Additive Schwarz Algorithms for Parabolic Convection-Diffusion Equations *

Xiao-Chuan Cai †

Abstract

In this paper, we consider the solution of linear systems of algebraic equations that arise from parabolic finite element problems. We introduce three additive Schwarz type domain decomposition methods for general, not necessarily selfadjoint, linear, second order, parabolic partial differential equations and also study the convergence rates of these algorithms. The resulting preconditioned linear system of equations is solved by the generalized minimal residual method. Numerical results are also reported.

Key words Schwarz's alternating method, domain decomposition, parabolic convection-diffusion equation, finite elements.

AMS(MOS) subject classifications. 65N30, 65F10

1 Introduction

Domain decomposition techniques are powerful iterative methods for solving linear systems of equations that arise from discretizing partial differential equations. A systematic theory has been developed for elliptic finite element problems in the past few years, see [2, 3, 4, 5, 6, 7, 8, 9, 12, 13, 21] etc. In this paper, we are interested in solving the finite element parabolic convection-diffusion problems, obtained by using implicit schemes, such as the backward Euler scheme and the Crank-Nicolson scheme, in the time variable, and Galerkin approximation in the space variables. At a fixed time level, the resulting equation is equivalent to an elliptic problem which depends on a time step parameter. This suggests that we might apply the methods, originally proposed for elliptic equations, to the parabolic cases. The central mathematical question is then to estimate how the convergence rate depends on the space mesh and the time step parameters, especially in the case when the latter parameter is relatively large. Other domain decomposition methods for parabolic problems can be found in [11, 17, 20] and references therein.

*Appeared in Numer. Math., Vol. 60, (1990) pp. 41-62.

†Department of Mathematics, University of Kentucky, Lexington, KY 40506. This work was supported in part by the National Science Foundation under Grant NSF-CCR-8903003 at the Courant Institute, New York University and in part by the National Science Foundation under contract number DCR-8521451 and ECS-8957475 at Yale University.

The outline of this paper is as follows. In the remainder of this section, we present the continuous and the discrete parabolic convection-diffusion equation, and also discuss some of their basic properties. In section 2, we outline an abstract additive Schwarz theory; cf. [8], without providing any detailed analysis. In section 3, we introduce three additive Schwarz type algorithms, which are instances of the abstract theory. The first one is an straightforward generalization of the additive Schwarz method for elliptic equation [9, 13], the second one is obtained by eliminating the coarse mesh space and using only local function spaces to take the advantage of the existance of a time step parameter. This algorithm is therefore more suitable for parallel computers than the first one. The third one is an non-overlapping domain decomposition method, but nevertheless fits into our general theory of additive Schwarz methods. Finally, in section 4, we report on some of our numerical experiments.

Throughout this paper, c and C, with or without subscripts, denote generic, strictly positive constants. Their values may be different at different occurrences, but are independent of the mesh parameters H, h and τ, which will be introduced later.

1.1 A parabolic convection-diffusion problem

We consider the following parabolic convection-diffusion problem: Find $u(x,t)$, such that

$$\left\{\begin{array}{rcll} \partial u/\partial t \ + \ Lu & = & f & \text{in} \quad \Omega \times [0,T], \\ u(x,t) & = & 0 & \text{on} \quad \partial\Omega \times [0,T], \\ u(x,0) & = & u_0(x) & \text{in} \quad \Omega, \end{array}\right. \tag{1}$$

where $\Omega \subset \mathbb{R}^d$, $d = 2$ or 3, is a polygonal domain with boundary $\partial\Omega$. L is a strongly elliptic operator, which has the following form

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

All coefficients are sufficiently smooth and $a_{ij}(x) = a_{ji}(x)$ for all i, j and x.

A weak formulation of equation (1) is: Find $u(x,t) \in H_0^1(\Omega)$, $u(x,0) = u_0(x)$ in Ω, such that

$$(\frac{\partial u}{\partial t}, v) \ + \ B(u,v) \ = \ (f,v), \ \ \forall v \in H_0^1(\Omega), \ \forall t \in [0,T].$$

The bilinear form $B(\cdot,\cdot)$ is defined as

$$B(u,v) = \sum_{i,j=1}^{d} \int_\Omega a_{ij}\frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_j}dx + \sum_{i=1}^{d}\int_\Omega b_i\frac{\partial u}{\partial x_i}vdx \ + \int_\Omega cuvdx$$

and the linear functional

$$(f,v) \ = \ \int_\Omega fvdx.$$

We assume that the bilinear form is
(1) bounded, i.e. $| B(u,v) | \leq C\|u\|_{H_0^1(\Omega)}\|v\|_{H_0^1(\Omega)}, \ \ \forall u,v \in H_0^1(\Omega)$, and
(2) elliptic, i.e. $| B(u,u) | \geq c\|u\|^2_{H_0^1(\Omega)}, \ \ \forall u \in H_0^1(\Omega)$.

The existence and uniqueness of the solution of the weak parabolic convection-diffusion equation present no problems, see [19]. We use two types of time discretization, namely, a backward Euler scheme and a Crank-Nicolson scheme. Let Δt_n be the n^{th} time step, M the number of steps and $\sum_{n=1}^{M} \Delta t_n = T$. For the first scheme, the time discrete problem is

$$\left(\frac{u^n - u^{n-1}}{\Delta t_n}, v \right) + B(u^n, v) = (f, v), \quad \forall v \in H_0^1(\Omega),$$

with $u^0(x,t) = u_0(x)$ and $n = 1, \cdots, M$. For the second scheme, we have

$$\left(\frac{u^n - u^{n-1}}{\Delta t_n}, v \right) + B\left(\frac{u^n + u^{n-1}}{2}, v\right) = (f, v), \quad \forall v \in H_0^1(\Omega),$$

with $u^0(x,t) = u_0(x)$ and $n = 1, \cdots, M$. Both schemes lead to the following problem: For a given function $g_{n-1} \in H^{-1}(\Omega)$, find $w \in H_0^1(\Omega)$, such that

$$D_\tau(w,v) \equiv (w,v) + \tau B(w,v) = (g_{n-1}, v), \ \ \forall v \in H_0^1(\Omega), \tag{2}$$

where τ is the time step parameter. For the backward Euler scheme

$$\begin{aligned} w &= u^n - u^{n-1}, \\ \tau &= \Delta t_n, \\ (g_{n-1}, v) &= \tau((f,v) - B(u^{n-1}, v)), \end{aligned}$$

and for the Crank-Nicolson scheme,

$$\begin{aligned} w &= u^n - u^{n-1}, \\ \tau &= \Delta t_n/2, \\ (g_{n-1}, v) &= \tau(2(f,v) - B(u^{n-1}, v)). \end{aligned}$$

The stability of both schemes is well understood, see [19]. In this paper, we focus on the study of fast iterative algorithms for solving the linear systems at each time step. In general, $D_\tau(\cdot,\cdot)$ is not symmetric. For technical reasons, it is convenient to separate the symmetric and the skewsymmetric parts of the bilinear form. We therefore introduce the bilinear forms $A_\tau(u,v) = 1/2(D_\tau(u,v) + D_\tau(v,u))$, which is symmetric, and $N_\tau(u,v) = 1/2(D_\tau(u,v) - D_\tau(v,u))$, which is skewsymmetric. It is easy to prove that $N_\tau(\cdot,\cdot)$ is bounded in the following sense: There exists a constant C, such that

$$| N_\tau(u,v) | \leq C\tau\|u\|_{H_0^1(\Omega)}\|v\|_{L^2(\Omega)}, \ \ \forall u,v \in H_0^1(\Omega).$$

Define the τ-norm by $\|\cdot\|^2_{\tau(\Omega)} = \|\cdot\|^2_{L^2(\Omega)} + \tau\|\cdot\|^2_{H_0^1(\Omega)}$. The following lemma gives the boundness and positive definiteness of $D_\tau(\cdot,\cdot)$.

Lemma 1 *There exist constants C and $c > 0$, which are independent of τ, such that*

(1) $| D_\tau(u,v) | \le C\|u\|_{\tau(\Omega)}\|v\|_{\tau(\Omega)}, \ \ \forall u,v \in H_0^1(\Omega)$

(2) $D_\tau(u,u) \ge c\|u\|^2_{\tau(\Omega)}, \ \ \forall u \in H_0^1(\Omega)$

(3) $| A_\tau(u,v) | \le C\|u\|_{\tau(\Omega)}\|v\|_{\tau(\Omega)}, \ \ \forall u,v \in H_0^1(\Omega)$

(4) $A_\tau(u,u) \ge c\|u\|^2_{\tau(\Omega)}, \ \ \forall u \in H_0^1(\Omega)$.

The proof follows immediately from the boundedness and ellipticity of $B(\cdot,\cdot)$.

If we denote $\|u\|_{A_\tau} = \sqrt{A_\tau(u,u)}$, then we know from this lemma that the A_τ-norm is equivalent to the τ-norm. We shall use the A_τ-norm in our discussion. It can be shown, cf Grisvard [18] and Nečas [22], that the solution of $B(u,v)$ is $H^{1+\gamma}(\Omega)$-regular in the following sense. For any $g \in L^2(\Omega)$, there exists a solution $u \in H^{1+\gamma}(\Omega) \cap H_0^1(\Omega)$ of

$$B(v,u) \ = \ (g,v), \ \ \forall v \in H_0^1(\Omega)$$

with

$$| \ u \ |_{H^{1+\gamma}(\Omega)} \le C\|g\|_{L^2(\Omega)},$$

where $\gamma \in [1/2, 1]$. Based on this regularity result, we have

Lemma 2 *For any $g \in L^2(\Omega)$, the equation*

$$D_\tau(v,u) \ = \ (g,v), \ \ \forall v \in H_0^1(\Omega)$$

has a solution $u \in H^{1+\gamma}(\Omega) \cap H_0^1(\Omega)$ that satisfies $| \ u \ |_{H^{1+\gamma}(\Omega)} \le C/\tau \|g\|_{L^2(\Omega)}$.

Proof: Consider the equality

$$(u,u) + \tau B(u,u) = (g,u).$$

Since $B(u,u) \ge 0$, we have $\|u\|_{L^2(\Omega)} \le \|g\|_{L^2(\Omega)}$, which implies that $\|u - g\|_{L^2(\Omega)} \le 2\|g\|_{L^2(\Omega)}$. On the other hand, we have

$$B(v,u) = ((g-u)/\tau, v), \quad \forall v \in H_0^1(\Omega).$$

By using the regularity result of $B(\cdot,\cdot)$, we complete the proof. □

1.2 A finite element approximation

We solve equation (2) by a conformal Galerkin finite element method. For simplicity, we use piecewise linear triangular elements in $\mathbb{R}^2$ and the corresponding tetrahedral elements in $\mathbb{R}^3$. In this subsection, we introduce a two level triangulation of Ω, the corresponding finite element spaces and the Galerkin equation.

For a given polygonal region $\Omega \in \mathbb{R}^d$, let $\{\Omega_i\}$ be a shape regular finite element triangulation of Ω and H the maximal diameter of these Ω_i's. We sometimes refer Ω_i as a substructure and $\{\Omega_i\}$ as the coarse mesh or H-level triangulation of Ω.

We further divide each substructure Ω_i into smaller triangles, denoted as τ_i^j, $j = 1, \cdots$. We assume that $\{\tau_i^j\}$ form a shape regular finite element triangulation of Ω and h is the maximal diameter of τ_i^j. We call $\{\tau_i^j\}$ the fine mesh or h-level triangulation of Ω.

We next define the piecewise linear finite element function spaces over both H-level and h-level triangulations of Ω.

$$V^H = \{v^H \mid \text{ continuous on } \Omega,\ v^H \mid_{\Omega_i} \text{ linear on } \Omega_i,\ \forall i,\ v^H = 0 \text{ on } \partial\Omega\}$$

and

$$V^h = \{v^h \mid \text{ continuous on } \Omega,\ v^h \mid_{\tau_i^j} \text{ linear on } \tau_i^j,\ \forall i,\ v^h = 0 \text{ on } \partial\Omega\}$$

It is obvious that $V^H \subset V^h$. We also use the notations

$$\Lambda^h = \{x \mid \in \text{interior nodes of } h - \text{level subdivision}\}$$

and

$$\Lambda^H = \{x \mid \in \text{interior nodes of } H - \text{level subdivision}\}.$$

To obtain a fully discrete problem, we discretize equation (2) in space by using a Galerkin finite element method and the subspace $V^h \subset H_0^1(\Omega)$. The Galerkin approximation of equation (2) reads as follows: Find $\mathbf{u}^h \in V^h$, such that

$$D_\tau(\mathbf{u}^h, v^h) = (g, v^h), \ \ \forall v^h \in V^h. \tag{3}$$

Here g is different for different time discretizations. The existence and uniqueness of $\mathbf{u}^h$ has been extensively studied in the literature, see [1, 19]. By using the nodal basis functions, equation (3) can be transformed into a linear system of equations, which is usually large and sparse, but not well conditioned. It is well known that the efficiency of any iterative methods to be used to solve this linear system of equations depends strongly on the conditioning of the stiffness matrix. The main focus of this paper is to provide some effective preconditioning techniques for solving this equation.

2 An abstract theory for the additive Schwarz method

In this section, we present an abstract theory developed in [8] for the additive Schwarz method. Interested reader should see [8] for detailed proof. Let V be a Hilbert space of real functions defined on $\Omega \subset \mathbb{R}^d$ with an inner product $(u, v)_V$ and the corresponding norm $\|u\|_V$. Let ω be a subdomain of Ω. We define $(u, v)_{V(\omega)}$ to be the restriction of the inner product to ω.

Let $\mathcal{B}(\cdot,\cdot)$ be a bilinear form on $V \times V$ and $\mathcal{F}(\cdot)$ a linear functional on V such that

• $\mathcal{B}(\cdot,\cdot)$ is continuous; in particular $\mid \mathcal{B}(u, v) \mid \leq C\|u\|_{V(\omega)}\|v\|_{V(\omega)}, \ \ \forall u, v \in V$, where $\omega = \{\text{supp } u\} \cap \{\text{supp } v\}$.

• $\mathcal{B}(\cdot,\cdot)$ is V-elliptic, i.e. $\mathcal{B}(u, u) \geq c\|u\|_V^2, \ \ \forall u \in V$.

• $\mathcal{F}(\cdot)$ is continuous, i.e. $\mid \mathcal{F}(u) \mid \leq C\|u\|_V, \ \ \forall u \in V$.

We define $\mathcal{A}(u, v) = 1/2(\mathcal{B}(u, v) + \mathcal{B}(v, u))$, a symmetric bilinear form, and $\mathcal{N}(u, v) = 1/2(\mathcal{B}(u, v) - \mathcal{B}(v, u))$, which is skewsymmetric.

It follows from the assumptions on $\mathcal{B}(u, v)$ that $\mathcal{A}(u, v)$ is elliptic and continuous in the same sense as $\mathcal{B}$. This implies that the norm corresponding to $\mathcal{A}(\cdot,\cdot)$ is equivalent to the V-norm. In the following, we can therefore use $(\cdot,\cdot)_\mathcal{A}$, instead of $(\cdot,\cdot)_V$.

We are interested in solving the problem that reads as follows: Find $\mathbf{u} \in V$, such that

$$\mathcal{B}(\mathbf{u}, v) \;=\; \mathcal{F}(v), \quad \forall v \in V. \tag{4}$$

Let $V_i, \quad i = 0, \cdots, N$, be subspaces of V, such that $V \;=\; V_0 + \cdots + V_N$, i.e., for any $v \in V$, there exist $v_i \in V_i, \quad i = 0, \cdots, N$, such that $v = v_0 + \cdots + v_N$. Moreover, we assume that there exists a constant C_0^2 such that

$$\sum_{i=0}^{N} \|v_i\|^2_{V(\omega_i)} \leq C_0^2 \|v\|^2_V, \quad \forall v \in V, \tag{5}$$

where ω_i is the support of V_i, i.e. the union of the supports of all functions in V_i.

This is usually called the bounded decomposition lemma. Note that the constant C_0^2 may depend on the number of subregions N and also on some other parameters of V, which may be introduced in particular applications.

We also assume that there exists a constant C_ω, such that

$$\sum_{i=0}^{N} \|v\|^2_{V(\omega_i)} \leq C_\omega \|v\|^2_V, \quad \forall v \in V. \tag{6}$$

In fact this constant C_ω is the maximal number of $\omega_i^{'}$s to which any point $x \in \Omega$ can belong.

For each subspace V_i, $0 \leq i \leq N$, we define a projection $P_i^{\mathcal{B}} \;=\; P_{V_i}^{\mathcal{B}} : \;\; V \;\longrightarrow V_i$, with respect to the bilinear form $\mathcal{B}(\cdot, \cdot)$, as the solution of

$$\mathcal{B}(P_i^{\mathcal{B}} u, v) \;=\; \mathcal{B}(u, v), \;\; \forall v \in V_i.$$

Let us denote $P^{\mathcal{B}} = P_0^{\mathcal{B}} + \cdots + P_N^{\mathcal{B}} : V \longrightarrow V$ and $\mathbf{b} \;=\; P_0^{\mathcal{B}}\mathbf{u} + \cdots + P_N^{\mathcal{B}}\mathbf{u}$, where each component $P_i^{\mathcal{B}}\mathbf{u}$ can be obtained, without *a priori* knowledge the solution $\mathbf{u}$, by solving

$$\mathcal{B}(P_i^{\mathcal{B}}\mathbf{u}, v) \;=\; \mathcal{F}(v), \quad \forall v \in V_i.$$

The additive Schwarz algorithm can be stated as: Find the solution $\mathbf{u}$ of equation (4) by solving the following equation.

$$P^{\mathcal{B}}\mathbf{u} \;=\; \mathbf{b}. \tag{7}$$

We often refer to this equation as the derived equation with respect to the bilinear form $\mathcal{B}(\cdot, \cdot)$ and the decomposition $\{V_i\}$, see [8, 13, 14]. The following theorem is easy to establish.

Theorem 1 *If the operator $P^{\mathcal{B}}$ is invertible, then the equations (4) and (7) have the same solution.*

In practice, the operator $P^{\mathcal{B}}$ can be constructed in terms of the original stiffness matrix and the inverse of some small matrices. The inverse of the small matrices are imbedded, by zero, in a larger matrix. The explicit matrix for $P^{\mathcal{B}}$ is normally not known. However, the matrix-vector-product $P^{\mathcal{B}}u$, which is all that is needed, can be computed by solving a linear system of equations for each subregion. In the following theorem, we summerize the bounds of the operator $P^{\mathcal{B}}$, which determine the convergence rate of the iterative method used to solve (7). The result was established in [8].

Theorem 2 *(1) There exists a constant C, such that*

$$\|P^{\mathcal{B}}u\|_{\mathcal{A}} \leq CC_{\omega}\|u\|_{\mathcal{A}}, \quad \forall u \in V.$$

(2) There exists a constant $c > 0$, independent of C_0, such that

$$\|P^{\mathcal{B}}u\|_{\mathcal{A}} \geq cC_0^{-2}\|u\|_{\mathcal{A}}, \quad \forall u \in V.$$

(3) If there exists $0 < \delta < 1$, such that $\mid \mathcal{N}(u, P^{\mathcal{B}}u) \mid \leq \delta\mathcal{B}(u, P^{\mathcal{B}}u)$, $\forall u \in V$, then

$$(u, P^{\mathcal{B}}u)_{\mathcal{A}} \geq c(1-\delta)C_0^{-2}(u,u)_{\mathcal{A}}, \ \forall u \in V,$$

where $c > 0$ is independent of C_0 and δ.

It is interesting to note that under similar assumptions for the bilinear forms and the skewsymmetric part, an abstract theory for multiplicative Schwarz methods can also be established, see [10].

3 Methods for parabolic convection-diffusion problems

In this section, we apply the abstract additive Schwarz method to parabolic convection-diffusion problems. Three algorithms of this type are discussed, namely, an additive Schwarz method (ASM) for problems in $\mathbb{R}^2$ and $\mathbb{R}^3$, a modified ASM for problems in $\mathbb{R}^2$ and an iterative substructuring method (ISM) for problems in $\mathbb{R}^2$. The convergence rates are given in Theorem 3 - 5.

3.1 An additive Schwarz method for problems in $\mathbb{R}^2$ and $\mathbb{R}^3$

The additive Schwarz method was introduced by Dryja and Widlund [13] for elliptic problems. In this section, we generalize the method to parabolic convection-diffusion problems.

We first introduce our basic decomposition of Ω and the corresponding projections. We first extend each subregion Ω_i, introduced in section 1.2, to obtain $\Omega_i^{'}$, such that $\Omega_i \subset \Omega_i^{'}$ and such that there exists a constant $\alpha > 0$ such that

$$\text{distance}(\partial\Omega_i^{'}, \partial\Omega_i \cap \Omega) \geq \ \alpha H, \ \ \forall i.$$

We suppose that $\partial\Omega_i^{'}$ does not cut through any h-level elements. We make the same constructions for the subregions that meet the boundary except that we cut off the parts that are outside Ω. To simplify the notations, we also denote $\Omega_0^{'} = \Omega$.

It is easy to see that the finite element space V^h can be decomposed into the sum of the coarse mesh function space V^H and a number of spaces which are supported only in subregions Ω_i', i.e. $V^h = V_0^h + V_1^h + \cdots + V_N^h$, where $V_0^h = V^H$ and $V_i^h = V^h \cap H_0^1(\Omega_i')$.

Let $P_i^{D_\tau}$ be the projection from V^h to V_i^h with respect to the bilinear form $D_\tau(\cdot,\cdot)$, and $P^{D_\tau} = P_0^{D_\tau} + \cdots + P_N^{D_\tau}$. We obtain the derived equation with respect to the bilinear form $D_\tau(\cdot,\cdot)$ and the decomposition $\{V_i^h\}$,

$$P^{D_\tau} u^h = g_{n-1}'. \tag{8}$$

Here g_{n-1}' can be computed without *a priori* knowledge of the solution u^h as described in the abstract theory. We will prove that this derived system is uniformly well conditioned under certain conditions, i.e., that the condition number of P^{D_τ} does not change if (1) we refine the fine mesh size h to increase the accuracy; (2) we refine the coarse mesh size H, increasing the number of subproblems for parallel computing purpose; (3) we increase the time step τ.

Theorem 3 *(1) There exists a constant C_p, such that*

$$\|P^{D_\tau} u^h\|_{A_\tau} \le C_p \|u^h\|_{A_\tau}, \quad \forall u^h \in V^h.$$

(2) There exists a constant $c > 0$, such that

$$\|P^{D_\tau} u^h\|_{A_\tau} \ge c \|u^h\|_{A_\tau}, \quad \forall u^h \in V^h.$$

(3) If $c_{H,\tau} = \max\{H, H^\gamma \sqrt{H^2/\tau + 1}\}$ is small enough, i.e. $0 < c_{H,\tau} \le \tilde{c}_0$, then there exists a constant $c_p(\tilde{c}_0) > 0$, such that

$$(u^h, P^{D_\tau} u^h)_{A_\tau} \ge c_p(\tilde{c}_0)(u^h, u^h)_{A\tau}, \ \ \forall u^h \in V^h.$$

Remarks:

(a) The problem is selfadjoint if the first order terms of L vanish and then it can be seen easily that the operator P^{D_τ} is symmetric with respect to A_τ-norm. The standard conjugate gradient method in A_τ-norm can therefore be used. Theorem 3 shows that the derived linear system is optimal for the conjugate gradient method in the sense that the rate of convergence does not depend on the mesh parameters H and h, nor on the time step size τ.

(b) Since, in general, (8) is a nonsymmetric but positive definite system, we can use the GMRES [16, 23] method to solve it. We use the $(\cdot,\cdot)_{A_\tau}$ inner product.

(c) In general, $\tilde{c}_0$ depends on the coefficients of the first order terms in L, the ellipticity constant of D_τ, the bounds on D_τ and also on the geometry of the domain Ω, which is reflected in γ. We do not have an explicit relation between $\tilde{c}_0$ and the skewsymmetric part of L. From the proof of the lower bound, we know that as the skewsymmetric coefficients in L increase, $\tilde{c}_0$ decreases.

In order to prove the theorem, we need only to show that all the assumptions of Theorem 2 hold and, in addition, that all the constants that appear in that abstract result are independent of the mesh parameters H, h and τ. We begin by establishing some lemmas, which contain most of the basic results.

The following lemma is well known.

Lemma 3 *There exist two constants $c > 0$ and C, which depend only on the shape regularity of the finite element subdivision of Ω, such that*

$$ch^d \sum_{x_i \in \Lambda^h} (u^h(x_i))^2 \le \|u^h\|^2_{L^2(\Omega)} \le Ch^d \sum_{x_i \in \Lambda^h} (u^h(x_i))^2, \ \forall u^h \in V^h.$$

The statement is also true if we replace V^h by V^H and h by H.

Lemma 4 *For all $u^h \in V^h$, there exist $u_i^h \in V_i^h$, $i = 0, 1, \cdots, N$, such that*

$$u^h \ = \ \sum_{i=0}^{N} u_i^h$$

and there exists a constant C, which is independent of h and H, such that

$$\sum_{i=0}^{N} \|u_i^h\|^2_{A_\tau(\Omega_i')} \le C\|u^h\|^2_{A_\tau}.$$

Proof: A proof can be found in [14] for the H_0^1 norm. Thus, our only task is to show that the same estimate holds for the L^2 norm. The construction of u_i^h, $i = 0, 1, \cdots, N$ is the same as that of [14]. Let Q_H be the L^2 projection into the coarse mesh space defined as

$$(Q_H u^h, v^H) \ = \ (u^h, v^H), \ \ \forall v^H \in V^H.$$

We take $u_0^h = Q_H u^h$, $w^h = u^h - Q_H u^h$, and then set $u_i^h = I_h(\theta_i(w^h))$, where $\{\theta_i\}$ is a partition of unity and θ_i belongs to $C_0^\infty(\Omega_i')$. It can be arranged so that $\nabla\theta_i$ is bounded by $const/H$. I_h is a interpolation operator which uses the function values at the h-level nodes only. Because $Q_H u^h$ is the L^2 projection, we have

$$\|Q_H u^h\|^2_{L^2(\Omega)} \le \|u^h\|^2_{L^2(\Omega)}. \tag{9}$$

Next, we prove the $L^2(\Omega)$ boundness for the other u_i^h. By using Lemma 3 for each Ω_i',

$$\sum_{i=1}^{N} \|I^h(\theta_i w^h)\|^2_{L^2(\Omega_i')} \le Ch^2 \sum_{i=1}^{N} (\sum_{x_j \in \Lambda^h \cap \Omega_i'} ((\theta_i w^h)(x_j))^2).$$

Since $| \theta_i | \le 1$, $i = 1, \cdots, N$, and $\{\Omega_i'\}$ is a finite covering of Ω, the right hand side can be bounded by $h^2 \sum_{x_i \in \Lambda^h} (w^h(x_i))^2$. Using Lemma 3 again, we obtain

$$\sum_{i=1}^{N} \|u_i^h\|^2_{L^2(\Omega_i')} \le C\|w^h\|^2_{L^2(\Omega)} \le C\|u^h\|^2_{L^2(\Omega)}. \tag{10}$$

We conclude the proof by combining the estimates (9) and (10) and the estimate in the H_0^1 norm.$\square$

Lemma 5 *For any $w \in H_0^1(\Omega) \cap H^{1+\gamma}(\Omega)$, there exists a $w^H \in V^H$, such that*

$$\|w - w^H\|_{\tau(\Omega)} \leq CH^\gamma \sqrt{H^2 + \tau} \mid w \mid_{H^{1+\gamma}(\Omega)} .$$

Proof: For any given w, let w^H be the solution of

$$a(w^H, v) \; = \; a(w, v), \quad \forall v \in V^H,$$

where $a(u,v) = \int_\Omega \nabla u \nabla v d\Omega$. By classical finite element theory, we know that

$$\|w - w^H\|_{H_0^1(\Omega)} \leq CH^\gamma \mid w \mid_{H^{1+\gamma}(\Omega)},$$

$$\|w - w^H\|_{L^2(\Omega)} \leq CH^{1+\gamma} \mid w \mid_{H^{1+\gamma}(\Omega)} .$$

Hence, we have $\|w - w^H\|_{\tau(\Omega)} \leq CH^\gamma \sqrt{H^2 + \tau} \mid w \mid_{H^{1+\gamma}(\Omega)}$. $\Box$

Lemma 6 *There exists a constant C, which is independent of H and τ, such that*

$$\|u^h - P_0^{D_\tau} u^h\|_{L^2(\Omega)} \leq CH^\gamma (\sqrt{H^2 + \tau}/\tau) \|u^h\|_{\tau(\Omega)}, \;\; \forall u^h \in V^h.$$

Proof: We first establish a bound for the τ-norm. It is easy to verify that

$$D_\tau(u^h - P_0^{D_\tau} u^h, u^h - P_0^{D_\tau} u^h) = D_\tau(u^h - P_0^{D_\tau} u^h, u^h).$$

Using Lemma 1, we get the estimate

$$\|u^h - P_0^{D_\tau} u^h\|_{\tau(\Omega)} \leq C \|u^h\|_{\tau(\Omega)}.$$

The $H_0^1(\Omega)$ (or τ-norm) estimate and Nitsche's trick give the $L^2(\Omega)$ estimate. We have

$$\|u^h - P_0^{D_\tau} u^h\|_{L^2(\Omega)} = \sup_{\|v\|_{L^2(\Omega)} \neq 0} \frac{(u^h - P_0^{D_\tau} u^h, v)}{\|v\|_{L^2(\Omega)}}.$$

For any fixed $v \in L^2(\Omega)$, we form an auxiliary problem: Find $w \in H^{1+\gamma}(\Omega) \cap H_0^1(\Omega)$, such that

$$D_\tau(\phi, w) \; = \; (\phi, v), \;\; \forall \phi \in H_0^1(\Omega).$$

By Lemma 2, we have $\mid w \mid_{H^{1+\gamma}(\Omega)} \leq C/\tau \|v\|_{L^2(\Omega)}$. Let $\phi = u^h - P_0^{D_\tau} u^h \;\; \in H_0^1(\Omega)$. Then

$$(u^h - P_0^{D_\tau} u^h, v) = D_\tau(u^h - P_0^{D_\tau} u^h, w).$$

Let w^H be the V^H approximation of w obtained in Lemma 5. Since $w^H \in V^H$, we have

$$\begin{aligned} \mid D_\tau(u^h - P_0^{D_\tau} u^h, w) \mid \;\; &= \;\; \mid D_\tau(u^h - P_0^{D_\tau} u^h, w - w^H) \mid \\ &\leq \;\; C \|u^h - P_0^{D_\tau} u^h\|_{\tau(\Omega)} \|w - w^H\|_{\tau(\Omega)} \\ &\leq \;\; CH^\gamma \sqrt{H^2 + \tau} \|u^h\|_{\tau(\Omega)} \mid w \mid_{H^{1+\gamma}(\Omega)} . \end{aligned}$$

Combining the above results, we obtain

$$\|u^h - P_0^{D_\tau} u^h\|_{L^2(\Omega)} \leq CH^\gamma (\sqrt{H^2 + \tau}/\tau) \|u^h\|_{\tau(\Omega)}. \Box$$

In the next lemma, we estimate the contribution from the skewsymmetric part $N_\tau(\cdot, \cdot)$. We show that it is a lower order term compared with the symmetric part, and that it can be controlled if the coarse mesh size is fine enough.

Lemma 7 *If* $\max\{H, H^{\gamma}\sqrt{H^2/\tau+1}\}$ *is small enough, there exists a constant* $0<\delta<1$*, such that*

$$| N_\tau(u^h, P^{D_\tau}u^h) | \le \delta D_\tau(u^h, P^{D_\tau}u^h), \quad \forall u^h \in V^h.$$

Proof: Since $N_\tau(P_i^{D_\tau}u^h, P_i^{D_\tau}u^h) = 0$, it is easy to verify that

$$| N_\tau(u^h, P^{D_\tau}u^h) | \le \sum_{i=0}^{N} | N_\tau(P_i^{D_\tau}u^h, u^h - P_i^{D_\tau}u^h) | .$$

Therefore, we need only to estimate the right hand side of the above inequality. Since the coarse mesh projection $P_0^{D_\tau}$ is special, we consider it separately using Lemma 6.

(1) $i = 0$; By Lemma 6

$$\begin{aligned} | N_\tau(P_0^{D_\tau}u^h, u^h - P_0^{D_\tau}u^h) | &\le C\tau \|P_0^{D_\tau}u^h\|_{H_0^1(\Omega)} \|u^h - P_0^{D_\tau}u^h\|_{L^2(\Omega)} \\ &\le CH^{\gamma}\sqrt{H^2+\tau}\|P_0^{D_\tau}u^h\|_{H_0^1(\Omega)} \cdot \|u^h\|_{\tau(\Omega)}. \end{aligned}$$

It is easy to see that $\sqrt{\tau}\|v\|_{H_0^1(\Omega)} \le \|v\|_{\tau(\Omega)}$. Hence,

$$| N_\tau(P_0^{D_\tau}u^h, u^h - P_0^{D_\tau}u^h) | \le CH^{\gamma}\sqrt{H^2/\tau+1}\|P_0^{D_\tau}u^h\|_{\tau(\Omega)}\|u^h\|_{\tau(\Omega)}.$$

From the definition of $P_0^{D_\tau}$, we have

$$D_\tau(P_0^{D_\tau}u^h, P_0^{D_\tau}u^h) = D_\tau(u^h, P_0^{D_\tau}u^h).$$

We thus have, by Lemma 1, $\|P_0^{D_\tau}u^h\|_{\tau(\Omega)} \le C\|u^h\|_{\tau(\Omega)}$.

Therefore, the first term can be bounded as follows

$$| N_\tau(P_0^{D_\tau}u^h, u^h - P_0^{D_\tau}u^h) | \le CH^{\gamma}\sqrt{H^2/\tau+1}\|u^h\|^2_{\tau(\Omega)}.$$

(2) $i \ne 0$:

$$\begin{aligned} | N_\tau(P_i^{D_\tau}u^h, u^h - P_i^{D_\tau}u^h) | &\le C\tau\|P_i^{D_\tau}u^h\|_{L^2(\Omega_i')}(\|u^h\|_{H_0^1(\Omega_i')} + \|P_i^{D_\tau}u^h\|_{H_0^1(\Omega_i')}) \\ &\le C\sqrt{\tau}\|P_i^{D_\tau}u^h\|_{L^2(\Omega_i')}(\|u^h\|_{\tau(\Omega_i')} + \|P_i^{D_\tau}u^h\|_{\tau(\Omega_i')}). \end{aligned}$$

The first factor can be estimated by using Friedrich's inequality. Since $P_i^{D_\tau}u^h \in H_0^1(\Omega_i')$ and the diameter of Ω_i' is of order H, we have

$$\|P_i^{D_\tau}u^h\|_{L^2(\Omega_i')} \le CH\|P_i^{D_\tau}u^h\|_{H_0^1(\Omega_i')}.$$

By the definition of $P_i^{D_\tau}$,

$$D_\tau(P_i^{D_\tau}u^h, P_i^{D_\tau}u^h) \;=\; D_\tau(u^h, P_i^{D_\tau}u^h).$$

We obtain, by Lemma 1, $\|P_i^{D_\tau}u^h\|_{\tau(\Omega_i')} \le C\|u^h\|_{\tau(\Omega_i')}$.

Combining these inequalities, we obtain

$$| N_\tau(P_i^{D_\tau} u^h, u^h - P_i^{D_\tau} u^h) | \leq CH \|u^h\|^2_{\tau(\Omega_i')}.$$

Putting the results in (1) and (2) together, we have

$$| N_\tau(P^{D_\tau} u^h, u^h) | \leq C \max\{H, H^\gamma \sqrt{H^2/\tau + 1}\} \sum_{i=0}^{N} \|u^h\|^2_{\tau(\Omega_i')}, \tag{11}$$

which also holds in the A_τ-norm. It is easy to see that

$$\sum_{i=0}^{N} \|P_i^{D_\tau} u^h\|^2_{A_\tau(\Omega_i')} \; = \; D_\tau(u^h, P^{D_\tau} u^h).$$

By using the decomposition lemma, we have

$$\begin{aligned} \|u^h\|^2_{A_\tau} &= \textstyle\sum_{i=0}^N D_\tau(P_i^{D_\tau} u^h, u_i^h) \\ &\leq C \textstyle\sum_{i=0}^N \|P_i^{D_\tau} u^h\|_{A_\tau(\Omega_i')} \|u_i^h\|_{A_\tau(\Omega_i')} \\ &\leq C \sqrt{\textstyle\sum_{i=0}^N \|u_i^h\|^2_{A_\tau(\Omega_i')}} \sqrt{\textstyle\sum_{i=0}^N \|P_i^{D_\tau} u^h\|^2_{A_\tau(\Omega_i')}}. \end{aligned}$$

Therefore, we obtain

$$\|u^h\|^2_{A_\tau} \leq C \sum_{i=0}^{N} \|P_i^{D_\tau} u^h\|^2_{A_\tau(\Omega_i')}.$$

Hence

$$\|u^h\|^2_{A_\tau} \leq C D_\tau(u^h, P^{D_\tau} u^h). \tag{12}$$

We complete the proof by combining estimates (11) and (12). □

The proof of Theorem 3 follows from Theorem 2 and the lemmas in this section.

3.2 The modified additive Schwarz method for problems in $\mathbb{R}^2$

In this subsection, we propose a modified version of ASM obtained by dropping the coarse mesh space V_0^h, which provided the global transportation of information. We show that in some situations, with precise conditions given in the following theorem, the global space is not necessary. This is in contrast to the elliptic case, for fast convergence, this alternative algorithm is more suitable for parallel computers.

Let us define $\tilde{P}^{D_\tau} = P_1^{D_\tau} + \cdots + P_N^{D_\tau}$, where the $P_i^{D_\tau}$ are the same as in the previous subsection. We have the following theorem.

Theorem 4 *(1) There exists a constant $C_{\tilde{p}}$, such that*

$$\|\tilde{P}^{D_\tau} u^h\|_{A_\tau} \leq C_{\tilde{p}} \|u^h\|_{A_\tau}, \quad \forall u^h \in V^h.$$

(2) There exists a constant $c > 0$, such that

$$\|\tilde{P}^{D_\tau} u^h\|_{A_\tau} \geq c(1 + \tau/H^2)^{-1} \|u^h\|_{A_\tau}, \quad \forall u^h \in V^h.$$

(3) If $c_{H,\tau} = H(1 + \tau/H^2)$ is small enough, i.e. $0 < c_{H,\tau} \le \tilde{c}_0$, then there exists a constant $c_{\tilde{p}}(\tilde{c}_0) > 0$, such that

$$(u^h, \tilde{P}^{D_\tau} u^h)_{A_\tau} \ge c_{\tilde{p}}(\tilde{c}_0)(u^h, u^h)_{A\tau}, \ \ \forall u^h \in V^h.$$

Remarks:

(a) For symmetric problems, parts (1) and (2) of this theorem show that if the factor τ/H^2 is small, the elimination of the coarse mesh space does not lead to slow convergence. This suggests that we can use the modified ASM when we have a relatively small time step or large substructures.

For nonsymmetric problems, we need that both H and τ/H^2 to be small in order to obtain the fast convergence. We cannot choose τ and H independently.

(b) This theorem has been established only in $\mathbb{R}^2$. In higher dimension, the fine mesh size h enters our bounds.

Theorem 4 can be proved by using the results of Theorem 2, a new decomposition lemma given below and a new estimate of the skewsymmetric part. We begin with the decomposition lemma.

Lemma 8 *For all $u^h \in V^h$, there exist $u_i^h \in V_i^h$, $i = 1, \cdots, N$, such that*

$$u^h \ = \ \sum_{i=1}^{N} u_i^h$$

and, there exists a constant C, which is independent of h, H and τ, such that

$$\sum_{i=1}^{N} \|u_i^h\|^2_{A_\tau(\Omega_i')} \le C(1 + \tau/H^2)\|u^h\|^2_{A_\tau}.$$

Proof: We first construct the decomposition and then we do the estimates in both the $H_0^1(\Omega)$ and the $L^2(\Omega)$ norm. Let $\{\theta_i, i = 1, \cdots, N\}$ be the partition of unity, defined in the proof of Lemma 4, of Ω and denote $u_i^h \ = \ I_h(\theta_i u^h)$. For each subregion Ω_i', we have, in the sense of equivalent norms, that

$$| \ u_i^h \ |^2_{H_0^1(\Omega_i')} = \sum ((\theta_i u^h)(x_l) - (\theta_i u^h)(x_m))^2,$$

where the sum is taken over all adjacent pairs of nodal points x_l and x_m in Ω_i'. Let $K \subset \Omega_i'$ be a single element and x_l, $x_m \in \bar{K}$. Let

$$\bar{\theta}_{ilm} = 1/2(\theta_i(x_l) + \theta_i(x_m)).$$

We then have

$$(\theta_i u^h)(x_l) - (\theta_i u^h)(x_m) \ =$$
$$(\theta_i(x_l) - \bar{\theta}_{ilm})u^h(x_l) - (\theta_i(x_m) - \bar{\theta}_{ilm})u^h(x_m) + \bar{\theta}_{ilm}(u^h(x_l) - u^h(x_m)),$$

which can be bounded from above by

$$C(h/H \max_{\bar{K}}\{| \ u^h(x) \ |\} + | \ u^h(x_l) - u^h(x_m) \ |).$$

By squaring this estimate, using the triangle inequality and summing over all $K \subset \Omega_i^{'}$, we obtain

$$
\begin{aligned}
\mid u_i^h \mid^2_{H_0^1(\Omega_i')} &\leq C(H^{-2} \textstyle\sum_{K \subset \Omega_i'} \max_{\bar{K}} \{\mid u^h(x) \mid\}^2 \cdot h^2 + \mid u^h \mid^2_{H_0^1(\Omega_i')}) \\
&\leq C(H^{-2} \|u^h\|^2_{L^2(\Omega_i')} + \mid u^h \mid^2_{H_0^1(\Omega_i')}).
\end{aligned}
$$

Using Friedrich's inequality, it is easy to establish that

$$
\sum_{i=1}^N \|u_i^h\|^2_{H_0^1(\Omega_i')} \leq (1 + CH^2) \sum_{i=1}^N \mid u_i^h \mid^2_{H_0^1(\Omega_i')} .
$$

Therefore,

$$
\sum_{i=1}^N \|u_i^h\|^2_{H_0^1(\Omega_i')} \leq C(H^{-2} \|u^h\|^2_{L^2(\Omega)} + \mid u^h \mid^2_{H_0^1(\Omega)}) \leq C H^{-2} \|u^h\|^2_{H_0^1(\Omega)}.
$$

By combining the estimate given for the L^2 norm in the proof of Lemma 4 and the definition of the τ-norm, we complete the proof. □

In the next lemma, we estimate the contribution from the skewsymmetric part $N_\tau(\cdot, \cdot)$.

Lemma 9 *There exists a constant C, which is independent of h, H and τ, such that*

$$
\mid N_\tau(u^h, \tilde{P}^{D_\tau} u^h) \mid \leq CH(1 + \tau/H^2) D_\tau(u^h, \tilde{P}^{D_\tau} u^h), \ \ \forall u^h \in V^h.
$$

The proof is very similar, replacing the decomposition Lemma 4 by Lemma 8, to that of Lemma 7 part (2).

3.3 An iterative substructuring method for problems in $\mathbb{R}^2$

We present a domain decomposition method which appears to be of non-overlapping type, but nevertheless fits into the additive Schwarz framework. In the first two algorithms, we extended each H-level substructure Ω_i to a larger region $\Omega_i^{'}$, which defined our subspace and the corresponding projection. $\{\Omega_i^{'}\}$ is thus the basic decomposition of the domain. Now, instead of extending each region, we combine each pair of adjacent substructures. Denote by Γ_{ij} the common edge of two adjacent substructures Ω_i and Ω_j, and let $\Omega_{ij} = \Omega_i \cup \Gamma_{ij} \cup \Omega_j$. We are now going to use $\{\Omega_{ij}\}$ as the basic decomposition of Ω, on which we will define our projections. This will eventually lead to an iterative substructuring algorithm. Let $\Omega_{00} = \Omega$ and denote

$$
\Lambda^E = \{(i,j) \mid x_i, x_j \in \Lambda^H, x_i, x_j \text{ adjacent, or } i = j = 0\}.
$$

For each $(0,0) \neq (i,j) \in \Lambda^E$, we define $V_{ij}^h \equiv H_0^1(\Omega_{ij}) \cap V^h$ and also $V_{00}^h = V^H$. Denote by $P_{ij}^{D_\tau}$ the projection associated with V_{ij}^h using the bilinear form $D_\tau(\cdot, \cdot)$ It is easy to verify that

$$
V^h = \sum_{(i,j) \in \Lambda^E} V_{ij}^h.
$$

Let us also denote

$$P^{D_\tau} \;=\; \sum_{(i,j)\in\Lambda^E} P_{ij}^{D_\tau} : \, V^h \longrightarrow V^h$$

and the derived equation can thus be defined as

$$P^{D_\tau}\mathbf{u} \;=\; \mathbf{b}^{D_\tau}.$$

In the next theorem, we present the bounds for the operator P^{D_τ}. We use the notation

$$\nu(h,H,\tau) \;=\; (1 + H^2/\tau + \log H/h)(1 + \log H/h).$$

Theorem 5 *(1) There exists a constant C_p, such that*

$$\|P^{D_\tau}u^h\|_{A_\tau} \le C_p\|u^h\|_{A_\tau}, \quad \forall u^h \in V^h.$$

(2) There exists a constant $c > 0$, such that

$$\|P^{D_\tau}u^h\|_{A_\tau} \ge c/\nu(h,H,\tau)\|u^h\|_{A_\tau}, \quad \forall u^h \in V^h.$$

(3) There exists a constant $c_p > 0$, such that for $H\sqrt{\nu}$ small enough,

$$(u^h, P^{D_\tau}u^h)_{A_\tau} \ge c_p/\nu(h,H,\tau)(u^h,u^h)_{A\tau}, \;\; \forall u^h \in V^h.$$

As before, we need only to prove that the assumptions of Theorem 2 hold and study how the constants depend on the mesh parameters h, H and τ. The proofs look more complicated because of the lack of overlap between the substructures, but the idea behind the proof is similar. We need to establish a bounded decomposition lemma for V^h functions. The bounds depend mildly on the parameters h and H.

We first give a lemma, which plays an important role in the traditional theory for iterative substructuring algorithms. Variations of this result, which dates back at least to 1966, are given in Bramble [3], Bramble, Pasciak and Schatz [4] and Yserentant [25].

Lemma 10 *Let Ω_i be a substructure, $u^h \in V_i^h$ and let $\overline{u_i^h} = \frac{1}{area(\Omega_i)}\int_{\Omega_i} u^h d\Omega$. Then*

$$\|u^h - \overline{u_i^h}\|^2_{L^\infty(\Omega_i)} \le C(1 + \log H/h) \mid u^h \mid^2_{H^1(\Omega_i)} .$$

Lemma 11 *For any $u^h \in V^h$, there exist $u_{ij}^h \in V_{ij}^h$, $(i,j) \in \Lambda^E$, such that $u^h \;=\; \sum_{(i,j)\in\Lambda^E} u_{ij}^h$ and the decomposition is bounded as follows*

(a) $\sum_{(i,j)\in\Lambda^E} \|u_{ij}^h\|^2_{H_0^1(\Omega_{ij})} \le C(1 + \log H/h)^2\|u^h\|^2_{H_0^1(\Omega)}$

(b) $\sum_{(i,j)\in\Lambda^E} \|u_{ij}^h\|^2_{L^2(\Omega_{ij})} \le C(\|u^h\|^2_{L^2(\Omega)} + H^2(1 + \log H/h)\|u^h\|^2_{H_0^1(\Omega)})$.

Proof: The construction of the decomposition and the proof of (a) were given by Dryja and Widlund in [14]. We shall focus on (b) only. Let I_H be the interpolation operator using the function values at the H-level nodes only. For a given $u^h \in V^h$, we estimate the $L^2(\Omega)$ norm of $u_{00}^h \;=\; I_H u^h$. Let us consider one substructure at a time.

Assume that Ω_i has vertices T_1, T_2, T_3 and denote $\alpha_i = \frac{1}{\text{area}(\Omega_i)} \int_{\Omega_i} u^h d\Omega$. Then, we have

$$\|I_H u^h\|^2_{L^2(\Omega_i)} \leq C(\|I_H u^h - \alpha_i\|^2_{L^2(\Omega_i)} + \|\alpha_i\|^2_{L^2(\Omega_i)}).$$

Since the function $I_H u^h - \alpha_i$ is linear in the region Ω_i, a straightforward calculation shows that the L^2 norm in Ω_i can be bounded by

$$CH^2((I_H u^h - \alpha_i)^2 \mid_{T_1} + (I_H u^h - \alpha_i)^2 \mid_{T_2} + (I_H u^h - \alpha_i)^2 \mid_{T_3}).$$

This expression is, using Lemma 10, bounded by $CH^2(1 + \log H/h) \mid u^h \mid^2_{H^1(\Omega_i)}$. We bound the $\|\alpha_i\|_{L^2(\Omega_i)}$ term, by using Schwarz's inequality

$$\|\alpha_i\|^2_{L^2(\Omega_i)} \ = \ \int_{\Omega_i} (\frac{1}{\text{area}(\Omega_i)} \int_{\Omega_i} u^h d\Omega)^2 d\Omega \ \leq \ \|u^h\|^2_{L^2(\Omega_i)}.$$

Therefore, we obtain

$$\|I_H u^h\|^2_{L^2(\Omega_i)} \leq C(\|u^h\|^2_{L^2(\Omega_i)} + H^2(1 + \log H/h) \mid u^h \mid^2_{H^1(\Omega_i)}).$$

By summing these inequalities, using Friedrich's inequality and replacing the H^1 seminorm by the H^1 norm, we obtain

$$\|I_H u^h\|^2_{L^2(\Omega)} \leq C(\|u^h\|^2_{L^2(\Omega)} + H^2(1 + \log H/h)\|u^h\|^2_{H^1_0(\Omega)}).$$

Now consider the case $(i, j) \neq (0, 0)$. Recalling the construction of [14], $u^h_{ij} = I_h(\theta_{ij}(u^h - I_H u^h))$, where $\{\theta_{ij}\}$ is a partition of unity of $\Omega \setminus \{H\text{-level nodes}\}$. More precise, as described in [14], θ_{ij} must equal to zero at all $H-$level nodes and equal to one at all nodal points in the interior of Γ_{ij}. Let us denote $\Lambda^h_{ij} = \Omega_{ij} \cap \Lambda^h$. By Lemma 3, we have

$$\|u^h_{ij}\|^2_{L^2(\Omega_{ij})} \leq Ch^2 \sum_{x_k \in \Lambda^h_{ij}} ((I_h(\theta_{ij}(u^h - I_H u^h)))(x_k))^2.$$

Since $\mid \theta_{ij} \mid \leq 1$, and x_k is a nodal point, the interpolation operator I_h can be removed. Therefore the right hand side can be bounded by

$$Ch^2 \sum_{x_k \in \Lambda^h_{ij}} ((u^h - I_H u^h)(x_k))^2.$$

Taking into account that the value at a nodal point contributes at most three times, we obtain

$$\sum_{(i,j) \in \Lambda^E \setminus (0,0)} \|u^h_{ij}\|^2_{L^2(\Omega_{ij})} \leq Ch^2 \sum_{x_k \in \Lambda^h} ((u^h - I_H u^h)(x_k))^2.$$

We can bound the right hand side, using Lemma 3, by $C\|u^h - I_H u^h\|^2_{L^2(\Omega)}$, which can then be bounded by $C(\|u^h\|^2_{L^2(\Omega)} + H^2(1 + \log H/h)\|u^h\|^2_{H^1_0(\Omega)})$. $\square$

An immediate consequence of this lemma is the following:

Lemma 12 *The decomposition in Lemma 11 is bounded in the A_τ-norm, i.e. there exists a constant C, independent of H, h and τ, such that*

$$\sum_{(i,j)\in\Lambda^E} \|u_{ij}^h\|^2_{\tau(\Omega_{ij})} \le C\nu(h,H,\tau)\|u^h\|^2_{\tau(\Omega)}.$$

Lemma 13 *There exists a constant C, independent of h, H and τ, such that*

$$\mid N_\tau(u^h, P^{D_\tau}u^h) \mid \le C\max\{H, H^\gamma\sqrt{H^2/\tau+1}\}\nu(h,H,\tau)D_\tau(u^h, P^{D_\tau}u^h), \quad \forall u^h \in V^h.$$

Proof: It is easy to see that the coarse mesh projection $P_{00}^{D_\tau}$ is identical to the coarse mesh projection $P_0^{D_\tau}$ defined previously. It follows from the proof of Lemma 7 part (1) and (2), that we have

$$\mid N_\tau(P^{D_\tau}u^h, u^h) \mid \le C\max\{H, H^\gamma\sqrt{H^2/\tau+1}\} \sum_{(i,j)\in\Lambda^E} \|u^h\|^2_{\tau(\Omega_{ij})}. \tag{13}$$

By using the same argument as when equation (12) was proved and the decomposition lemma 12, we have

$$\|u^h\|^2_{A_\tau} \le C\nu(h,H,\tau)D_\tau(u^h, P^{D_\tau}u^h).\square$$

The proof of Theorem 5 is completed by using Theorem 2, the lemmas in this section and the assumption that $\max\{H, H^\gamma\sqrt{H^2/\tau+1}\}\nu(h,H,\tau)$ is small enough.

4 Numerical results

In this section, we present some numerical results with model problems. To specify our model problems, we need only to give the elliptic parts of the parabolic operators. We consider the following linear second order elliptic operator defined on $\Omega = [0,1]\times[0,1] \subset \mathbb{R}^2$,

$$Lu = -\frac{\partial}{\partial x}(\xi\frac{\partial u}{\partial x}) - \frac{\partial}{\partial y}(\eta\frac{\partial u}{\partial y}) + \alpha\frac{\partial u}{\partial x} + \beta\frac{\partial u}{\partial y} + \gamma u = f,$$

with the homogenous Dirichlet boundary condition. The coefficients are specified as follows.

Example 0. $\xi = 1$, $\eta = 1$ and $\alpha = \beta = \gamma = 0$. This is a selfadjoint problem, used to test the iterative substructuring algorithm. f is chosen so that the solution is $u = xe^{xy}sin(\pi x)sin(\pi y)$.

Example 1. $\xi = 1 + x^2 + y^2$, $\eta = e^{xy}$, $\alpha = 5(x+y)$, $\beta = 1/(1+x+y)$ and $\gamma = 0$. u is the same as in Example 0.

Example 2. The coefficients are chosen as $\xi = \sigma$, $\eta = \sigma$, $\alpha = 1$, $\beta = 1$ and $\gamma = 1$. σ will be specified later. u is the same as in Example 0.

The H-level subdivision of Ω is as shown in Figure 1. We further divide each subregion in a similar fashion into h-level triangles, which are not shown in the figure.

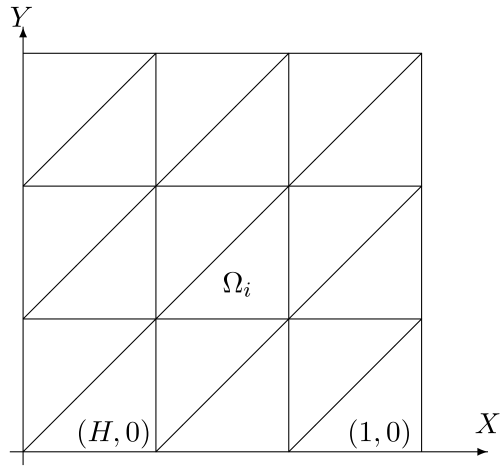

Figure 1. H−level subdivision

We use **ovlp** to denote the size of the overlap, i.e. **ovlp** = distance$(\partial\Omega_i', \partial\Omega_i \cap \Omega)$. In our Fortran program, all the subproblems are solved exactly by the band solver from LINPACK. The stopping criterion for the GMRES method is $\|r_i\|_{A_\tau}/\|r_0\|_{A_\tau} \le 10^{-4}$, where r_i is the residual after the i^{th} step. The programs were run in single precision on a CONVEX C-1 computer at New York University.

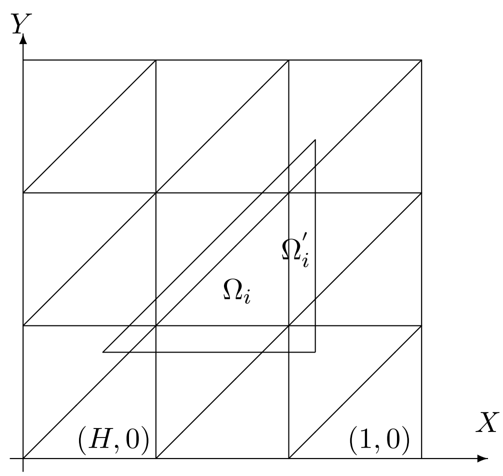

Figure 2. extended subregions

We assume that the time step has the form

$$\tau = h^{\epsilon},$$

where $\epsilon > 0$. The main issue here is to determine how the rate of convergence (number of iterations) depends on ϵ. In the case of $\epsilon = 2.0$, the stiffness matrix is well conditioned with a condition number independent of the mesh parameters h and τ. In general, the smaller ϵ is (the larger the time step is) the more ill-conditioned is the problem. If ϵ is very small, the accuracy of the solution is lost. We include some test results by using unusually large time step τ to show the robustness of the algorithm. In our experiments, we have used $\epsilon \in [0.25, 1.5]$.

4.1 Tests of the additive Schwarz method

	ϵ	1.5	1.0	0.5	0.25
Example 0	h=1/45,H=1/3,**ovlp**=5h	10	12	12	13
	h=1/60,H=1/3,**ovlp**=6h	10	12	13	13
	h=1/60,H=1/10,**ovlp**=2h	10	11	12	12
	h=1/45,H=1/3,**ovlp**=h	15	18	18	19
Example 1	h=1/15,H=1/3,**ovlp**=2h	12	13	14	14
	h=1/30,H=1/3,**ovlp**=4h	12	14	14	14
	h=1/45,H=1/3,**ovlp**=6h	12	14	15	15
	h=1/60,H=1/3,**ovlp**=8h	11	14	15	15
	h=1/15,H=1/5,**ovlp**=1h	11	12	12	12
	h=1/30,H=1/5,**ovlp**=2h	12	14	14	14
	h=1/45,H=1/5,**ovlp**=3h	12	13	14	15
	h=1/60,H=1/5,**ovlp**=4h	11	13	15	15
Example 2 $\sigma = \sqrt{2}/30$	h=1/45,H=1/3,**ovlp**=5h	11	17	28	35
	h=1/45,H=1/9,**ovlp**=2h	9	12	14	16
	h=1/45,H=1/15,**ovlp**=1h	7	8	10	10
	h=1/60,H=1/10,**ovlp**=2h	9	11	14	15
	h=1/60,H=1/15,**ovlp**=1h	7	9	10	11
	h=1/60,H=1/20,**ovlp**=1h	6	8	10	10
Example 2 $\sigma = \sqrt{2}/100$	h=1/60,H=1/15,**ovlp**=1h	8	11	18	20
	h=1/60,H=1/20,**ovlp**=1h	7	9	13	16

This table shows that the number of iterations is almost independent on the mesh parameters as well as the time step size, especially for symmetric problems. For nonsymmetric, see Theorem 3. For nonsymmetric problems, a relatively small coarse mesh size is needed for reducing the total number of iterations, which confirms Theorem 3 part(3). The larger the Reynold's number, which is roughly the ratio of the first order term coefficients and the second order term coefficients, is, the smaller the coarse mesh size is needed. As in the stationary case, using less overlap can increase the total number of iterations. This can be seen from the last row of Example 0.

4.2 Tests of the modified additive Schwarz method

	ϵ	1.5	1.0	0.5	0.25
Example 0	h=1/45,H=1/3,**ovlp**=5h	10	12	14	15
	h=1/60,H=1/3,**ovlp**=6h	9	12	15	16
	h=1/60,H=1/10,**ovlp**=2h	12	24	34	35
	h=1/45,H=1/3,**ovlp**=h	12	21	27	28
Example 1	h=1/15,H=1/3,**ovlp**=2h	13	15	15	15
	h=1/30,H=1/3,**ovlp**=4h	13	14	16	16
	h=1/45,H=1/3,**ovlp**=6h	13	15	16	17
	h=1/60,H=1/3,**ovlp**=8h	12	15	17	17
	h=1/15,H=1/5,**ovlp**=1h	14	18	19	20
	h=1/30,H=1/5,**ovlp**=2h	15	20	23	24
	h=1/45,H=1/5,**ovlp**=3h	13	19	24	25
	h=1/60,H=1/5,**ovlp**=4h	13	18	23	25

	ϵ	1.5	1.0	0.75	0.5
Example 2 $\sigma = \sqrt{2}/30$	h=1/45,H=1/3,**ovlp**=5h	13	22	31	42
	h=1/45,H=1/5,**ovlp**=3h	12	18	23	32
	h=1/45,H=1/9,**ovlp**=2h	10	14	19	30
	h=1/45,H=1/15,**ovlp**=1h	7	12	22	36
	h=1/60,H=1/10,**ovlp**=2h	10	13	20	33
	h=1/60,H=1/15,**ovlp**=1h	8	13	24	41
Example 2 $\sigma = \sqrt{2}/100$	h=1/45,H=1/9,**ovlp**=2h	10	19	31	+50
	h=1/45,H=1/15,**ovlp**=1h	7	13	25	48
	h=1/60,H=1/15,**ovlp**=1h	7	13	25	+50
	h=1/60,H=1/20,**ovlp**=1h	6	12	24	+50

This table shows that if the factor τ/H^2 is small, the results are quite satisfactory compared with the ASM which also uses a coarse mesh. This is true especially for symmetric and mildly nonsymmetric problems; cf. Example 0 and Example 1. However, if τ/H^2 is large, i.e. the time step or the number of substructures large, the number of iterations can be large; cf. the third row of Example 0. For the problems with high Reynold's number, the algorithm becomes very sensitive to the parameter ϵ; cf. the last colum of the second table. See also Theorem 4.

We also note that using less overlap can increases the total number of iterations. This can be seen by comparing the first and last rows in Example 0.

By comparing with the ASM which uses coarse mesh, we see that this simpler algorithm is more sensitive to the mesh parameters and the overlapping factor.

4.3 Tests of the iterative substructuring method

	ϵ	1.5	1.0	0.5	0.25
Example 0	h=1/45,H=1/3	14	14	14	14
	h=1/60,H=1/3	15	15	14	14
	h=1/60,H=1/10	11	12	12	13
Example 1	h=1/45,H=1/5	18	17	17	17
	h=1/45,H=1/9	14	14	14	14
	h=1/60,H=1/5	18	17	17	18
	h=1/60,H=1/10	13	14	14	14
Example 2 $\sigma = \sqrt{2}/30$	h=1/45,H=1/5	15	25	35	38
	h=1/45,H=1/9	13	18	20	22
	h=1/45,H=1/15	10	12	13	14
	h=1/60,H=1/5	15	22	36	41
	h=1/60,H=1/10	13	17	20	22
	h=1/60,H=1/15	11	14	15	16
Example 2 $\sigma = \sqrt{2}/100$	h=1/45,H=1/15	11	15	20	22
	h=1/60,H=1/15	12	18	24	27
	h=1/60,H=1/20	10	14	17	18

For the symmetric problems, we see that the number of iterations is insensitive to the fine mesh size h, but depends on the coarse mesh size H. The number of iterations is reduced when the coarse mesh size is reduced. This is not true for elliptic problems. The algorithm is not very sensitive to the time step parameter.

For nonsymmetric problems, the algorithm becomes more sensitive to the time step parameter and the coarse mesh size is more important for controlling the number of iterations.

Acknowledgement: I wish to thank my thesis advisor Professor O. Widlund for his inspiring guidance and constant encouragement. I would like also to thank professor M. Dryja for his carefully reading of an early draft.

References

[1] I. Babuška, The Mathematical Foundations of the Finite Element Method, with Applications to Partial Differential Equations (edited by A. D. Aziz), Academic Press, New York and London, 1972.

[2] P. E. Bjørstad and O. B. Widlund, Iterative Methods for the Solution of Elliptic Problems on Regions Partitioned into Substructures, SIAM J. NUMER. ANAL. Vol 23. No. 6, Dec. 1986, pp. 1093-1120.

[3] J. H. Bramble, A Second Order Finite Difference Analogue of the First Biharmonic Boundary Value Problem, Numer. Math. Vol. 9, pp. 236– 249, 1966.

[4] J. H. Bramble, J. E. Pasciak, and A. H. Schatz, The construction of preconditioners for elliptic problems by substructuring, I, Math Comp., Vol. 47, pp. 103-134, 1986.

[5] J. H. Bramble, J. E. Pasciak, and A. H. Schatz, The construction of preconditioners for elliptic problems by substructuring, II, Math. Comp., Vol. 49, pp. 1 - 16, 1987.

[6] J. H. Bramble, J. E. Pasciak, and A. H. Schatz, The construction of preconditioners for elliptic problems by substructuring, III, Math. Comp. Vol. 51, pp. 415 - 430, 1988.

[7] J. H. Bramble, J. E. Pasciak, and A. H. Schatz, The construction of preconditioners for elliptic problems by substructuring, IV, Math. Comp. Vol. 53, pp. 1 - 24, 1989.

[8] X. -C. Cai, Some Domain Decomposition Algorithms for Nonselfadjoint Elliptic and Parabolic Partial Differential Equations, Ph.D. thesis, Courant Institute, Sept. 1989.

[9] X.- C. Cai, An Additive Schwarz Algorithm for Nonselfadjoint Elliptic Equations, in Third International Symposium on Domain Decomposition Methods for Partial Differential Equations, T. Chan, R. Glowinski, J. Périaux, and O. Widlund, eds., SIAM, Philadelphia 1990.

[10] X. -C. Cai and O. B. Widlund, Multiplicative Schwarz Algorithms for Nonsymmetric and Indefinite Elliptic and Parabolic Problems, 1990. (to appear).

[11] C. Dawson, Q. Du and T. F. Dupont, A Finite Difference Domain Decomposition Algorithm for Numerical Solution of the Heat Equation, Tech. Rep., 89 - 09, Univ. of Chicago, Nov. 1989.

[12] M. Dryja, An Additive Schwarz Algorithm for Two- and Three- Dimensional Finite Element Elliptic Problems, in Domain Decomposition Methods for Partial Differential Equations II, T. Chan, R. Glowinski, G. A. Meurant, J. Périaux, and O. Widlund, edt., SIAM, 1989.

[13] M. Dryja and O. B. Widlund, An Additive Variant of the Schwarz Alternating Method for the Case of Many Subregions, Tech. Rep. 339, Dept. of Comp. Sci., Courant Institute, 1987.

[14] M. Dryja and O. B. Widlund, Some Domain Decomposition Algorithms for Elliptic Problems, in Iterative Methods for Large Linear Systems, L. Hayes and D. Kincaid, eds., Academic Press, San Diego, California 1989.

[15] M. Dryja and O. B. Widlund, Towards a Unified Theory of Domain Decomposition Algorithms for Elliptic Problems, in Third International Symposium on Domain Decomposition Methods for Partial Differential Equations, T. Chan, R. Glowinski, J. Périaux, and O. Widlund, eds., SIAM, Philadelphia 1990.

[16] S. C. Eisenstat, H. C. Elman and M. H. Schultz, Variational Iterative Methods for Nonsymmetric System of Linear Equations, SIAM J. Numer. Anal. Vol. 20, April 1983.

[17] R. E. Ewing, R. D. Lazarov, J. E. Pasciak, and P. S. Vassilevski, Finite Element Methods for Parabolic Problems with Time Steps Variable in Space, Tech. Rep. # 1989-05, Inst. for Sci. Comp. at Univ. of Wyoming.

[18] P. Grisvard, Elliptic Problems in Nonsmooth Domains, Pitman, Boston, MA, 1985.

[19] C. Johnson, Numerical Solution of Partial Differential Equation by the Finite Element Method, Cambridge University Press, 1987.

[20] Yu. A. Kuznetsov, Domain Decomposition Methods for Time-Dependent Problems, May 1989, preprint.

[21] P. L. Lions, On the Schwarz Alternating Method. I, in R. Glowinski, G. H. Golub, G. A. Meurant, and J. Périaux, edt., Domain Decomposition Methods for Partial Differential Equations, SIAM, Philadelphia, 1988.

[22] J. Nečas, Sur la Coercivité des Formes Sesquilinéaires, Elliptiques, Rev. Roumaine Math. Pures Appl., Vol. 9, pp. 47 – 69, 1964.

[23] Y. Saad and M. H. Schultz, GMRES: A Generalized Minimal Residual Algorithm for Solving Nonsymmetric Linear Systems, SIAM J. Sci. Stat. Comp., Vol. 7, pp. 865-869, 1986.

[24] O. B. Widlund, Some Domain Decomposition and Iterative Refinement Algorithms for Elliptic Finite Element Problems, J. Comp. Math., Vol. 7, pp. 200 - 208, 1989, Proceedings of the China - U. S. Seminar on Boundary Integral Equations and Boundary Element Methods in Physics and Engineering, held at the Xi'an Jiaotong University, Xi'an, The People's Republic of China, December 27, 1987 - January 1, 1988.

[25] H. Yserentant, On the Multi-level Splitting of Finite Element Spaces, Num. Math., 49, pp. 379 - 412, 1986.