Numer. Math. 63, 521–539 (1992)

Numerische Mathematik

Multilevel Schwarz methods

Xuejun Zhang[1]
Department of Computer Sciences, University of Maryland, College Park, MD 20742, USA

Received April 20, 1992

Summary. We consider the solution of the algebraic system of equations which result from the discretization of second order elliptic equations. A class of multilevel algorithms are studied using the additive Schwarz framework. We establish that the condition number of the iteration operators are bounded independent of mesh sizes and the number of levels. This is an improvement on Dryja and Widlund's result on a multilevel additive Schwarz algorithm, as well as Bramble, Pasciak and Xu's result on the BPX algorithm. Some multiplicative variants of the multilevel methods are also considered. We establish that the energy norms of the corresponding iteration operators are bounded by a constant less than one, which is independent of the number of levels. For a proper ordering, the iteration operators correspond to the error propagation operators of certain V-cycle multigrid methods, using Gauss-Seidel and damped Jacobi methods as smoothers, respectively.

Mathematics Subject Classification (1991): 65F10, 65N30

1. Introduction

Multilevel methods, such as multigrid methods, are some of the most efficient methods of solving large systems of linear equations arising from the finite element or finite difference discretization of elliptic PDES; cf. Hackbusch [12], McCormick [13] and the references therein. Recently, with the increasing interest in parallel computation, several new multilevel methods have been developed and analyzed, e.g. Yserentant's hierarchical basis method [21], the hierarchical basis multigrid method of Bank et al. [1], the parallel multilevel preconditioners developed in Bramble, Pasciak and Xu [7] and Xu [19], and the multilevel additive Schwarz methods of Dryja and Widlund [10].

* This work was supported in part by the National Science Foundation under Grants NSF-CCR-8903003 at Courant Institute of Mathematical Sciences, New York University and NSF-ASC-8958544 at Department of Computer Science, University of Maryland.

We consider second order, self-adjoint, uniformly elliptic differential equations on two or three-dimensional polygonal domains, approximated by continuous, piecewise linear finite elements. We use multilevel Schwarz algorithms to solve the resulting linear system and estimate the condition number of the additive algorithms and the energy norm of the error propagation operator of the multiplicative algorithms. When the additive algorithms are used, an equivalent equation is solved by an iterative method such as the conjugate gradient method. In each iteration, a number of independent problems corresponding to the subdomains are solved. The size of all the subproblems can be very small. Dryja and Widlund have shown that the condition number of a multilevel additive Schwarz (MAS) operator grows at most quadratically with the number of levels; cf. [10]. Similar results for the BPX algorithm were established in Bramble et al. [7]. In this paper, we improve the results for a class of multilevel methods. We show that the condition number of the MAS operator is bounded by a constant independent of mesh sizes and the number of levels. We note that Peter Oswald [16] has obtained a similar result for the BPX algorithm using Besov space theory. For alternative proofs of Oswald's result, see Bramble and Pasciak [5], Xu [20] and Bornemann and Yserentant [4].

The rest of the paper is organized as follows. In Sect. 2, we describe a class of multilevel additive Schwarz algorithms. In Sect. 3, we establish a bound for the condition number of the iteration operator of the algorithm. In Sect. 4, we describe a variant of the algorithm. We construct a very special decomposition of the space, and show that it can also be regarded as a multilevel diagonal scaling algorithm. In the case of constant coefficients and uniform triangulation, this algorithm is identical, up to a constant multiple, to the main algorithm developed in Bramble et al. [7]. As a consequence, we also obtain an improved result for the BPX algorithm. In Sect. 5, we consider some multilevel multiplicative Schwarz schemes and establish that the energy norms of the iterative operators are bounded by a constant less than one, independent of the mesh size and the number of levels. In Sect. 6, we report on some numerical results for the multilevel additive Schwarz method. For a discussion of implementations of the algorithms on parallel computers, see Bjørstad et al. [2] and Bjørstad and Skogen [3], who implemented multilevel additive Schwarz algorithms on a MasPar MP-1, a massively parallel, SIMD machine. The use of approximate solvers for the subproblems are also discussed in those papers.

The present paper is based in part on chapter 3 of the author's thesis [23], see also [22]. The result for the upper bound of the eigenvalues of the additive operator was obtained in early 1991 and announced by Widlund [18] at the 5th domain decomposition conference in early May, 1991, held at Norfolk, VA.

2. Multilevel additive Schwarz methods

Consider the following second order elliptic problem

$$\begin{cases} -\sum_{i,j=1}^{d} \dfrac{\partial}{\partial x_i} a_{ij} \dfrac{\partial}{\partial x_j} u = f & \text{in } \Omega , \\ u = 0 & \text{on } \partial\Omega . \end{cases}$$

Here Ω is a bounded polygonal region in $R^d, d = 2$ or 3. The matrix $\{a_{ij}\}$ is symmetric and positive definite, i.e. its eigenvalues $\lambda_i(x) > 0, \forall x \in \Omega$. The variational form is: Find $u \in H_0^1(\Omega)$ such that

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

where

$$a(u, v) = \int_\Omega \sum_{i,j=1}^d a_{ij} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} dx \qquad \text{and} \qquad f(v) = \int_\Omega fv\, dx .$$

We describe the method and carry out the analysis for Poisson's equation. However there is no difficulty in carrying out the analysis for more general second order problems. In particular, we can obtain a good upper bound as long as $\lambda_d(x)/\lambda_1(x)$ is uniformly bounded in Ω and the coefficients $\{a_{ij}\}$ do not change very much inside individual subdomains.

We define a sequence of nested triangulations $\{\mathcal{T}^l\}_{l=1}^L$. We start with a coarse triangulation $\mathcal{T}^1 = \{\tau_i^1\}_{i=1}^{N_1}$, where τ_i^1 represents an individual triangle. The successively finer triangulations $\mathcal{T}^l = \{\tau_i^l\}_{i=1}^{N_l}$ are defined by dividing individual triangles in the set $\mathcal{T}^{l-1}$ into several triangles. We make a similar construction for three dimensional problems. We assume that all the triangulations are shape regular. Let $h_i^l = diameter(\tau_i^l)$, $h_l = \max_i h_i^l$ and $h = h_L$. We also assume that there exists a constant $\gamma < 1$ and a constant C, such that if an element τ_i^{l+k} of level $l + k$ is contained in an element τ_j^l of level l, then

$$\frac{diam(\tau_i^{l+k})}{diam(\tau_j^l)} \leqq C\gamma^k .$$

For a uniform refinement, with each triangle divided into k^2 equal triangles, $\gamma = 1/k$ and $C = 1$. We denote by $\mathcal{N}^l$ and $\mathcal{E}^l$ the sets of nodes and edges induced by the triangulations $\mathcal{T}^l$, and by $\mathcal{E}^l(S)$ the edges of the subset S.

Let $V^l, l = 1, \ldots, L$, be the space of continuous piecewise linear elements associated with the triangulation $\mathcal{T}^l$. The finite element solution $u_h \in V^L$ satisfies

$$a(u_h, \phi_h) = f(\phi_h), \quad \forall \phi_h \in V^L . \tag{1}$$

We assume that there are $L - 1$ sets of overlapping subdomains $\{\hat{\Omega}_i^l\}_{i=1}^{N_l}, l = 2, 3, \ldots, L$. Thus, on each level, there is an overlapping decomposition $\Omega = \bigcup_{i=1}^{N_l} \hat{\Omega}_i^l$. We make the following assumption about the sets $\{\hat{\Omega}_i^l\}$.

Assumption 2.1. *On each level, the decomposition* $\Omega = \bigcup_{i=1}^{N_l} \hat{\Omega}_i^l$ *satisfies*:

(a) $\partial\hat{\Omega}_i^l$ *aligns with the boundaries of level* l *elements, i.e.* $\hat{\Omega}_i^l$ *is the union of level* l *elements. The diameter*$(\hat{\Omega}_i^l) = O(h_{l-1})$.

(b) *The subdomains* $\{\hat{\Omega}_i^l\}_{i=1}^{N_l}$ *form a finite covering of* Ω, *with a covering constant* N_c, *i.e. we can color* $\{\hat{\Omega}_i^l\}_{i=1}^{N_l}$, *using at most* N_c *colors in such a way that subdomains of the same color are disjoint.*

(c) *There exists a partition of unity* $\{\theta_i^l\}$, *associated with* $\{\hat{\Omega}_i^l\}_{i=1}^{N_l}$, *which satisfies*

$$\sum_i \theta_i^l = 1, \quad \textit{with } \theta_i^l \in H_0^1(\Omega_i^l) \cap C^0(\Omega_i^l), 0 \leqq \theta_i^l \leqq 1 \textit{ and } |\nabla\theta_i^l| \leqq C/h_{l-1} .$$

The first condition simply means that the restriction of the triangulation $\mathcal{T}^l$ to a subdomain $\hat{\Omega}_i^l$ defines a triangulation for $\hat{\Omega}_i^l$ and that the finite element problem on $\hat{\Omega}_i^l$ is well defined. The second condition is used when establishing the upper bound of the spectrum of the additive Schwarz operator. The last condition is used for the lower bound of the spectrum.

One way of constructing subdomains $\{\hat{\Omega}_i^l\}_{i=1}^{N_l}, l = 2, \ldots, L$, with the above properties, is described in Dryja and Widlund [8, 9]. Each element τ_i^{l-1} is extended to a larger region $\hat{\tau}_i^{l-1}$ in such a way that $ch_i^{l-1} \leqq dist(\partial\hat{\tau}_i^{l-1}, \partial\tau_i^{l-1}) \leqq Ch_i^{l-1}$. We align $\partial\hat{\tau}_i^{l-1}$ with the boundaries of level l triangles and cut off the part of $\hat{\tau}_i^{l-1}$ that is outside Ω and use $\hat{\tau}_i^{l-1}$ as the subdomains $\hat{\Omega}_i^l$. Another way of constructing $\{\hat{\Omega}_i^l\}$ is given in Sect. 4.

Let $N_1 = 1$, $V_1^1 = V^1$ and $V_i^l = V^l \cap H_0^1(\hat{\Omega}_i^l)$ for $i = 1, \ldots, N_l, l = 2, \ldots, L$. The finite element space $V^h = V^L$ is represented as a sum

$$V^L = \sum_{l=1}^{L} V^l = \sum_{l=1}^{L} \sum_{i=1}^{N_l} V_i^l.$$

Operators $A: V^L \to V^L$ and $A_{V_i^l}: V_i^l \to V_i^l$, are defined by

$$(Au, \phi) = a(u, \phi), \quad \forall \phi \in V^L, \qquad (A_{V_i^l}u, \phi) = a(u, \phi), \quad \forall \phi \in V_i^l.$$

Let $P_{V_i^l}: V^L \to V_i^l$, $Q_{V_i^l}: V^L \to V_i^l$, be projections defined by

$$a(P_{V_i^l}u, \phi) = a(u, \phi), \quad \forall \phi \in V_i^l,$$

$$(Q_{V_i^l}u, \phi) = (u, \phi), \quad \forall \phi \in V_i^l.$$

The preconditioner B_{MAS}^{-1} and the L-level additive Schwarz operator P_{MAS} are defined by

$$B_{\text{MAS}}^{-1} = \sum_{l=1}^{L} \sum_{i=1}^{N_l} A_{V_i^l}^{-1} Q_{V_i^l}$$

$$P_{\text{MAS}} = B_{\text{MAS}}^{-1} A = \sum_{l=1}^{L} \sum_{i=1}^{N_l} P_{V_i^l}.$$

Algorithm 2.1 (MAS). Find the solution u_h of the finite element equation (1) by solving iteratively the equation

$$P_{\text{MAS}} u_h = f_{\text{MAS}} \stackrel{\text{def}}{=} B_{\text{MAS}}^{-1} f. \tag{2}$$

Note that $f_{\text{MAS}} = B_{\text{MAS}}^{-1} f = \sum_l \sum_i f_i^l$ where $f_i^l = A_{V_i^l}^{-1} Q_{V_i^l}^{-1} f$ are the solutions for the finite element problems

$$a(f_i^l, \phi_h) = a(P_{V_i^l}u, \phi_h) = f(\phi_h), \quad \forall \phi_h \in V_i^l. \tag{3}$$

It is easy to see that for f_i^l given by (3), equations (1) and (2) are equivalent. To find u_h, we first compute the right hand side f_{MAS} by solving (3), and we then use the conjugate gradient (CG) algorithm to solve the system (2). In each iteration, we need to compute $P_{V_i^l} v_h$ for a given $v_h \in V^h$ by solving the equation

$$a(P_{V_i^l} v_h, \phi_h) = a(v_h, \phi_h), \quad \forall \phi_h \in V_i^l.$$

This is a finite element equation on $\hat{\Omega}_i^l$ with mesh size h_l, and $dim(V_i^l) \approx c(h_{l-1}/h_l)^d$. Thus the size of all such problems can be very small.

3. Condition number estimate

When analyzing the CG algorithm for a linear system, the crucial issue is the condition number of the iteration operator. Dryja and Widlund [10] have established the following estimates for the spectrum of P_{MAS}:

$$C_1 L^{-1} a(u, u) \leqq a(P_{\text{MAS}} u, u) \leqq C_2 L\, a(u, u), \quad \forall u \in V^L . \tag{4}$$

Thus $\kappa(P_{\text{MAS}}) \leqq C_2 C_1^{-1} L^2$, i.e. the condition number of P_{MAS} grows at most quadratically with the number of levels. In this section, we improve the bounds in (4) by eliminating the dependence on L.

Theorem 3.1. *The multilevel additive Schwarz operator P_{MAS} satisfies*

$$C_1 a(u, u) \leqq a(P_{\text{MAS}} u, u) \leqq C_2\, a(u, u), \quad \forall u \in V^L .$$

All the constants are independent of $\{h_l\}$ and L.

Lemma 3.1. *Let V be a Hilbert space, V_i be subspaces of V and $V = \sum V_i$. Let P_{V_i} be the projections from V onto V_i, and $P = \sum_i P_{V_i}$. Then*

$$\lambda_{\min}(P) = \min_u \frac{a(u, u)}{a(P^{-1} u, u)} = \min_u \frac{a(u, u)}{\min\limits_{\sum u_i = u} \sum\limits_i a(u_i, u_i)}$$

and

$$\lambda_{\max}(P) = \max_u \frac{a(u, u)}{a(P^{-1} u, u)} = \max_u \frac{a(u, u)}{\min\limits_{\sum u_i = u} \sum\limits_i a(u_i, u_i)} .$$

Proof. For $u = \sum_i u_i$, $u_i \in V_i$, we have

$$a(P^{-1} u, u) = \sum_i a(P^{-1} u, u_i) = \sum_i a(P_{V_i} P^{-1} u, u_i)$$
$$\leqq a(P^{-1} u, u)^{1/2} \Big(\sum_i a(u_i, u_i) \Big)^{1/2} .$$

Thus,

$$a(P^{-1} u, u) = \min_{\sum u_i = u} \sum_i a(u_i, u_i) , \tag{5}$$

since the minimum is achieved for $u_i = P_{V_i} P^{-1} u$. The lemma follows from (5). □

Remark 3.1. If we can find constants C_1 and C_2 such that there exists a decomposition of $u = \sum_i u_i$ satisfying

$$C_1 \sum_i a(u_i, u_i) \leqq a(u, u), \quad \forall u \in V ,$$

and $\forall u \in V$ and for any decomposition of $u = \sum_i u_i$, we have

$$a(u, u) \leqq C_2 \sum_i a(u_i, u_i) ,$$

then, it follows from Lemma 3.1 that $C_1 \leqq \lambda_{\min}(P) \leqq \lambda_{\max}(P) \leqq C_2$. The first part, known as Lions' lemma, is very important in estimating the minimum eigenvalue of P; cf. Dryja and Widlund [8–10] and Nepomnyaschikh [14] for different variants of this result.

For $u, v \in V$, let $\cos(u, v) = a(u, v)/\|u\|_a \|v\|_a$. For V_1, V_2, two nontrivial subspaces of V, we define the cosine of the angle between V_1 and V_2 by

$$\cos(V_1, V_2) \stackrel{\text{def}}{=} \sup_{u_1 \in V_1, u_2 \in V_2} \cos(u_1, u_2)\,. \tag{6}$$

Let $\theta_{ij} = \cos(V_i, V_j)$, $\Phi = \{\theta_{ij}\}$, and let $u = \sum_i u_i$, $u_i \in V_i$, be any decomposition of u. Then,

$$a(u, u) = \sum_{i,j} a(u_i, u_j) \leqq \sum_{i,j} \theta_{ij} |u_i|_a |u_j|_a \leqq \|\Theta\|_2 \sum_i a(u_i, u_i)\,.$$

It follows from Lemma 3.1 that

$$\lambda_{\max}(P) \leqq \|\Theta\|_2\,.$$

Thus, to establish an upper bound for P, it suffices to estimate $\|\Theta\|_2$ for the corresponding space decomposition.

It is obvious that we always have $\cos(V_1, V_2) \leqq 1$, however, stronger results often hold.

Lemma 3.2. *Let $l < k$, $i = 1, 2, \ldots, N_l$, $j = 1, 2, \ldots, N_k$. Then*

$$\cos(V_i^l, V_j^k) \leqq C\gamma^{|k-l-1|d/2}\,.$$

Proof. For $u_i^l \in V_i^l$, $u_j^k \in V_j^k$, we have

$$a(u_i^l, u_j^k) \leqq a_{\Omega_j^k}^{1/2}(u_i^l, u_i^l) a(u_j^k, u_j^k)^{1/2}\,.$$

On an element $\tau^l \subset \Omega_i^l$, $|\nabla u_i^l| = const.$; thus,

$$a_{\hat{\Omega}_j^k \cap \tau^l}(u_i^l, u_i^l) = \frac{mes(\hat{\Omega}_j^k \cap \tau^l)}{mes(\tau^l)} a_{\tau^l}(u_i^l, u_i^l) \leqq C \frac{h_{k-1}^d}{h_l^d} a_{\tau^l}(u_i^l, u_i^l)$$

$$\leqq C\gamma^{d(k-l-1)} a_{\tau^l}(u_i^l, u_i^l)\,.$$

Summing over $\tau^l \subset \Omega_i^l$, we obtain

$$a_{\hat{\Omega}_j^k}(u_i^l, u_i^l) \leqq C\gamma^{d(k-l-1)} a_{\Omega_i^l}(u_i^l, u_i^l) = C\gamma^{d(k-l-1)} a(u_i^l, u_i^l)\,.$$

The lemma follows from the above inequalities. □

Let $\theta_{ij}^{lk} = \cos(V_i^l, V_j^k)$. The matrix,

$$\Theta = \{\theta_{ij}^{lk}\}_{i \leqq N_l,\, j \leqq N_k,\, l,k \leqq L}\,,$$

characterize relations between the subspaces V_i^l. We note that Θ can be partitioned into a L by L block matrix:

$$\Theta = \{\Theta^{lk}\}_{1 \leqq l,k \leqq L}\,,$$

where $\Theta^{lk} = \{\theta_{ij}^{lk}\}_{i \leqq N_l, j \leqq N_k}$ are N_l by N_k submatrices. If we replace the submatrices Θ^{lk} by their l_2-norms, $\|\Theta^{lk}\|_2$, we obtain a L by L matrix

$$\tilde{\Theta} = \{\|\Theta^{lk}\|_2\}_{1 \leqq l, k \leqq L} .$$

The proof of the following lemma is elementary.

Lemma 3.3. *Let $A = \{A_{ij}\}_{1 \leqq i \leqq m, 1 \leqq j \leqq n}$, where the A_{ij} are m_i by n_j submatrices. Let $\tilde{A} = \{\|A_{ij}\|_2\}_{1 \leqq i \leqq m, 1 \leqq j \leqq n}$. Then,*

$$\|A\|_2 \leqq \|\tilde{A}\|_2 .$$

The following lemma provides an estimate of $\|\Theta^{lk}\|_2$.

Lemma 3.4. *For $1 \leqq l, k \leqq L$, we have*

$$\|\Theta^{lk}\|_2 \leqq CN_c \sqrt{\gamma}^{|k-l-1|} . \tag{7}$$

Proof. It is clear that $\|\Theta^{ll}\|_2 \leqq \|\Theta^{ll}\|_1 \leqq N_c$. Assume $l < k$. We first note that $u_i^l \in V^l$ is zero outside Ω_i^l and piecewise harmonic inside Ω_i^l. If $\Omega_j^k \cap \mathscr{E}^l(\Omega_i^l) = \emptyset$, then Ω_j^k is contained either in $\Omega \backslash \Omega_i^l$, where $u_i^l = 0$, or in an element of level l, where u_i^l is a harmonic function; therefore $\cos(V_i^l, V_j^k) = 0$. This implies that $\theta_{ij}^{lk} \neq 0$ only if $\Omega_j^k \cap \mathscr{E}^l(\Omega_i^l) \neq \emptyset$. Thus most elements of matrix Θ^{lk} are zeros and the number of nonzero θ_{ij}^{lk} per row (fixed i) is $O((h_l/h_{k-1})^{d-1}) \leqq C\gamma^{-(k-l-1)(d-1)}$, the number of nonzero θ_{ij}^{lk} per column (fixed j) is bounded by CN_c. Therefore,

$$\sum_{j=1}^{N_k} \operatorname{sign}(\theta_{ij}^{lk}) \leqq C\gamma^{-(k-l-1)(d-1)} \qquad \text{and} \qquad \sum_{i=1}^{N_l} \operatorname{sign}(\theta_{ij}^{lk}) \leqq CN_c .$$

The first inequality implies that

$$\|row_i(\Theta^{lk})\|_2 = \left\{\sum_{j=1}^{N_2} (\theta_{ij}^{lk})^2\right\}^{1/2} \leqq \left\{C\gamma^{d(k-l-1)} \sum_{j=1}^{N_k} \operatorname{sign}(\theta_{ij}^{lk})\right\}^{1/2}$$
$$\leqq \{C\gamma^{d(k-l-1)}\gamma^{-(k-l-1)(d-1)}\}^{1/2} \leqq C\sqrt{\gamma}^{k-l-1} .$$

The second inequality implies that

$$\|\Theta^{lk}\|_2 \leqq CN_c \max_i \|row_i(\Theta^{lk})\|_2 .$$

Inequality (7) follows from these estimates. □

Lemma 3.5. *$\|\Theta\|_2$ can be estimated by*

$$\|\Theta\|_2 \leqq \|\tilde{\Theta}\|_2 \leqq \|\tilde{\Theta}\|_1 \leqq CN_c \frac{1}{1-\sqrt{\gamma}} .$$

Proof. The first inequality follows from Lemma 3.3 with $A = \Theta$, $A_{ij} = \Theta^{lk}$ and $\tilde{A} = \tilde{\Theta}$. Since $\tilde{\Theta}$ is symmetric, the second inequality holds. The last inequality follows from Lemma 3.4 and the definition of the l_1-norm. □

Remark 3.2. The assumption that V^l are linear, bilinear or trilinear elements is not crucial, Lemma 3.5 also holds for general conforming elements.

The upper bound in Theorem 3.1 now follows easily from Lemma 3.5 and Remark 3.1.

Remark 3.3. Although the proof of the upper bound is given for the model problem, it is easy to see that it works for any uniform elliptic operator. Since we can confine our study to one subdomain at a time, we also see that the upper bound is independent of jumps in the coefficients between the subdomains.

To establish the lower bound, we first assume that Ω is convex. We can then use Nitsche's trick to show that the H^1-projection $P_{V^l}: H^1(\Omega) \to V^l$, satisfies the approximation property

$$\|P_{V^l}u - u\|_{L^2(\Omega)} \leqq Ch_l |u|_{H^1(\Omega)}, \quad \forall u \in H^1(\Omega) . \tag{8}$$

Let $P_{V^0} = 0$, $u^l = (P_{V^l} - P_{V^{l-1}})u = (I - P_{V^{l-1}})u^l$. It follows from (8) that

$$\|u^l\|_{L^2(\Omega)} \leqq C h_{l-1} |u^l|_{H^1(\Omega)} . \tag{9}$$

We use the H^1-orthogonal decomposition

$$u = P_{V^L}u = \sum_{l=1}^{L} u^l, \quad u \in V^L \tag{10}$$

and further decompose u^l as

$$u^l = \sum_{i=1}^{N_l} u_i^l, \quad \text{with } u_i^l \equiv \Pi^l(\theta_i^l u^l) \in V_i^l .$$

Here $\Pi^l = \Pi^{h_l}$ is the standard nodal value interpolation operator from $C(\Omega)$ onto V^l and $\{\theta_i^l\}$ a partition of unity as in assumption 2.1. It can be shown that (cf. [8])

$$\begin{aligned}
|u_i^l|^2_{H^1(\hat{\Omega}_i^l)} &= |\Pi^l(\theta_i u^l)|^2_{H^1(\hat{\Omega}_i^l)} \\
&\leqq C(|\theta_i|^2_{L^\infty(\Omega)} |u^l|^2_{H^1(\hat{\Omega}_i^l)} + |\theta_i|^2_{W^{1,\infty}(\Omega)} \|u_i^l\|^2_{L^2(\hat{\Omega}_i^l)}) \\
&\leqq C(|u^l|^2_{H^1(\hat{\Omega}_i^l)} + (1/h_{l-1}^2)\|u_i^l\|^2_{L^2(\hat{\Omega}_i^l)}) .
\end{aligned}$$

Summing over i, using the finite covering property of $\{\hat{\Omega}_i^l\}$, and inequality (9), we obtain

$$\begin{aligned}
\sum_i |u_i^l|^2_{H^1(\Omega)} = \sum_i |u_i^l|^2_{H^1(\hat{\Omega}_i^l)} &\leqq C\sum_i \{|u^l|^2_{H^1(\hat{\Omega}_i^l)} + 1/h_{l-1}^2 \|u^l\|^2_{L^2(\hat{\Omega}_i^l)}\} \\
&\leqq C\{|u^l|^2_{H^1(\Omega)} + 1/h_{l-1}^2 \|u^l\|^2_{L^2(\Omega)}\} \leqq C|u^l|^2_{H^1(\Omega)} .
\end{aligned}$$

Summing over l, $1 \leqq l \leqq L$, and using the orthogonality of u^l, we get

$$\sum_{l=1}^{L} \sum_i |u_i^l|^2_{H^1(\Omega)} \leqq C|u|^2_{H^1(\Omega)} .$$

The lower bound for P_{MAS} now follows from Lemma 3.1.

In the general case, we consider a larger convex region $\tilde{\Omega}$ that contains Ω. We extend the triangulations $\mathcal{T}^l$ to $\tilde{\Omega}$ and denote by $\tilde{V}^l$ the corresponding finite element space defined on $\tilde{\Omega}$ with zero trace on $\partial\tilde{\Omega}$.

For $u \in V^L$, we define $\tilde{u} \in \tilde{V}^L$ by

$$\tilde{u} = \begin{cases} u & x \in \bar{\Omega} \\ 0 & x \in \tilde{\Omega} \backslash \bar{\Omega} \end{cases}.$$

Let $\tilde{u}^l = (P_{\tilde{V}^l} - P_{\tilde{V}^{l-1}})\tilde{u} \in \tilde{V}^l$. Then

$$\tilde{u} = \sum_{l=1}^{L} \tilde{u}^l, \quad \text{in } \tilde{\Omega} \qquad \text{and} \qquad \sum_l \tilde{u}^l = 0 \quad \text{on } \partial\Omega .$$

Since $\tilde{\Omega}$ is convex, we have H^2-regularity, which implies (cf. inequality (9))

$$\sum_l h_{l-1}^{-2} \|\tilde{u}^l\|_{L^2(\tilde{\Omega})}^2 \leqq C \sum_l |\tilde{u}^l|_{H^1(\tilde{\Omega})}^2 = C|\tilde{u}|_{H^1(\tilde{\Omega})}^2 = C|u|_{H^1(\Omega)}^2 . \tag{11}$$

Let $\Omega^l = \bigcup_{\tau \in \mathcal{T}^l, \bar{\tau} \cap \partial\Omega = \emptyset} \tau$ be the union of the interior elements of $\mathcal{T}^l$ and $\Omega \backslash \Omega^l = \bigcup_{\tau \in \mathcal{T}^l, \bar{\tau} \cap \partial\Omega \neq \emptyset} \tau$ be the union of the boundary elements of $\mathcal{T}^l$. We note that

$$\Omega^1 \subset \cdots \subset \Omega^L \subset \Omega^{L+1} \equiv \Omega .$$

We decompose $\tilde{u}^l$ as

$$\tilde{u}^l = u_{\mathrm{I}}^l + u_B^l, \quad u_{\mathrm{I}}^l \in V^l, \; u_B^l \in \tilde{V}^l ,$$

where

$$u_{\mathrm{I}}^l = \begin{cases} \tilde{u}^l & x \in \mathcal{N}^l(\bar{\Omega}^l) \\ 0 & x \in \mathcal{N}^l(\tilde{\Omega} \backslash \bar{\Omega}^l) \end{cases} \quad \text{and} \quad u_B^l = \tilde{u}^l - u_{\mathrm{I}}^l = \begin{cases} 0 & x \in \mathcal{N}^l(\bar{\Omega}^l) \\ \tilde{u}^l & x \in \mathcal{N}^l(\tilde{\Omega} \backslash \bar{\Omega}^l) \end{cases}.$$

Using the discrete norm, it is easy to show that

$$\|u_{\mathrm{I}}^l\|_{L^2(\Omega)}^2 \leqq C h_l^d \sum_{x \in \mathcal{N}^l(\bar{\Omega})} |u_I^l(x)|^2 = C h_l^d \sum_{x \in \mathcal{N}^l(\Omega)} |\tilde{u}^l(x)|^2 \leqq C \|\tilde{u}^l\|_{L^2(\Omega)}^2 , \tag{12}$$

$$\|u_B^l\|_{L^2(\Omega)}^2 \leqq C h_l^d \sum_{x \in \mathcal{N}^l(\bar{\Omega})} |u_B^l(x)|^2 = C h_l^d \sum_{x \in \mathcal{N}^l(\partial\bar{\Omega})} |\tilde{u}^l(x)|^2 \leqq C \|\tilde{u}^l\|_{L^2(\tilde{\Omega})}^2 . \tag{13}$$

Lemma 3.6. *Let $u_B = \sum_l u_B^l$. Then $u_B = 0$ on $\partial\Omega \cup \partial\Omega^1$ and*

$$|u_B|_{H^1(\Omega)}^2 \leqq C|u|_{H^1(\Omega)}^2 . \tag{14}$$

Proof. We note that u and u_I^l vanish on $\partial\Omega$, thus $u_B = u - \sum_l u_I^l = 0$ on $\partial\Omega$. In addition, $u_B^l = 0$ on $\partial\Omega^1$, thus $u_B = \sum_l u_B^l = 0$ on $\partial\Omega^1$. We decompose u_B^l further as $u_B^l = \sum_i v_i^l$, $v_i^l \in V_i^l$. By using Lemma 3.1 and the upper bound of P_{MAS}, we obtain

$$|u_B|_{H^1(\Omega)}^2 \leqq C \sum_l \sum_i |v_i^l|_{H^1(\Omega)}^2 \leqq C \sum_l \sum_i h_l^{-2} \|v_i^l\|_{L^2(\Omega)}^2 \leqq \sum_l h_l^{-2} \|u_B^l\|_{L^2(\Omega)}^2 .$$

Inequality (14) now follows from (13) and (11). □

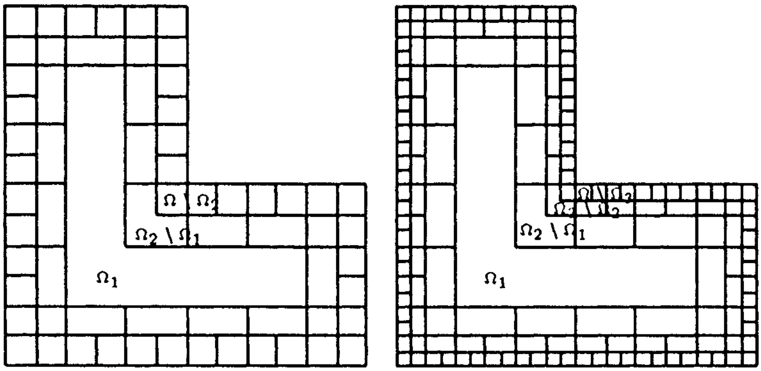

Fig. 1. Composite grid

We decompose the strip $\Omega\backslash\Omega^1 = \Omega^{L+1}\backslash\Omega^1$ into a number of thinner strips (cf. Fig. 1).

$$\Omega\backslash\Omega^1 = (\Omega\backslash\Omega^L) \cup (\Omega^L\backslash\Omega^{L-1}) \cup \cdots \cup (\Omega^2\backslash\Omega^1) = \bigcup_{l=1}^{L}(\Omega^{l+1}\backslash\Omega^l) .$$

The restrictions of $\mathcal{T}^l$ to $\Omega\backslash\Omega^l$, $l = 1, \ldots, L$, define a composite grid on the strip region $\Omega\backslash\Omega^1$. Note that

$$u_B|_{\Omega^l} = \sum_{i=1} w^i \in V^l, \quad l = 2, \ldots, L .$$

Thus, u_B is piecewise linear with respect to this composite grid. This special structure of u_B implies that restricted to region $\Omega\backslash\Omega^1$, there is a unique decomposition of u_B

$$u_B = \sum w^l, \quad w^l \in V^l .$$

Furthermore, $w^l = (\Pi^l - \Pi^{l-1})u_B$. We now show that this unique decomposition is a good decomposition of u_B. We observe that

$$\Pi^1 u_B \equiv 0, \qquad \text{and} \qquad \Pi^l u_B = u_B, \quad x \in \bar{\Omega}^l .$$

This implies that

$$w^l = (\Pi^l - \Pi^{l-1})u_B = 0 \quad \text{in } \Omega^l \cap \Omega^{l-1} = \Omega^{l-1} ,$$

and thus

$$\|w^l\|^2_{L^2(\Omega)} = \|w^l\|^2_{L^2(\Omega\backslash\Omega^{l-1})} \leqq 2(\|\Pi^{l-1}u_B\|^2_{L^2(\Omega\backslash\Omega^{l-1})} + \|\Pi^l u_B\|^2_{L^2(\Omega\backslash\Omega^{l-1})}) .$$

Since $\Pi^{l-1}u_B|_{\partial\Omega^{l-1}} = u_B$ and $\Pi^{l-1}u_B|_{\partial\Omega} = 0$, we have

$$\|\Pi^{l-1}u_B\|^2_{L^2(\Omega\backslash\Omega^{l-1})} \leqq C \sum_{x\in\mathcal{N}^{l-1}(\partial\Omega^{l-1})} h^d_{l-1}|\Pi^{l-1}u_B(x)|^2 \tag{15}$$

$$= C \sum_{x\in\mathcal{N}^{l-1}(\partial\Omega^{l-1})} h^d_{l-1}|u_B(x)|^2$$

$$\leqq C\|u_B\|^2_{L^2(\Omega^{l-1}\backslash\Omega^{l-2})} .$$

Similarly,

$$\|\Pi^l u_B\|^2_{L^2(\Omega\backslash\Omega^l)} \leqq C\|u_B\|^2_{L^2(\Omega^l\backslash\Omega^{l-1})} . \tag{16}$$

Since $\Omega\backslash\Omega^{l-1} = (\Omega\backslash\Omega^l) \cup (\Omega^l\backslash\Omega^{l-1})$ and $\Pi^l u_B|_{\Omega^l\backslash\Omega^{l-1}} = u_B$, we obtain, using (16),

$$\|\Pi^l u_B\|^2_{L^2(\Omega\backslash\Omega^{l-1})} = \|\Pi^l u_B\|^2_{L^2(\Omega\backslash\Omega^l)} + \|u_B\|^2_{L^2(\Omega^l\backslash\Omega^{l-1})} \leqq C\|u_B\|^2_{L^2(\Omega^l\backslash\Omega^{l-1})} . \tag{17}$$

Combining the inequalities (15) and (17) , we obtain

$$\|w^l\|^2_{L^2(\Omega)} = \|w^l\|^2_{L^2(\Omega\backslash\Omega^{l-1})} \leqq C\|u_B\|^2_{L^2(\Omega^l\backslash\Omega^{l-2})} .$$

Multiplying by h_l^{-2}, summing over l, using Hardy's inequality (see e.g. [11] or [15]) and inequality (14), we obtain

$$\begin{aligned}\sum_l h_l^{-2}\|w^l\|^2_{L^2(\Omega)} &\leqq C\sum_l h_l^{-2}\|u_B\|^2_{L^2(\Omega^l\backslash\Omega^{l-2})} \leqq C\sum_l h_l^{-2}\|u_B\|^2_{L^2(\Omega^l\backslash\Omega^{l-1})} \\ &\leqq C\int \frac{u_B^2(x)}{dist(x,\partial\Omega)^2}\,dx \leqq C|u_B|^2_{H^1(\Omega)} \leqq C|u|^2_{H^1(\Omega)} .\end{aligned}$$

Let $u^l = u_1^l + w^l$. Then

$$u = \sum \tilde{u}^l = \sum u_1^l + \sum u_B^l = \sum u_1^l + \sum w^l = \sum u^l, \quad u^l \in V^l ,$$

and

$$\sum_l h_{l-1}^{-2}\|u^l\|^2_{L^2(\Omega)} \leqq 2\sum_l h_{l-1}^{-2}\|u_1^l\|^2_{L^2(\Omega)} + 2\sum_l h_{l-1}^{-2}\|w^l\|^2_{L^2(\Omega)} \leqq C|u|^2_{H^1(\Omega)} .$$

The rest of the proof is identical to that in the case of a convex domain.

Combining our arguments with the extension theorem for finite element functions (cf. Widlund [17]), our approach can be applied to other types of boundary conditions as well.

Remark 3.4. Define a discrete norm $\|\cdot\|_O$ by

$$\|u\|^2_O = \inf_{\sum_{l=1}^L u^l = u} h_{l-1}^{-2}\|u^l\|^2_{L^2}, \quad u^l \in V^l . \tag{18}$$

Using Besov space theory, Oswald [16] has established the following norm equivalency: There exist two constants C_1 and C_2 such that

$$C_1\|u\|^2_O \leqq |u|^2_{H^1} \leqq C_2\|u\|^2_O . \tag{19}$$

Using (19), Oswald proved that the condition number of the BPX operator is $O(1)$.

In [5], Bramble and Pasciak give a different proof of Oswald's result using the properties of L^2 and H^1 projections, see also Xu [20]. Bornemann and Yserentant [4] give another proof using Peetre's K-method.

4. A multilevel diagonal scaling

We begin this section by constructing a special decomposition of the domain Ω. We show that this decomposition, and the corresponding decomposition of the finite element subspaces, satisfies Assumption 2.1. We then demonstrate that the algorithm is a multilevel diagonal scaling, i.e. a natural generalization of the regular diagonal scaling. For problems with constant coefficients and uniform triangulations, the multilevel diagonal scaling algorithm is identical, up to a constant multiple, to the BPX algorithm of Bramble et al. [7]. In the general case, BPX with diagonal scaling results in an MAS algorithm.

Let $\{\mathcal{T}^l\}_{l=1}^L$ be a nested sequence of triangulations, with $\mathcal{T}^{l+1}$ obtained from $\mathcal{T}^l$ by dividing the triangles (rectangles) of $\mathcal{T}^l$ into four triangles (rectangles). In three dimensions, we make a similar construction. We consider piecewise linear, bilinear or trilinear elements, respectively. As in the previous section, the finite element space associated with $\mathcal{T}^l$ is denoted by V^l. Let ϕ_i^l be a nodal basis function of V^l, and associate with each ϕ_i^l a subdomain $\hat{\Omega}_i^l = \text{supp}\{\phi_i^l\}$. We choose $V_i^l = \text{span}\{\phi_i^l\} = V^l \cap H_0^1(\hat{\Omega}_i^l)$ and obtain the decomposition

$$V^L = \sum_{l=1}^{L} \sum_{i=1}^{N_l} V_i^l .$$

The corresponding projections are $Q_{V_i^l}: V^L \xrightarrow{L^2} V_i^l$ and $P_{V_i^l}: V^L \xrightarrow{H^1} V_i^l$. Since $\dim(V_i^l) = 1$, we have, $A_{V_i^l} = \lambda_i^l = a(\phi_i^l, \phi_i^l)/(\phi_i^l, \phi_i^l)$. Using the preconditioner $B_{\text{MAS}}^{-1} = \sum_{l=1}^L \sum_{i=1}^{N_l} 1/\lambda_i^l Q_{V_i^l}$ and the MAS operator $P_{\text{MAS}} = \sum_{l=1}^L \sum_i P_{V_i^l}$, we define an additive Schwarz algorithm

Algorithm 4.1 (MAS). Find the finite element solution $u_h \in V^L$ by solving iteratively the equation

$$P_{\text{MAS}} u_h = f_{\text{MAS}} \stackrel{\text{def}}{=} B_{\text{MAS}}^{-1} f$$

In this special case, we note that $P_{V_i^l} u = a(u, \phi_i^l)/a(\phi_i^l, \phi_i^l)\phi_i^l$ and $f_i^l = 1/\lambda_i^l Q_{V_i^l} f = f(\phi_i^l)/a(\phi_i^l, \phi_i^l)\phi_i^l$. Thus the additive Schwarz equation can be given explicitly by

$$\sum_{l=1}^{L} \sum_{i=1}^{N_l} \frac{a(u, \phi_i^l)}{a(\phi_i^l, \phi_i^l)} \phi_i^l = \sum_{l=1}^{L} \sum_{i=1}^{N_l} \frac{f(\phi_i^l)}{a(\phi_i^l, \phi_i^l)} \phi_i^l .$$

Recall that the BPX preconditioner is given by $B^{-1}v = \sum_{l=1}^L \sum_{i=1}^{N_l} (v, \phi_i^l)\phi_i^l$, therefore $B^{-1}Av = \sum_{l=1}^L \sum_{i=1}^{N_l} a(v, \phi_i^l)\phi_i^l$. It is clear that if we drop the terms $a(\phi_i^l, \phi_i^l)$ from the additive Schwarz equation, we obtain the BPX algorithm.

Define the degree of a vertex x_i as the number of edges incident to it, and the degree of a triangulation $\mathcal{T}^h$ as the maximum of the degrees of its vertices. It is easy to see that the overlapping subdomains $\{\hat{\Omega}_i^l\}_{i=1}^{N_l}$ form a finite covering of Ω with a covering constant less than or equal to degree $(\mathcal{T}^l) + 1$. We also see that on each level $\{\hat{\Omega}_i^l\}$ provides relatively generous overlap. The nodal basis $\{\phi_i^l\}$ can also be used as a partition of unity. The optimal convergence properties of the algorithm follow from Theorem 3.1.

Let K_l be the stiffness matrix associated with V^l and $D_l = diag(K_l)$. Let Π_l: $V^l \to V^L$, $(l \leqq L)$ be the standard interpolation (prolongation) operator, and Π_l^t: $V^L \to V^l$ be the adjoint operator of Π_l. Algorithm 4.1 can then be written as: Find the solution of $K_L x = b$ by solving the preconditioned system

$$B_L^{-1} K_L x = B_L^{-1} b ,$$

where

$$B_L^{-1} = \Pi_1 K_1^{-1} \Pi_1^t + \Pi_2 D_2^{-1} \Pi_2^t + \cdots + \Pi_{L-1} D_{L-1}^{-1} \Pi_{L-1}^t + D_L^{-1}.$$

We note that K_1^{-1} can be replaced by any good preconditioner B_1 of K_1.

If we replace the matrices D_l by identity matrices, we obtain the BPX algorithm. However, since the diagonal elements contain information on the shapes of the triangles and the coefficients of the problems, we expect that the multilevel diagonal preconditioner will work better in practice for non-model problems, since it more closely reflects the properties of the problem.

5. Multiplicative variants

In this section, we discuss some multiplicative variant of the multilevel Schwarz methods, which correspond to certain multigrid methods. In particular, we want to estimate the norm of the following operators

$$E_G = \prod_{l=1}^{L} \prod_{i=1}^{N_l} (I - P_{V_i^l}),$$

$$E_J = \prod_{l=1}^{L} (I - T_l) \stackrel{\text{def}}{=} \prod_{l=1}^{L} (I - \eta \sum_{i=1}^{N_l} P_{V_i^l}),$$

where η is a damping factor such that $\|T_l\| \leqq \omega < 2$.

The algorithms can, in matrix form, be regarded as the error propagation operator of the following schemes:

$$u_l = u_{l-1} + \eta \Pi_l D_l^{-1} \Pi_l^t (f - K_L u_l), \quad l = 1, \ldots, L,$$

or

$$u_l = u_{l-1} + \Pi_l (D_l - L_l)^{-1} \Pi_l^t (f - K_L u_l), \quad l = 1, \ldots, L,$$

where u_0 is the current approximation, D_l, $-L_l$ and L_l^t are the diagonal, lower and upper triangular parts of K_l, respectively. Π_l and Π_l^t are defined as in Sect. 4.

The product in the above expression can be arranged in any order; different orders result in different schemes. When the product is arranged in the natural order, the operators E_G and E_J correspond to the error propagation operators of V-cycle multigrid methods using Gauss-Seidel and damped Jacobi method as smoothers, respectively. We will show that the energy norm of $\|E_G\|$ and $\|E_J\|$ is bounded by a constant less than one, independent of the number of levels and the order of the product. We refer to [5, 6, 20] for general techniques for analyzing multiplicative schemes.

Let $T_i, i = 1, \ldots, N$, be symmetric, semi-positive definite operators, with respect to $a(\cdot,\cdot)$ and let $\|T_i\| \leqq \omega < 2$. Let $T = \sum_{i=1}^{N} T_i$, $E_0 u = u$, $E_i u = (I - T_i) E_{i-1} u$ and $E = E_N$. We first establish bounds for $\|E\|$ in terms of T_i. Then, we apply these estimates to E_G and E_J.

The proof of the following lemma can be found in [6]; the case $T_i = P_i$ is trivial, and equality holds with $\omega = 1$.

Lemma 5.1. *For any $u \in V$,*

$$\sum_{i=1}^{N} a(T_i E_{i-1} u, E_{i-1} u) \leqq \frac{1}{2 - \omega} \{a(u, u) - a(Eu, Eu)\}.$$

Let $\Theta_T = \{\theta_T^{ij}\}$, where $\theta_T^{ii} = 1$ and $\theta_T^{ij}, i \neq j$ are given by

$$\theta_T^{ij} = \sup_{u,v} \frac{a(T_i u, T_j v)}{a(T_i u, u)^{1/2} a(T_j v, v)^{1/2}} . \tag{20}$$

We note that when the T_i are projections, definitions (20) and (6) are identical. As a result of the definition (20), we have

$$a(T_i u, T_j v) \leqq \theta_T^{ij} a(T_i u, u)^{1/2} a(T_j v, v)^{1/2}, \quad i \neq j . \tag{21}$$

and

$$a(T_i u, v) \leqq a(T_i u, u)^{1/2} a(T_i v, v)^{1/2} = \theta_T^{ii} a(T_i u, u)^{1/2} a(T_i v, v)^{1/2}. \tag{22}$$

Lemma 5.2. *For Θ_T defined by* (20), *we have*

$$a(Tu, u) \leqq \|\Theta_T\|_2^2 \sum_{i=1}^{N} a(T_i E_{i-1} u, E_{i-1} u), \quad \forall u \in V.$$

Proof. u can be written as:

$$u = T_1 u + (I - T_1) u = T_1 E_0 u + E_1 u = \cdots = \sum_{j=1}^{i-1} T_j E_{j-1} u + E_{i-1} u . \tag{23}$$

By using (23), (21) and (22), we obtain,

$$\begin{aligned} a(T_i u, u) &= a(T_i u, E_{i-1} u) + \sum_{j=1}^{i-1} a(T_i u, T_j E_{j-1} u) \\ &\leqq \theta_T^{ii} a(T_i u, u)^{1/2} a(T_i E_{i-1} u, E_{i-1} u)^{1/2} \\ &\quad + \sum_{j=1}^{i-1} \theta_T^{ij} a(T_i u, u)^{1/2} a(T_j E_{j-1} u, E_{j-1} u)^{1/2} \\ &= a(T_i u, u)^{1/2} \left\{ \sum_{j=1}^{i} \theta_T^{ij} a(T_j E_{j-1} u, E_{j-1} u)^{1/2} \right\}. \end{aligned}$$

Thus, for $1 \leqq i \leqq N$, we have

$$a(T_i u, u)^{1/2} \leqq \left\{ \sum_{j=1}^{i} \theta_T^{ij} a(T_j E_{j-1} u, E_{j-1} u)^{1/2} \right\} \leqq \left\{ \sum_{j=1}^{N} \theta_T^{ij} a(T_j E_{j-1} u, E_{j-1} u)^{1/2} \right\} .$$

The lemma follows by taking the l_2-norm. □

As a consequence of Lemma 5.1 and Lemma 5.2, we have

$$a(u, u) \leqq \frac{1}{\lambda_{\min}(T)} a(Tu, u) \leqq \frac{1}{2 - \omega} \frac{\|\Theta_T\|_2^2}{\lambda_{\min}(T)} (\|u\|_a^2 - \|Eu\|_a^2) .$$

Solving for $\|Eu\|$, we obtain

Lemma 5.3.

$$\|Eu\|_a^2 \leqq \left(1 - (2 - \omega) \frac{\lambda_{\min}(T)}{\|\Theta_T\|_2^2}\right) \|u\|_a^2 .$$

The next lemma gives a lower bound of $\|E\|$.

Lemma 5.4.

$$1-\frac{2\lambda_{\min}(T)}{2-\omega+\lambda_{\min}(T)} \leqq \|E\| .$$

Proof. Using identity (23) and the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}
\|u\|_a^2 = a(u,u) &= a\left(u, Eu + \sum_{i=1}^N T_i E_{i-1} u\right) \\
&\leqq \|u\|_a \|Eu\|_a + \sum_{i=1}^N a(T_i u, u)^{1/2} a(T_i E_{i-1} u, E_{i-1} u)^{1/2} \\
&\leqq \|u\|_a \|Eu\|_a + a(Tu, u)^{1/2} \left[\sum_{i=1}^N a(T_i E_{i-1} u, E_{i-1} u)\right]^{1/2} \\
&\leqq \|u\|_a \|Eu\|_a + a(Tu, u)^{1/2} \left[\frac{\|u\|_a^2 - \|Eu\|_a^2}{2-\omega}\right]^{1/2} .
\end{aligned}$$

Dividing both sides by $\|u\|_a^2$ results in

$$1 \leqq \frac{\|Eu\|_a}{\|u\|_a} + \left[\frac{1}{2-\omega}\right]^{1/2} \left[\frac{a(Tu,u)}{a(u,u)}\right]^{1/2} \left[1 - \frac{\|Eu\|_a^2}{\|u\|_a^2}\right]^{1/2} .$$

Simple algebra gives

$$(2-\omega)\frac{1-\|E\|}{1+\|E\|} \leqq (2-\omega) \frac{1-\dfrac{\|Eu\|_a}{\|u\|_a}}{1+\dfrac{\|Eu\|_a}{\|u\|_a}} \leqq \frac{a(Tu,u)}{a(u,u)} ,$$

which shows that

$$\lambda_{\min}(T) \geqq (2-\omega)\frac{1-\|E\|}{1+\|E\|} . \tag{24}$$

Solving for $\|E\|$, we obtain

$$\|E\| \geqq \frac{2-\omega-\lambda_{\min}(T)}{2-\omega+\lambda_{\min}(T)} = 1 - \frac{2\lambda_{\min}(T)}{2-\omega+\lambda_{\min}(T)} . \tag{25}$$

The lemma follows from (25). $\square$

Combining Lemma 5.3 and Lemma 5.4, we obtain

Theorem 5.1.

$$1-\frac{2\lambda_{\min}(T)}{2-\omega+\lambda_{\min}(T)} \leqq \|E\| \leqq \sqrt{1-(2-\omega)\frac{\lambda_{\min}(T)}{\|\Theta_T\|_2^2}} .$$

Theorem 5.1 implies that if $\lambda_{\min}(T) \to 0$ then $\|E\| = 1 - O(\lambda_{\min}(T)) \to 1$. If we can show that $\|E\| \leqq \delta_E < 1$, independently of the number of refinement levels, then, it follows from (24) that $\lambda_{\min}(T)$ is bounded from below by a constant independent of the number of levels.

To apply Theorem 5.1 to operators E_J and E_G, we need to estimate the corresponding $\lambda_{\min}(T)$ and $\|\Theta_T\|_2$.

We first estimate $E_G = \prod_{l=1}^{L}\prod_{i=1}^{N_l}(I - P_{V_i^l})$. In this case, $T_{\iota} = P_{V_j^l}$ for some subspace V_j^l. Thus, $T = P_{\text{MAS}}$ and $\Theta_T = \Theta = \{\cos(V_i^l, V_j^k)\}$. By Lemma 3.5, we know that $\|\Theta\|_2 \leqq C_2$ and that $\lambda_{\min}(T) \geqq C_1$. Thus, $\|E_G\| \leqq \delta_G < 1$, i.e. the V-cycle multigrid method using Gauss-Seidel as a smoother has a rate of convergence independent of the number of levels.

We now estimate $E_J = \prod_{l=1}^{L}(I - \eta\sum_i P_{V_i^l})$. We note that $T_l = \eta\sum_i P_{V_i^l}$. Thus, $T = \sum T_l = \eta P_{\text{MAS}}$ and $\lambda_{\min}(T) = \eta\lambda_{\min}(P_{\text{MAS}}) \geqq C$. To estimate $\|\Theta_T\|_2$, we note that

$$\begin{aligned} a(T_l u, T_k v) &= \eta^2 \sum_i \sum_j a(P_{V_i^l} u, P_{V_j^k} v) \\ &\leqq \eta^2 \sum_i \sum_j \theta_{ij}^{lk} |P_{V_i^l} u|_a |P_{V_j^k} v|_a \\ &\leqq \eta \|\Theta^{lk}\|_2 \left(\eta \sum_i |P_{V_i^l} u|_a^2\right)^{1/2} \left(\eta \sum_j |P_{V_j^k} v|_a^2\right)^{1/2} \\ &= \eta \|\Theta^{lk}\|_2 a(T_l u, u)^{1/2} a(T_k v, v)^{1/2} . \end{aligned}$$

By the definition of θ_T^{lk} and Lemma 3.4, we have

$$\theta_T^{lk} \leqq \eta \|\Phi^{lk}\|_2 \leqq C N_c \sqrt{\gamma}^{\,|k-l-1|}.$$

Therefore,

$$\|\Theta_T\|_2 \leqq \|\Theta_T\|_1 = \max_l \sum_k \theta_T^{lk} \leqq C N_c \frac{1 + \sqrt{\gamma}}{1 - \sqrt{\gamma}} .$$

Thus, $\|E_J\| \leqq \delta_J < 1$, i.e. the V-cycle multigrid method using damped Jacobi as a smoother, has a rate of convergence independent of the number of levels.

Remark 5.1. Similar techniques can be used to study the hierarchical basis multigrid method. Let E_{HBMG} be the error propagation operator of the hierarchical basis multigrid method, using Gauss-Seidel or damped Jacobi method as smoothers. Then

$$\|E_{\text{HBMG}}\|_a^2 \leqq 1 - C\frac{\lambda_{\min}(K_{\text{HB}})}{\|\Theta_T\|_2^2} \leqq 1 - CL^{-2} .$$

Here, K_{HB} is the hierarchical basis stiffness matrix.

6. Numerical experiments

In this section, we report on some numerical experiments with multilevel additive Schwarz methods. These experiments were carried out for Poisson's equation on a unit square with homogeneous Dirichlet boundary conditions

$$\begin{cases} -\Delta u = f(x) & \text{in} \quad \Omega , \\ u = 0 & \text{on} \quad \partial\Omega . \end{cases}$$

We divide the domain Ω into $k \times k$ square elements $\tau_i^1, i = 1, \ldots, k^2$, and obtain a triangulation $\mathcal{T}^1 = \{\tau_i^1\}$. We then divide each τ_i^1 into $k \times k$ squares to

Table 1. Properties of the multilevel additive Schwarz scheme, using bilinear elements

Level	Unknowns	h_{l-1}/h_l	ovlp ratio	$\kappa(P_{MAS})$	Iter. no.
2	$(4-1)^2$	2	1/2	4.5	5
3	$(8-1)^2$	2	1/2	7.2	11
4	$(16-1)^2$	2	1/2	9.3	17
5	$(32-1)^2$	2	1/2	10·9	20
6	$(64-1)^2$	2	1/2	11.8	21
7	$(128-1)^2$	2	1/2	12.6	23
2	$(9-1)^2$	3	1/3	4.7	9
3	$(27-1)^2$	3	1/3	7.1	16
4	$(81-1)^2$	3	1/3	8.5	19
5	$(243-1)^2$	3	1/3	9.5	21
2	$(16-1)^2$	4	1/4	5.1	13
3	$(64-1)^2$	4	1/4	7.3	17
4	$(256-1)^2$	4	1/4	8.5	20
2	$(25-1)^2$	5	1/5	5.7	14
3	$(125-1)^2$	5	1/5	7.8	17

obtain the triangulation $\mathscr{T}^2 = \{\tau_i^2\}$, etc. The length of an edge of τ_i^l is denoted by h_l and $h_l = (1/k)^l$. For $l = 2, \ldots, L$, we extend τ_i^{l-1} to a larger square $\hat{\tau}_i^{l-1}$. The overlap ratio

$$\text{overlap ratio} = \frac{dist\{\partial\hat{\tau}_i^{l-1}, \partial\tau_i^{l-1}\}}{h_{l-1}}$$

measures the width of $\hat{\tau}_i^{l-1}\backslash\tau_i^{l-1}$ in terms of h_{l-1}, the side of the square τ_i^{l-1}. We use Ω as our domain for $l = 1$, and $\Omega_i^l = \hat{\tau}_i^{l-1}$ as our subdomains for $l = 2, \ldots, L$.

In these experiments, we take $k = 2, 3, 4$ or 5, and $\hat{\tau}_i^{l-1}\backslash\tau_i^{l-1}$ is one element (h_l) wide, i.e. the overlap ratio is $1/k$. Therefore, we only need to solve very small linear systems of order 9, 16, 25 or 36, respectively. We use the conjugate gradient method to solve the system $P_{MAS}u_h = f_{MAS}$ iteratively. The last column of the table gives the number of iterations required to decrease the l_2-norm of the residual by a factor of $\varepsilon = 10^{-6}$.

In the next set of experiment, we report some numerical results for Algorithm 4.1; see also [7]. In Table 2, we report results for the linear elements

Table 2. BPX, using linear elements

Level	Unknowns	$\lambda_{min}(P_{MAS})$	$\lambda_{max}(P_{MAS})$	$\kappa(P_{MAS})$	Iter. no.
2	$(4-1)^2$	0.61	1.75	2.9	4
3	$(8-1)^2$	0.51	2.66	5.3	13
4	$(16-1)^2$	0.47	3.29	7.0	17
5	$(32-1)^2$	0.46	3.81	8.2	19
6	$(64-1)^2$	0.46	4.23	9.2	20
7	$(128-1)^2$	0.46	4.58	9.9	20
8	$(256-1)^2$	0.46	4.88	10.6	21
9	$(512-1)^2$	0.46	5.13	11.2	21

Table 3. BPX, using bilinear elements

Level	Unknowns	$\lambda_{\min}(P_{\mathrm{MAS}})$	$\lambda_{\max}(P_{\mathrm{MAS}})$	$\kappa(P_{\mathrm{MAS}})$	Iter. no.
2	$(4-1)^2$	0.82	1.7	2.1	3
3	$(8-1)^2$	0.77	2.3	3.0	8
4	$(16-1)^2$	0.76	2.7	3.6	11
5	$(32-1)^2$	0.76	3.1	4.0	11
6	$(64-1)^2$	0.76	3.3	4.4	13
7	$(128-1)^2$	0.75	3.6	4.7	14
8	$(256-1)^2$	0.75	3.8	5.0	14
9	$(512-1)^2$	0.75	3.9	5.3	14
10	$(1024-1)^2$	0.75	4.1	5.5	14
11	$(2048-1)^2$	0.75	4.2	5.6	14

discretization. In Table 3, we summarize the result for the bilinear elements discretization.

Acknowledgements. I would like to thank Professor Olof Widlund for suggesting the problem and for all of his excellent advice and critical comments on this manuscript. I would also like to thank Professors Xiao-Chuan Cai and Howard Elman for valuable suggestions.

References

1. Bank, R.E., Dupont, T.F., Yserentant, H. (1988): The hierarchical basis multigrid method. Numer. Math. **52**, 427–458
2. Bjørstad, P.E., Moe, R., Skogen, M. (1990): Parallel domain decomposition and iterative refinement algorithms. In: W. Hackbusch, ed., Parallel Algorithms for PDEs. Proceedings of the 6th GAMM-Seminar held in Kiel, Germany, January 19–21, 1990. Vieweg, Braunschweig, Wiesbaden
3. Bjørstad, P.E., Skogen, M. (1992): Domain decomposition algorithms of Schwarz type, designed for massively parallel computers. In: T.F. Chan, D.E. Keyes, G.A. Meurang, J.S. Scroggs, R.G. Voigt, eds., Fifth International Symposium on Domain Decomposition Methods for Partial Differential Equations, SIAM, Philadelphia, PA (to appear)
4. Bornemann, F., Yserentant, H. (1992): A basic norm equivalence for the theory of multilevel methods. Numer. Math. (in press)
5. Bramble, J.H., Pasciak, J.E. (1991): New estimates for multilever algorithms including the V-cycle. Tech. Report BNL-46730, Department of Applied Science, Brookhaven National Laboratory
6. Bramble, J.H., Pasciak, J.E., Wang, J., Xu, J. (1991): Convergence estimates for product iterative methods with applications to domain decomposition. Math. Comput. **57**, 1–21
7. Bramble, J.H., Pasciak, J.E., Xu, J. (1990): Parallel multilever preconditioners. Math. Comput. **55**, 1–22
8. Dryja, M., Widlund, O.B. (1987): An additive variant of the Schwarz alternating method for the case of many subregions. Tech. Report 339, also Ultracomputer Note 131, Department of Computer Science, Courant Insitute
9. Dryja, M., Widlund, O.B. (1989): Some domain decomposition algorithms for alliptic problems. In: L. Hayes, D. Kincaid, eds., Iterative methods for large linear systems. Academic Press, San Diego, California, pp. 273–291. Proceeding of the Conference on Iterative Methods for Larger Linear Systems held in Austin, Texas, October 19–21, 1988, to celebrate the sixty-fifth birthday of David M. Young, Jr.
10. Dryja, M., Widlund, O.B. (1987): Multilevel additive methods for elliptic finite element problems. In: W. Hackbusch, ed., Parallel algorithms for partial differential equations. Proceedings of the Sixth GAMM-Seminar, Kiel, January 19–21, 1990. Vieweg, Braunschweig

11. Hardy, G.H., Littlewood, J.E., Pólya, G. (1934): Inequalities. Cambridge University Press, Cambridge
12. Hackbusch, W. (ed.) (1985): Multi-grid methods and applications. Springer, Berlin Heidelberg New York
13. McCormick, S.F. (ed) (1987): Multigrid methods. SIAM, Philadelphia, PA
14. Nepomnyaschikh, S.V. (1987): On the application of the method of bordering for elliptic mixed boundary value problems and on the difference norms of $W_2^{1/2}(S)$. Tech. Report 106, Computing Center of the Siberian Branch of the USSR Academy of Sciences, Novosibirsk [in Russian]
15. Nečas, J. (1967): Les méthodes directes in théorie des équations elliptiques. Academia, Prague
16. Oswald, P. (1992): On discrete norm estimates related to multilevel preconditioners in the finite element method. In: Proc. Int. Conf. Theory of Functions, Varna **91** (to appear)
17. Widlund, O.B. (1987): An extension theorem for finite element spaces with three applications. In: W. Hackbusch, K. Witsch, eds., Numerical techniques in continuum mechanics. Notes on numerical fluid mechanics. vol. 16. Proceedings of the Second GAMM-Seminar, Kiel, January, 1986. Vieweg, Braunschweig Weisbaden, pp. 110–122
18. Widlund, O.B. (1993): Some Schwarz methods for symmetric and nonsymmetric elliptic problem. In: T.F. Chan, D.E. Keyes, G.A. Meurant, J.S. Scroggs, R.G. Voigt, eds., Fifth Conference on Domain Decomposition Methods for Partial Differential Equations (to appear)
19. Xu, J. (1989): Theory of Multilevel Methods. Ph.D. thesis, Cornell University
20. Xu, J. (1990): Iterative methods by space decomposition and subspace correction. Tech. report, Penn State University, University Park, PA (revised, 1992) (accepted by SIAM Rev.)
21. Yserentant, H. (1986): On the multi-level splitting of finite element spaces. Numer. Math. **49**, 412
22. Zhang, X. (1991): Multilevel additive Schwarz methods. Tech. Report 582, Courant Institute of Mathematical Sciences, Department of Computer Science
23. Zhang, X. (1991): Studies in domain decomposition: Multilevel methods and the biharmonic Dirichlet problem. PhD thesis, Courant Institute, New York University