opt}, we will verify (in Step 7) that if we choose any $\beta < \beta_{opt}$, then $\rho(\beta) \geq$

[2]In practice, we tried the three possibilities $W(\beta_{opt}) = \{w(1), w_{\min}\}, \{w(1), w(m)\}, \{w_{\min}, w(m)\}$. It turns out that the first one is the correct one. In analyzing the first case, the proof that the remaining two cases do not give optimal values of β arises naturally.

$r(w(1), h, \beta) > r(w(1), h, \beta_{opt}) = \rho(\beta_{opt})$, and hence no $\beta < \beta_{opt}$ is optimal. A similar argument applies to the case $\beta > \beta_{opt}$.

Step 3. We now consider the solution(s) of the equioscillation problem $W(\beta_{opt}) = \{w(1), w_{\min}\}$, and we show that $r(w(1)) > 0$ and $r(w_{\min}) < 0$, and that $r(w(1)) + r(w_{\min}) = 0$. Since $W(\beta) = \{w(1), w_{\min}\}$, we must have that $w_{\min} > w(1)$. If we had that $r(w(1)) = r(w_{\min})$, the mean value theorem would yield another critical point in the interval $(w(1), w_{\min})$. Therefore, it must be that $r(w(1)) + r(w_{\min}) = 0$. We now check that $r(w_{\min}) < 0$. Indeed, $r(w)$ is negative when w is large, and $r(+\infty) = 0$. If we had $r(w_{\min}) > 0$, there would be a $w' > w_{\min}$ such that $r(w') < 0$ is minimized, creating another critical point. Since $w_{\min}$ is the only critical point, it must be that $r(w_{\min}) < 0$. Hence, $r(w(1)) > 0$.

Step 4. We show that there is a unique solution to the equioscillation problem $W(\beta_{opt}) = \{w(1), w_{\min}\}$, characterized by $r(w(1)) + r(w_{\min}) = 0$. From (5.11), we see that $\frac{\partial r}{\partial \beta}(w(1), h, \beta) < 0$, and likewise,

$$\frac{\partial(r(w_{\min}(\beta, h), h, \beta))}{\partial \beta} = \frac{\partial r}{\partial \beta}(w_{\min}, h, \beta) + \overbrace{\frac{\partial r}{\partial w}(w_{\min}, h, \beta)}^{=0} \frac{\partial w_{\min}}{\partial \beta}(\beta, h) < 0.$$

Combining the facts that $r(w(1)) > 0$ and $r(w_{\min}) < 0$ are both decreasing in β, there is a unique value of $\beta = \beta_{opt}$ such that $r(w(1)) + r(w_{\min}) = 0$; this β_{opt} will minimize $\rho(L, h, \beta_{opt})$ under the assumption that $W(\beta) = \{w(1), w_{\min}\}$.

Step 5. We give an asymptotic formula for the unique β_{opt} solving the equioscillation problem $W(\beta_{opt}) = \{w(1), w_{\min}\}$. To this end, we make the ansatz[3] $\beta = 1 - c(h/L)^{2/3}$, and we find that

$$r(w(1)) = 1 - \frac{2\pi}{c}\left(\frac{h}{L}\right)^{1/3} + O(h^{2/3}) \text{ and} \tag{5.13}$$

$$r(w_{\min}) = -1 + 4\sqrt{c}\left(\frac{h}{L}\right)^{1/3} + O(h^{2/3}). \tag{5.14}$$

Hence, the equioscillation occurs when $c = (\pi/2)^{2/3}$.

Step 6. We now show that the equioscillation $W(\beta_{opt}) = \{w(1), w_{\min}\}$ occurs when $w_{\min} \in (w(1), w(m))$. Let $\beta_{opt} = 1 - c(h/L)^{2/3} + O(h)$. Then, from (5.12) and (5.3),

$$w_{\min} = \frac{\sqrt{c}}{L^{1/3} h^{2/3}} + O(h^{2/3}) < w(m) = \frac{\operatorname{arccosh}(3)}{h} + O(h),$$

provided that h is sufficiently small.

Step 7. If $\beta < \beta_{opt}$, then $\rho(\beta) > \rho(\beta_{opt})$. Indeed, we see from (5.11) that $\frac{\partial r}{\partial \beta}(w_1, h, \beta) < 0$. Hence, if $\beta < \beta_{opt}$, then $\rho(\beta) \geq r(w_1, h, \beta) > r(w_1, h, \beta_{opt}) = \rho(\beta_{opt})$. A similar argument shows that if $\beta > \beta_{opt}$, then $\rho(\beta) > \rho(\beta_{opt})$.

We therefore conclude that the β_{opt} minimizing $\rho(\beta)$ is the unique solution to the equioscillation problem $W(\beta_{opt}) = \{w(1), w_{\min}\}$, and its asymptotic expansion is given by (5.8). We compute a series expansion of $\rho(L, h, 1 - c(h/L)^{2/3})$ to obtain (5.9). ∎

This shows that the O0 method converges at a rate similar to the OO0 method. In a practical problem where the domain is not a strip, or the partial differential equation

[3] This ansatz is inspired from the result obtained in the OO0 case; cf. [13] and references therein.

is not the Laplacian, if one wants to obtain the best possible convergence factor, then one should solve the nonlinear optimization problem (4.6). We now consider the convergence factor obtained when instead the linear minimization problem (4.7) is solved.

LEMMA 5.3. *For our model problem, the solution of the optimization problem*

$$\beta_0 = \arg\min_{\beta} \|(-\beta B_{12} + B_{13})\|$$

is

$$\beta_0 = \frac{3}{2}\,\frac{\left(\left(1+\sqrt{2}\right)^{2/3} - 1\right)s}{\sqrt[3]{1+\sqrt{2}}}, \tag{5.15}$$

where

$$s^{-1} = 2 - \cos\left(\frac{\pi h}{L}\right) + \sqrt{3 - \cos\left(\frac{\pi h}{L}\right)}\sqrt{1 - \cos\left(\frac{\pi h}{L}\right)}. \tag{5.16}$$

The resulting asymptotics are

$$\begin{aligned} \beta_0 &= 0.894\ldots - 2.8089\ldots(h/L) + O(h^2) \quad \textit{and} \\ \rho_0 &= 1 - 62.477\ldots(h/L) + O(h^2). \end{aligned} \tag{5.17}$$

We mention that the classical Schwarz iteration as implemented, e.g., using RAS, is obtained with $\beta = 0$, yielding the asymptotic convergence factor

$$1 - 9.42\ldots(h/L) + O(h^2).$$

In other words, our algorithm is asymptotically $62.477/9.42 \approx 6.6$ times faster than a classical Schwarz iteration in the sense that it will take about 6.6 iterations of a classical Schwarz method to equal one of our O0 method, with the parameter $\beta = \beta_0$, if h is small. An optimized Schwarz method such as OO0 would further improve the asymptotic convergence factor to $1 - ch^{1/3} + \ldots$ (where $c > 0$ is a constant). This indicates that one can gain significant performance by optimizing the nonlinear problem (4.6) instead of (4.7).

Proof of Lemma 5.3. By proceeding as in the proof of Lemma 5.1, we find that

$$\|(-\beta B_{12} + B_{13})\| = \max_{w \in \{w(1),\ldots,w(m)\}} |(\beta - e^{-wh})e^{-2wh}|.$$

We thus set

$$r_0(w) = r_0(w,\beta,h) = (\beta - e^{-wh})e^{-2wh}.$$

The function $r_0(w)$ has a single extremum at $w^* = w^*(h,\beta) = (1/h)\ln(3/(2\beta))$. We further find that

$$r_0(w^*) = \frac{4}{27}\beta^3,$$

independently of h. We look for an equioscillation by setting

$$r_0(w(1,L,h),\beta_0,h) = r_0(w^*(\beta_0,h),\beta_0,h);$$

that is,

$$\frac{4}{27}\beta_0^3 + s^{-2}\beta_0 + s^{-3} = 0.$$

Solving for the unknown β yields (5.3) and (5.15). Substituting $\beta = \beta_0$ and $w = w(1)$ into (5.2) and taking a series expansion in h gives (5.17). ∎

6. Multiple diagonal blocks. Our analysis so far has been restricted to the case of two (overlapping) blocks. We show in this section that our analysis applies to multiple (overlapping) blocks. To that end, we use a standard trick of Schwarz methods to handle the case of multiple subdomains if they can be colored with two colors.

Let $Q_1, \ldots, Q_p$ be restriction matrices, defined by taking rows of the $n \times n$ identity matrix I; cf. (2.6). Let $\tilde{Q}_1^T, \ldots, \tilde{Q}_p^T$ be the corresponding prolongation operators, such that

$$I = \sum_{k=1}^{p} \tilde{Q}_k^T Q_k;$$

cf. (2.9). Given the stiffness matrix A, we say that the domain decomposition is two-colored if

$$Q_i^T A Q_j = O \text{ for all } |i-j| > 1.$$

In this situation, if p is even, we can define

$$R_1 = \begin{bmatrix} Q_1 \\ Q_3 \\ \vdots \\ Q_{p-1} \end{bmatrix} \text{ and } R_2 = \begin{bmatrix} Q_2 \\ Q_4 \\ \vdots \\ Q_p \end{bmatrix}. \tag{6.1}$$

We make similar definitions if p is odd and also assemble $\tilde{R}_1$ and $\tilde{R}_2$ in a similar fashion.

The rows and columns of the matrices R_1 and R_2 could be permuted in such a way that (2.6) holds. (For an example of this, see Figure 7.6, bottom right.) Therefore, all the arguments in the previous sections, as in the rest of the paper, hold, mutatis mutandis. In particular, the optimal interface conditions D_1 and D_2 give an algorithm that converges in two steps, regardless of the number of subdomains (see also [16] for the required operators in the case of an arbitrary decomposition with cross points).

Although it is possible to "physically" reorder R_1 and R_2 in this way, it is computationally more convenient to work with the matrices R_1 and R_2 as defined by (6.1). We now outline the computations needed to obtain the optimal transmission matrices (or their approximations).

The matrices A_1 and A_2, defined by $A_i = R_i A R_i^T$, similar to (2.5), are block diagonal. The matrices E_1 and E_3 are defined by

$$E_i = \tilde{R}_i R_{3-i}^T,$$

and the matrices $B_3^{(1)}$ and $B_1^{(2)}$ are defined by

$$B_3^{(1)} = A_1^{-1} E_3 \text{ and } B_1^{(2)} = A_2^{-1} E_1;$$

cf. (3.4). Since the matrices A_1 and A_2 are block diagonal, we have retained the parallelism of the p subdomains.

We must now define the finer structures, such as B_{12} and A_{34}. We say that the kth row is in the kernel of X if $X e_k = 0$, where $e_k = [0, \ldots, 0, 1, 0, \ldots, 0]^T$ is the usual basis vector. Likewise, we say that the kth column is in the kernel of X if $e_k^T X = 0$.

We define the matrix B_{12} to be the rows of $B_1^{(2)}$ that are not in the kernel of $R_1\tilde{R}_2^T$. We define the matrix B_{13} to be the rows of $B_1^{(2)}$ that are in the kernel of $R_1R_2^T$, and we make similar definitions for B_{32} and B_{31}; cf. (3.4). The matrix A_{34} is the submatrix of A_2 whose rows are not in the kernel of $R_1\tilde{R}_2^T$, and whose columns are in the kernel of $R_1R_2^T$. We make similar considerations for the other blocks A_{ij}.

This derivation allows us to define transmission matrices defined by (4.7) in the case of multiple diagonal overlapping blocks.

7. Numerical experiments. We have three sets of numerical experiments. In the first set, we use an advection-reaction-diffusion equation with variable coefficients discretized using a finite difference approach. We consider all permutations of square- and L-shaped regions, two and four subdomains, and additive and multiplicative preconditioning, as iterations and used as preconditioners for GMRES with the following methods: nonoverlapping block Jacobi; overlapping block Jacobi (RAS); and our O0s, O0, O2, and optimal methods, and their ILU approximations. Our second set of experiments uses a space shuttle domain with a finite element discretization. We test our multiplicative preconditioners and iterations for two and eight subdomains. In our third set of experiments, we validate the asymptotic analysis of section 5 on a square domain.

7.1. Advection-reaction-diffusion problem. For these numerical experiments, we consider a finite difference discretization of a two-dimensional advection-reaction-diffusion equation of the form

$$\eta u - \nabla \cdot (a \nabla u) + b \cdot \nabla u = f, \tag{7.1}$$

where

$$a = a(x,y), \quad b = \begin{bmatrix} b_1(x,y) \\ b_2(x,y) \end{bmatrix}, \quad \eta = \eta(x,y) \geq 0, \tag{7.2}$$

with $b_1 = y - 1/2$, $b_2 = -(x - 1/2)$, $\eta = x^2 \cos(x+y)^2$, $a = (x+y)^2 e^{x-y}$. We consider two domain shapes—a square and an L-shaped region.[4]

Note that this problem is numerically challenging. The PDE is nonsymmetric, with significant advection. The diffusion coefficient approaches 0 near the corner $(x,y) = (0,0)$, which creates a boundary layer that does not disappear when $h \to 0$. Our numerical methods perform well despite these significant difficulties. In Figure 2.1 (right), we have plotted the solution to this problem with the forcing $f = 1$. The boundary layer is visible in the lower left corner. Since the boundary layer occurs at a point and the equation is elliptic in the rest of the domain, the ellipticity is "strong enough" that there are no oscillations appearing in the solution for the discretization described below.

For the finite difference discretization, we use $h = 1/21$ in each direction resulting in a banded matrix with $n = 400$ (square domain), $n = 300$ (L-shaped domain), and a semiband of size 20. We preprocess the matrix using the reverse Cuthill–McKee algorithm; see, e.g., [18]. This results in the matrix depicted in Figure 2.1 (left). In the same figure, we show the two subdomain partitions used, i.e., with $n_1 = n_4 = 180$ and $n_2 = n_3 = 20$.

[4] An anonymous reviewer points out that concave polygonal domains can give rise to polylogarithmic singularities on the boundary. However, our use of homogeneous Dirichlet conditions prevents this situation from arising. We also note [5], which studies the case of when such singularities really do occur.

Our results, summarized in Figures 7.1–7.4, are organized as follows. Each figure consists of four plots. In each figure, the top left plot summarizes the convergence histories of the various additive iterations, while the bottom left plot summarizes the convergence histories of the multiplicative iterations. The right plots give the corresponding convergence histories of the methods used as preconditioners for GMRES. Note that the plots of the iterative methods use the Euclidean norm of the error, while the plots of the GMRES methods show the Euclidean norm of the preconditioned residual. For the iterative methods, we use a random initial vector u^0 and zero forcing $f = 0$. For GMRES, we use random forcing; the initial vector $u^0 = 0$ is chosen by the GMRES algorithm.

The methods labeled "Nonoverlapping" and "Overlapping" correspond to the nonoverlapping block Jacobi and RAS preconditioners respectively; see section 2. Our new methods use $D_i = \beta_i I$ (O0s), $D_i =$ diagonal (O0), and $D_i =$ tridiagonal (O2); see section 4. As noted in section 4, it will be preferable to compute B_{ij} approximately in a practical algorithm. To simulate this, we have used an ILU factorization of blocks of A with threshold $\tau = 0.2/n_3$ to compute approximations $B_{ij}^{(ILU)}$ to the matrices B_{ij}. From those matrices, we have then computed O0s (ilu), O0 (ilu), O2 (ilu), and Optimal (ilu) transmission conditions D_1 and D_2.

We now discuss the results in Figure 7.1 (square domain, two subdomains) in detail. The number of iterations to reduce the norm of the error below 10^{-8} is two for the Optimal iteration (as predicted by our theory), 22 for the O0 iteration, and 12 for the O2 iteration, for the additive variants. We also test the O0s method with scalar matrices $D_i = \beta I$. This method is not well suited to the present problem

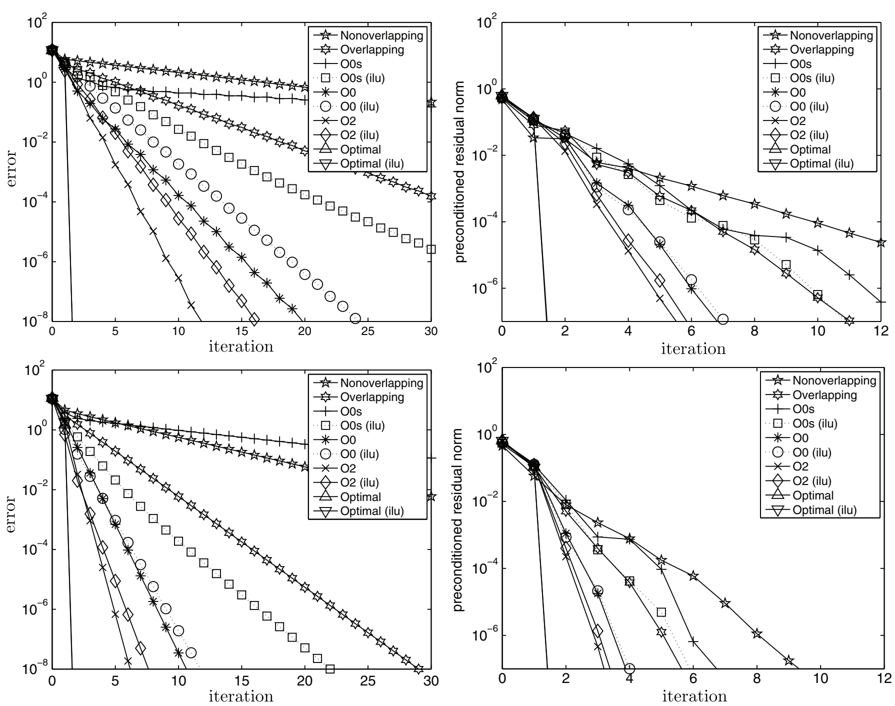

FIG. 7.1. *Square domain, two subdomains. Left: Iterative methods. Right: GMRES. Top: Additive. Bottom: Multiplicative.*

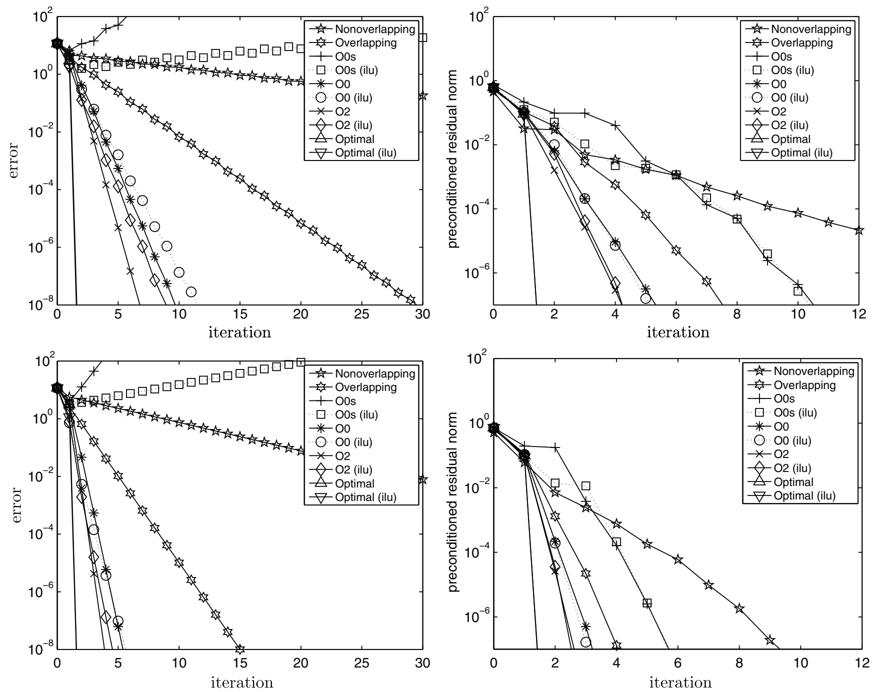

FIG. 7.2. *Square domain, four subdomains. Left: Iterative methods. Right: GMRES. Top: Additive. Bottom: Multiplicative.*

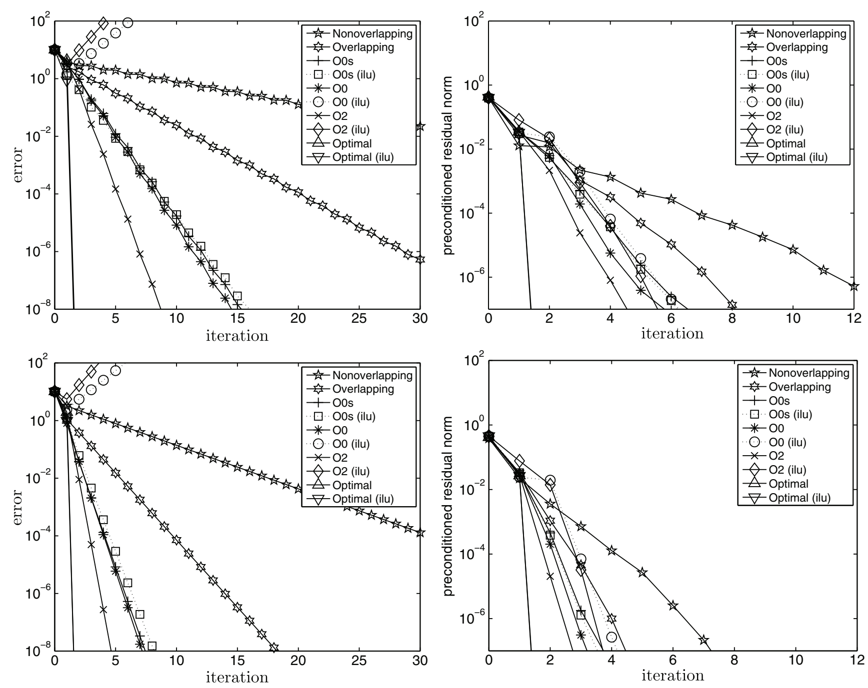

FIG. 7.3. *L-shaped domain, two subdomains. Left: Iterative methods. Right: GMRES. Top: Additive. Bottom: Multiplicative.*

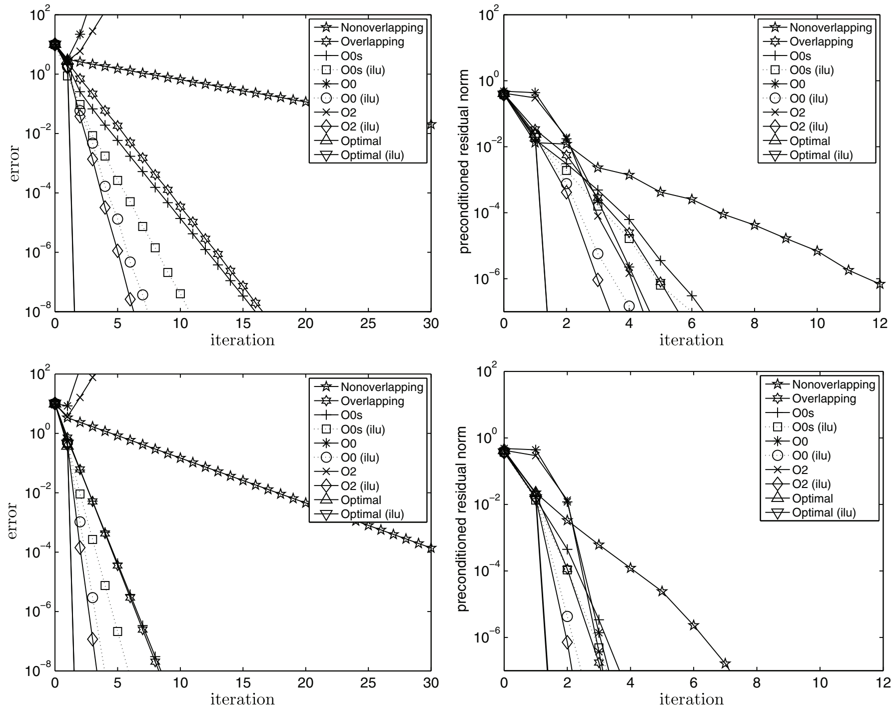

FIG. 7.4. *L-shaped domain, four subdomains. Left: Iterative methods. Right: GMRES. Top: Additive. Bottom: Multiplicative.*

because the coefficients vary significantly, and hence the diagonal elements are not well approximated by a scalar multiple of the identity. This explains why the O0s method requires over 200 iterations to converge. Both the optimal D_i and their ILU approximations converge in two iterations. For the iterative variants O2, O0, there is some deterioration of the convergence factor, but this results in only one extra iteration for the GMRES accelerated iteration. The multiplicative algorithms converge faster than the additive ones, as expected.

We also show in the same plots the convergence histories of block Jacobi (without overlap) and RAS (overlapping). In these cases, the number of iterations to reduce the norm of the error below 10^{-8} is 179 and 59, respectively. One sees that the new methods are much faster than the block Jacobi and RAS methods.

In Figure 7.2 (square domain, four subdomains) we also obtain good results, except that the O0s method now diverges. This is not unexpected, since the PDE has variable coefficients. Nevertheless, the O0s method works as a preconditioner for GMRES. The tests for the L-shaped region are summarized in Figures 7.3 and 7.4. Our methods continue to perform well with the following exceptions. The O0s methods continue to struggle due to the variable coefficients of the PDE. In Figure 7.4, we also note that the O2 iterations diverge. The reverse Cuthill–McKee ordering has rearranged the A_{22} and A_{33}, and these matrices are not tridiagonal. In the O2 algorithm, the sparsity pattern of the D_1 and D_2 matrices should be chosen to match the sparsity pattern of the A_{ii} blocks. In order to check this hypothesis, we reran the test without reordering the vertices with the reverse Cuthill–McKee algorithm. These experiments, summarized in Figure 7.5, confirm our theory: when the sparsity

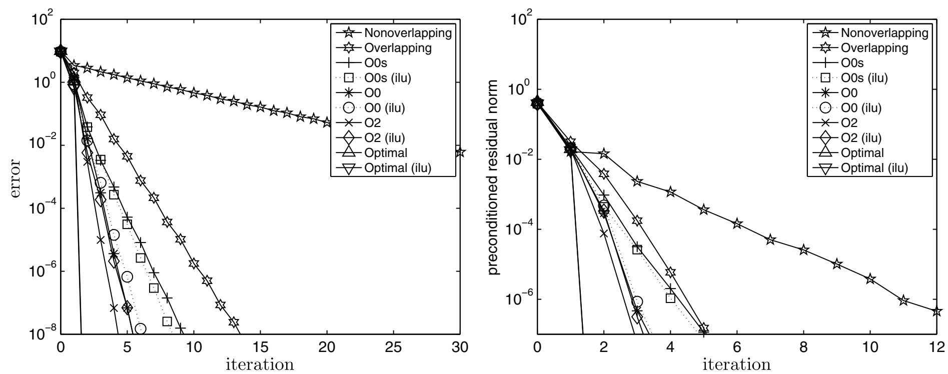

FIG. 7.5. *Iterative methods without reverse Cuthill–McKee reordering; L-shaped, additive, four subdomains. Left: Iterative method. Right: GMRES.*

structure of D_1 and D_2 match the corresponding blocks of A_1 and A_2, the method O2 converges rapidly.

7.2. Domain decomposition of a space shuttle with finite elements. In order to illustrate our methods applied to more general domains, we solved a Laplace problem on a space shuttle model (Figure 7.6, top left for the solution and top right for the domain decomposition). We have partitioned the shuttle into eight overlapping subdomains; the two rightmost subdomains are actually disconnected. This partition is done purely by considering the stiffness matrix of the problem and partitioning it into eight overlapping blocks (Figure 7.6, top right). Note that as before this matrix has been ordered in such a way that the bandwidth is minimized, using a reverse Cuthill–McKee ordering of the vertices; this produces the block structure found in Figure 7.6 (bottom left). In Figure 7.6 (bottom right), we show the block structure of the same matrix once all the odd-numbered blocks have been permuted to the 1st and 2nd block rows and columns and the even-numbered blocks have been permuted to the 3rd and 4th block rows and columns, as per section 6.

We have tested our multiplicative preconditioners on this space shuttle model problem with two and eight subdomains (Figure 7.7). We note that our optimized preconditioners (O0s, O0, O2, and Optimal) converge faster than traditional Schwarz preconditioners (Nonoverlapping and Overlapping), even using the ILU approximations $B_{12}^{(ILU)}$ and $B_{13}^{(ILU)}$ for the matrices B_{12} and B_{13}. The exception is the O0s (ILU) iteration with eight subdomains, which diverges. However, this preconditioner works well with GMRES.

Because we are solving a Laplacian, the diagonal entries of A are all of similar size. As a result, the $O0s$ preconditioner behaves in a manner very similar to the $O0$ preconditioner (and both preconditioners work well). The O2 preconditioner gives improved performance, and the Optimal preconditioner gives nilpotent iterations that converge in two steps independently of the number of subdomains. Similar results were obtained with additive preconditioners.

7.3. Scaling experiments for a rectangular region with an approximate Schur complement. We present an experimental study showing the effectiveness of ILU when used to approximate the solution of systems with the atomic blocks, as in (4.5). To that end, we consider a simpler PDE, namely, the Laplacian on the rectangle $(-1, 1) \times (0, 1)$; i.e., $a = 1$, $b = O$, $\eta = 0$ in (7.2). We use two

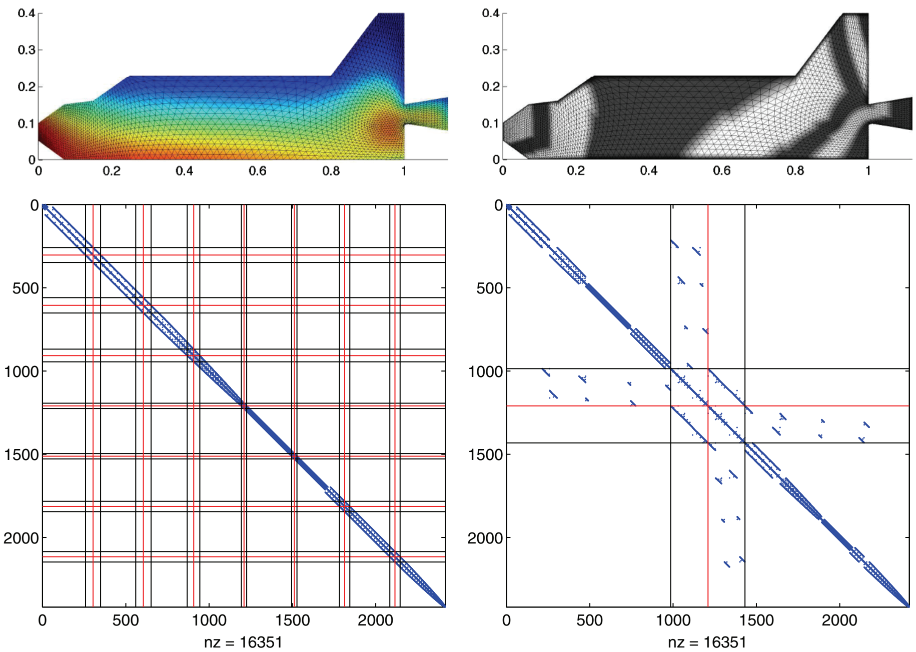

FIG. 7.6. *The space shuttle model and the solution (top left), the domain decomposition (top right), the corresponding matrix partitioning (bottom left), and the even-odd reordering of section* 6 *(bottom right).*

overlapping subdomains, which then correspond to two overlapping blocks in the band matrix. We consider several systems of equations of increasing order by decreasing the value of the mesh parameter h. We show that for this problem, the rate of convergence of our method when the ILU is used stays very close to that obtained with the exact solution of the systems with the atomic blocks. Furthermore, in the one-parameter approximations to the transmission matrices, the value of this parameter β computed with ILU is also very close to that obtained with the exact solutions. See Figure 7.8.

We mesh the rectangle $(-1,1)\times(0,1)$, with a regular mesh, with the mesh interval h. The vertices of the first subdomain are all the vertices in $(-1,h/2]\times(0,1)$, and the vertices of the second subdomain are all the vertices in $[-h/2,1)\times(0,1)$. We choose h in such a way that there are no vertices on the line $x=0$, but instead the interfaces are at $x=\pm h/2$. By ordering the vertices such that all the vertices in $x<-h/2$ occur first, then all the vertices with $x=-h/2$ occur second, then all the vertices with $x=h/2$ occur third, and finally all the vertices with $x>h/2$ occur fourth, we obtain a stiffness matrix of the form (2.1), with additionally $A_{13}=A_{42}=O$. More specifically, the matrix A is a finite matrix of the form (5.1). We use $D_i=\beta I$ and use the optimized parameter given by (4.8).

As with the other experiments, we computed using the matrices B_{ij}. Since B_{12},B_{13} are difficult to compute, we also used an ILU decomposition of A_2 to obtain approximations $B_{12}^{(ILU)}$ and $B_{13}^{(ILU)}$, which we then plug into (4.8). To obtain a good value of β, we used a drop tolerance of $1/(n_2+n_3+n_4)$, where $n_2+n_3+n_4$ is the dimension of A_2. Using this drop tolerance, we found that the L and U factors have approximately ten times as many nonzero entries as the matrix A. Since the two

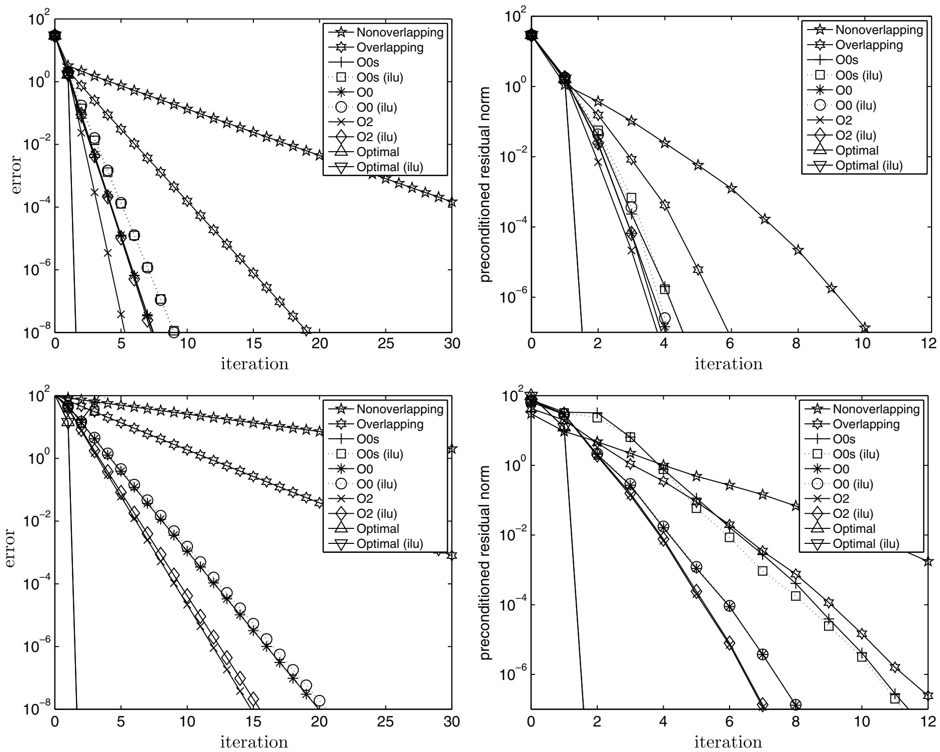

FIG. 7.7. *Convergence histories for the shuttle problem with two subdomains (top) and eight subdomains (bottom) with the multiplicative preconditioners used iteratively (left) or with GMRES acceleration (right).*

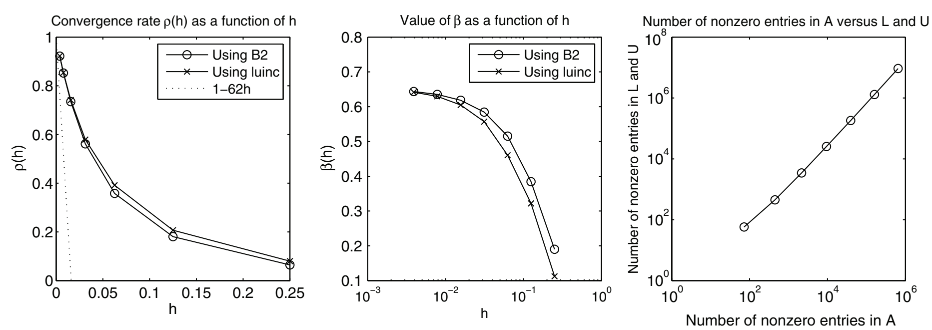

FIG. 7.8. *Convergence factor of the new iterative method, using an approximate calculation for β. Left: The convergence factor as a function of h for the value of β obtained from* (4.8) *as well as the value obtained by using the ILU approximation. We also plot the line $1-62h$; cf.* (5.9). *Middle: The two β parameters as a function of h. Right: The number of nonzero entries of the L and U factors, compared to the number of nonzero entries of A.*

subdomains are symmetrical, the value of β computed using (4.8) is the same for each subdomain.

Note that as $h \to 0$, the value of β approaches 0.65, whereas the theoretical calculation (5.17) approaches 0.894. The source of this disparity is the different domain geometries: the numerical experiments use a rectangle, while the analysis uses an infinite strip. For a more detailed analysis, see [12].

8. Concluding remarks. Inspired by the optimized Schwarz methods for the solution (and preconditioning) of PDEs on simple domains, we have presented an algebraic view of these methods. These new methods can be applied to banded and block banded matrices, again as iterative methods, and as preconditioners. The new method can be seen as the application of several local Schur complements. When these Schur complements are computed, the method and preconditioner are guaranteed to converge in two steps. The new formulation presents these Schur complements as solutions of nonlinear least squares problems. We can approximate the solution of these problems by solving closely related linear least squares problems. Further approximations are obtained by restricting the solution of these linear minimizations to matrices of certain sparsity patterns. When the matrix is not reordered, the blocks of A can be tridiagonal (or themselves decompose into tridiagonal blocks), which motivated our choice of tridiagonal matrices D_i for the O2 case. Other patterns might be more suitable for reordered matrices (it would make sense to use the sparsity pattern of A to guide the choice of sparsity pattern of D_i). Experiments show that these approximations, as well as the use of ILU to approximate the computation of the inverses in the Schur complement, produce very fast methods and preconditioners.

Acknowledgment. We are thankful to the many suggestions we obtained from two anonymous referees, which greatly helped to improve the first version of our manuscript.

REFERENCES

[1] M. Benzi, *Preconditioning techniques for large linear systems: A survey*, J. Comput. Phys., 182 (2002), pp. 418–477.

[2] M. Benzi, A. Frommer, R. Nabben, and D. B. Szyld, *Algebraic theory of multiplicative Schwarz methods*, Numer. Math., 89 (2001), pp. 605–639.

[3] X.-C. Cai and M. Sarkis, *A restricted additive Schwarz preconditioner for general sparse linear systems*, SIAM J. Sci. Comput., 21 (1999), pp. 792–797.

[4] P. Chevalier and F. Nataf, *Symmetrized method with optimized second-order conditions for the Helmholtz equation*, Contemp. Math., 218 (1998), pp. 400–407.

[5] C. Chniti, F. Nataf, and F. Nier, *Improved interface conditions for 2D domain decomposition with corners: A theoretical determination*, Calcolo, 45 (2008), pp. 111–147.

[6] J. Côté, M. J. Gander, L. Laayouni, and S. Loisel, *Comparison of the Dirichlet-Neumann and optimal Schwarz method on the sphere*, in R. Kornhuber et al., eds., Domain Decomposition Methods in Science and Engineering, Lect. Notes Comput. Sci. Eng. 40, Springer, Berlin, Heidelberg, 2004, pp. 235–242.

[7] V. Dolean, S. Lanteri, and F. Nataf, *Optimized interface conditions for domain decomposition methods in fluid dynamics*, Internat. J. Numer. Methods Fluids, 40 (2002), pp. 1539–1550.

[8] O. Dubois, M. J. Gander, S. Loisel, A. St-Cyr, and D. B. Szyld, *The optimized Schwarz method with a coarse grid correction*, SIAM J. Sci. Comput., 34 (2012), pp. A421–A458.

[9] E. Efstathiou and M. J. Gander, *Why restricted additive Schwarz converges faster than additive Schwarz*, BIT, 43 (2003), pp. 945–959.

[10] A. Frommer and D. B. Szyld, *Weighted max norms, splittings, and overlapping additive Schwarz iterations*, Numer. Math., 83 (1999), pp. 259–278.

[11] A. Frommer and D. B. Szyld, *An algebraic convergence theory for restricted additive Schwarz methods using weighted max norms*, SIAM J. Numer. Anal., 39 (2001), pp. 463–479.

[12] M. J. Gander, *On the influence of geometry on optimized Schwarz methods*, Bol. Soc. Esp. Mat. Apl., 53 (2011), pp. 71–78.

[13] M. J. Gander, *Optimized Schwarz methods*, SIAM J. Numer. Anal., 44 (2006), pp. 699–731.

[14] M. J. Gander, *Schwarz methods in the course of time*, Electron. Trans. Numer. Anal., 31 (2008), pp. 228–255.

[15] M. J. Gander, L. Halpern, and F. Nataf, *Optimized Schwarz methods*, in Proceedings of the Twelfth International Conference on Domain Decomposition, T. Chan et al., eds., DDM Press, Bergen, Norway, 2001, pp. 15–27.

[16] M. J. Gander and F. Kwok, *Optimal interface conditions for an arbitrary decomposition into subdomains*, in Domain Decomposition Methods in Science and Engineering XIX, Vol. 78, Y. Huang et al., eds., Springer, Berlin, 2010, pp. 101–108.

[17] M. J. Gander, F. Magoules, and F. Nataf, *Optimized Schwarz methods without overlap for the Helmholtz equation*, SIAM J. Sci. Comput., 24 (2002), pp. 38–60.

[18] A. George and J. W. Liu, *Computer Solution of Large Sparse Positive Definite Systems*, Prentice–Hall, Englewood Cliffs, NJ, 1981.

[19] L. Gerardo-Giorda and F. Nataf, *Optimized algebraic interface conditions in domain decomposition methods for strongly heterogeneous unsymmetric problems*, in Domain Decomposition Methods in Science and Engineering XVI, O. B. Widlund and D. E. Keyes, eds., Springer, Berlin, 2006, pp. 189–196.

[20] M. Griebel and P. Oswald, *On the abstract theory of additive and multiplicative Schwarz algorithms*, Numer. Math., 70 (1995), pp. 163–180.

[21] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed., The Johns Hopkins University Press, Baltimore, MD, 1996.

[22] S. Loisel, *Optimal and Optimized Domain Decomposition Methods on the Sphere*, Ph.D. thesis, Department of Mathematics, McGill University, Montreal, Canada, 2005.

[23] S. Loisel, J. Côté, M. J. Gander, L. Laayouni, and A. Qaddouri, *Optimized domain decomposition methods for the spherical Laplacian*, SIAM J. Numer. Anal., 48 (2010).

[24] F. Magoulès, F.-X. Roux, and S. Salmon, *Optimal discrete transmission conditions for a nonoverlapping domain decomposition method for the Helmholtz equation*, SIAM J. Sci. Comput., 25 (2004), pp. 1497–1515.

[25] F. Magoulès, F.-X. Roux, and L. Series, *Algebraic way to derive absorbing boundary conditions for the Helmholtz equation*, J. Comput. Acoustics, 13 (2005), pp. 433–454.

[26] F. Magoulès, F.-X. Roux, and L. Series, *Algebraic approximation of Dirichlet-to-Neumann maps for the equations of linear elasticity*, Comput. Methods Appl. Mech. Engrg., 195 (2006), pp. 3742–3759.

[27] M. Naumov and A. H. Sameh, *A tearing-based hybrid parallel banded linear system solver*, J. Comput. Appl. Math., 226 (2009), pp. 306–318.

[28] T. P. A. Mathew, *Domain Decomposition Methods for the Numerical Solution of Partial Differential Equations*, Lect. Notes Comput. Sci. Eng. 61, Springer, Berlin, Heidelberg, 2008.

[29] R. Nabben and D. B. Szyld, *Convergence theory of restricted multiplicative Schwarz methods*, SIAM J. Numer. Anal., 40 (2003), pp. 2318–2336.

[30] R. Nabben and D. B. Szyld, *Schwarz iterations for symmetric positive semidefinite problems*, SIAM J. Matrix Anal. Appl., 29 (2006), pp. 98–116.

[31] F. Nataf, F. Rogier, and E. de Sturler, *Optimal Interface Conditions for Domain Decomposition Methods*, Technical report CMAP-301, Centre de Mathematiques Appliquees, CNRS URA-756, Ecole Polytechnique, Paris, France, 1994.

[32] F. Nataf, F. Rogier, and E. de Sturler, *Domain decomposition methods for fluid dynamics*, in Navier-Stokes Equations and Related Nonlinear Problems, A. Sequeira, ed., Plenum Press, New York, 1995, pp. 367–376.

[33] F. Nier, *Remarques sur les algorithmes de décomposition de domaines*, in Séminaire Équations aux Dérivés Partielles de lÉcole Polytechnique, Centre des Mathematiques, Palaiseau, France, 1998–1999, Exposé IX.

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

[35] Y. Saad, *Iterative Methods for Sparse Linear Systems*, 2nd ed. SIAM, Philadelphia, 2003.

[36] V. Simoncini and D. B. Szyld, *Recent computational developments in Krylov subspace methods for linear systems*, Numer. Linear Algebra Appl., 14 (2007), pp. 1–59.

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

[38] A. St-Cyr, M. J. Gander, and S. J. Thomas, *Optimized multiplicative, additive, and restricted additive Schwarz preconditioning*, SIAM J. Sci. Comput., 29 (2007), pp. 2402–2425.

[39] E. de Sturler, *IBLU preconditioners for massively parallel computers*, in Domain Decomposition Methods in Science and Engineering, Proceedings of the Seventh International Conference on Domain Decomposition, D. E. Keys and J. Xu, eds., AMS, Providence, RI, 1994.

[40] E. de Sturler, *Incomplete block LU preconditioners on slightly overlapping subdomains for a massively parallel computer*, Appl. Numer. Math., 19 (1995), pp. 129–146.

[41] W. P. Tang, *Generalized Schwarz splittings*, SIAM J. Sci. Statist. Comput., 13 (1992), pp. 573–595.

[42] A. Toselli and O. Widlund, *Domain Decomposition Methods: Algorithms and Theory*, Springer Ser. Comput. Math. 34, Springer, Berlin, Heidelberg, 2005.

[43] R. S. Varga, *Matrix Iterative Analysis*, 2nd ed., Springer Ser. Comput. Math. 27, Springer, Berlin, Heidelberg, New York, 2000.

[44] D. M. Young, *Iterative Solution of Large Linear Systems*, Academic Press, New York, 1971.