Iterative Solution of Elliptic Finite Element Problems on Partially Refined Meshes and the Effect of Using Inexact Solvers

Hsuanjen Cheng
Courant Institute of Mathematical Sciences
New York University

October, 1993

A dissertation in the Department of Mathematics submitted to the faculty of the Graduate School of Arts and Sciences in partial fulfillment of the requirements for the degree of Doctor of Philosophy at New York University.

Approved: _____

Olof B. Widlund
Adviser

copyright © Hsuanjen Cheng
All Rights Reserved 1994

Abstract

We will consider some solution methods for the linear systems of algebraic equations which arise from second-order elliptic finite element problems. We first consider iterative refinement methods. Historically these methods were called FAC and AFAC methods. Optimal bounds of the condition number for the AFAC iterative operator are established by proving a strengthened Cauchy-Schwarz inequality using an interpolation theorem for Hilbert scales. This work completes earlier work by Dryja and Widlund. We also apply an extension theorem for finite element functions to get a weaker bound under a more general assumption. The optimality of the FAC method when spectrally equivalent inexact solvers are used is also proved by using similar techniques.

We next consider multilevel additive Schwarz methods with partial refinement. These algorithms are generalizations of the multilevel additive Schwarz methods developed by Dryja and Widlund and many others. We use quasi-interpolants to prove that this class of methods has an optimal condition number, which is independent of the selection of refinement subregions, under quite general assumptions.

The third algorithm, which we develop, is a domain decomposition parallel preconditioning method (DD-PPM). The preconditioner is closely related to the construction of an operator providing a uniformly bounded extension from the boundary of the nonoverlapping subdomains to the interior of these subdomains. This method is well suited for the computation of finite element solutions on multiprocessor computer architectures. We also obtain a new condition number estimate for general DD-PPM which is sharper than those previously known by using the framework of additive Schwarz methods.

Finally, we develop an iterative substructuring method which is a variation of Smith's vertex space method by using the similar idea of the previous algorithm. Some numerical experiments illustrate the difference between this algorithm and that of Smith.

Acknowledgements

I would like to thank my advisor, Olof Widlund, for spending so much time and energy on me.

I would also like to thank all my friends who gave me the encouragement and help .

I dedicate this dissertation to my parents and all my good friends.

Contents

1 Introduction **1**

1.1 An Overview 1

1.2 Sobolev Spaces 3

1.2.1 The Norms 3

1.2.2 Some Technical Lemmas 6

1.2.3 Hilbert Scale and Interpolation Theorems 8

1.3 Finite Element Approximation for Elliptic Problems 10

1.4 Substructuring Methods 12

1.5 The Conjugate Gradient Method and Preconditioning 15

2 Domain Decomposition Methods **18**

2.1 Multiplicative Schwarz Methods 18

2.2 Additive Schwarz Methods 23

2.3 Multigrid Methods 25

3 Iterative Refinement Methods **29**

3.1 Introduction 29

3.2 FAC and AFAC Formulations 30

3.3 Optimal Convergence Rate Estimates for FAC and AFAC 36

3.3.1 An Extension Theorem for Finite Element Functions 37

3.3.2 Convergence Rate Estimate for AFAC 39

3.3.3 Convergence Rate Estimate for FAC with Inexact Solvers 48

3.4 Numerical Results 53

4 Multilevel Additive Schwarz Methods with Partial Refinement 60

4.1 Introduction 60

4.2 General Multilevel Additive Schwarz Methods 61

4.2.1 Basic Two-Level Methods 61

4.2.2 General Multilevel Additive Schwarz Methods 63

4.3 Description of Multilevel Additive Schwarz Methods with Partial Refinement and the Condition Number Estimate 65

4.4 Numerical Results 73

5 Domain Decomposition Parallel Preconditioning Methods 77

5.1 Introduction 77

5.2 Description of a General DD-PPM 78

5.3 Abstract Spectral Equivalence Bound Estimates 79

5.4 A new algorithm and some of its variants 85

5.5 Numerical results 94

6 Iterative Substructuring Methods 102

6.1 Description of the Algorithm 102

6.2 Numerical Results 106

Chapter 1

Introduction

1.1 An Overview

Finite element discretizations of elliptic partial differential equations often result in quite large, sparse linear systems. Domain decomposition methods are often used to solve such problems. The domain is decomposed into overlapping or non-overlapping subdomains. In the first case the algorithms are often referred to as Schwarz methods, while the latter are called iterative substructuring methods. A domain decomposition method is thus an iterative method and is different from direct methods of solving linear systems based upon the Cholesky decomposition of the corresponding matrix. In each iteration step, a number of smaller subproblems, which correspond to the restriction of the original problems to subregions, and a problem with a coarse mesh which provides global transportation of information, are solved. These subproblems can often be solved in parallel. Therefore these methods are very promising for parallel computation.

The Schwarz alternating method, proposed by H. A. Schwarz [68] in 1869, is the first domain decomposition algorithm. This method involves solving the boundary value problems on two subdomains sequentially. Recent work by P. L. Lions [46] has interpreted the method for two subdomains cases in a variational framework. It is not ideally suitable for parallel computation because of the sequential nature of the method. In [29], Dryja and Widlund developed the additive Schwarz method which makes it possible to solve all the subproblems in parallel. Matsokin and Nepomnyaschikh [53] also obtained an algorithm which is equivalent to the additive Schwarz method on the interface of non-overlapping subdomains; see also Nepomnyaschikh [60]. We note that the additive Schwarz methods

can be regarded as an iterative method for an equivalent system with a better condition number.

Iterative substructuring methods for second order problems were developed by Bramble, Pasciak and Schatz [11],[14], Dryja and Widlund [31], Smith [71], and Dryja, Smith and Widlund [28]. It has been demonstrated that the Schwarz framework can be used to analyze many iterative substructuring methods. The best of these algorithms have condition numbers which are bounded independently of the number of subdomains and unknowns or grow like $(1 + \log(H/h))$. Here H is the characteristic diameter of the subdomains and h is the characteristic diameter of the elements.

In this dissertation, we will first study iterative refinement methods. These methods are devised to handle elliptic finite element problems defined on composite meshes using only solvers of problems with relative uniform mesh size. The first iterative refinement method was developed by Thomas and McCormick in [54]. They called it the Fast Adaptive Composite (FAC) method. A parallel implementation of FAC led to the introduction of asynchronous FAC (AFAC) methods; cf. Hart and McCormick [42]. We can apply the framework of multiplicative Schwarz methods to analyze FAC and that of additive Schwarz methods to analyze AFAC.

Multilevel methods, e.g. multigrid methods, are among the powerful methods of solving linear systems arising from elliptic problems; cf. Hackbusch [41] or McCormick [56]. Some new multilevel methods have been developed in the past few years, cf. Bramble, Pasciak and Xu [17], Dryja and Widlund [33], and Yserentant [84]. We develop the analog of some multilevel methods to the case of elliptic problems with composite meshes and prove the optimality without any regularity assumption for the corresponding continuous elliptic problems.

The domain decomposition parallel preconditioning method (DD-PPM) is based on block Gaussian elimination of the stiffness matrix. It is similar to the iterative substructuring method because the subdomains are non-overlapping. However, we not only need good preconditioners for Schur complement and each subproblems, but also a good approximate solver for inhomogeneous Dirichlet subproblems with a solution which has a uniformly bounded energy when compared with the discrete harmonic function having the same boundary value on the boundaries of subdomains. We remark that there are

no problems for overlapping domain decomposition methods if we replace the local subdomain solvers by approximate subdomain solvers, but that it is much more difficult to find appropriate inexact solvers of subdomains for nonoverlapping domain decomposition methods. Börgers [7] first analyzed the convergence of a Dirichlet DD-PPM method and gave a rough bound for the condition number. There has also been interesting work on these methods by Hasse, Langer and Meyer [37], [38] and Meyer [58].

This dissertation is organized as follows. In the remainder of Chapter 1, we review some basic definitions and useful results about Sobolev spaces including Hilbert scales and interpolation theorems. We also discuss some standard finite element discretizations for second order elliptic problems and iterative methods for solving linear systems of equations. In Chapter 2, we discuss the classical alternating Schwarz method, additive Schwarz methods and some basic multigrid theory. In Chapter 3, we describe some iterative refinement methods in their original form and discuss some variations. We then prove the optimality of these methods by using an extension theorem and an interpolation theorem based upon Hilbert scales of finite element functions. In Chapter 4, we discuss the multilevel additive Schwarz method of Dryja and Widlund and the BPX algorithm of Bramble, Pasciak, and Xu in the case of partial refinement. This work is an extension of that of Zhang [86]. However, we use a quasi-interpolant to prove the optimality of these methods in an approach different from Zhang's. In Chapter 5, we construct good DD-PPM preconditioners by using a norm-preserving extension of finite element functions, which is due to Matsokin [52]. We can then give a good condition number estimate for the DD-PPM methods using the framework of additive Schwarz methods. In Chapter 6, we apply an idea similar to that of Chapter 5 to derive a variation of Smith's vertex space method in two dimensional space and then compare the two algorithms in numerical experiments.

1.2 Sobolev Spaces

1.2.1 The Norms

Let Ω be a bounded, Lipschitz domain in R^n and let $L^p(\Omega)$ be the Banach space of p-th power integrable function on Ω.

We use the multi-index notation for partial derivatives. Let $\alpha \equiv (\alpha_1, \cdots, \alpha_n)$ be a

nonnegative multi-index representing the order of the partial differentiation, set $|\alpha| \equiv \alpha_1 + \cdots + \alpha_n$, and let

$$D^\alpha \equiv \frac{\partial^\alpha}{\partial x^\alpha} \equiv \frac{\partial^\alpha}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}$$

be the differentiation operator corresponding to α in weak sense.

The Sobolev space $W^{m,p}(\Omega)$ is defined by

$$W^{m,p}(\Omega) = \{u | D^\alpha u \in L^p(\Omega), \forall |\alpha| \leq m\}$$

with a norm given by

$$\|u\|^p_{W^{m,p}(\Omega)} = \sum_{|\alpha| \leq m} \|D^\alpha u\|^p_{L^p(\Omega)}.$$

When $p = 2$, $W^{m,p}(\Omega) \equiv H^m(\Omega)$ becomes a Hilbert space with the inner product

$$(u, v)_{H^m(\Omega)} = \sum_{|\alpha| \leq m} (D^\alpha u, D^\alpha v)_{L^2(\Omega)}.$$

It can be shown that the space $H^m(\Omega)$ is the completion of $C^\infty(\overline{\Omega})$ with respect to H^m-norm, cf. Adams [1] . Similarly, we can define $H_0^m(\Omega)$ as the completion of $C_0^\infty(\Omega)$ under the H^m-norm.

We will define Sobolev spaces of fractional order in a different way. If $0 < s < 1$ and $1 \leq p < \infty$, we can define

$$W^{s,p}(\Omega) = \{u \in L^p(\Omega) | \frac{|u(x) - u(y)|}{|x - y|^{s+n/p}} \in L^p(\Omega \times \Omega)\}.$$

These are Banach spaces with the norm given by

$$\|u\|^p_{W^{s,p}(\Omega)} = \|u\|^p_{L^p(\Omega)} + |u|^p_{W^{s,p}(\Omega)},$$

where the last term is the seminorm given by

$$|u|^p_{W^{s,p}(\Omega)} = \int_\Omega \int_\Omega \frac{|u(x) - u(y)|^p}{|x - y|^{n+ps}} dx\, dy.$$

As before, we use the notation $H^s(\Omega) = W^{s,2}(\Omega)$. Finally, the Sobolev space $W^{s,p}(\Omega)$ with $s \in R$ not an integer, $s > 1$ is defined as follows. Write $s = m + \sigma$, with $m =$ integer part of s and define

$$W^{s,p}(\Omega) = \{u \in W^{m,p}(\Omega) | D^\alpha u \in W^{\sigma,p}(\Omega), \forall \alpha \text{ with } |\alpha| = m\}.$$

Finally, we define $H_0^s(\Omega)$ as the completion of $C_0^\infty(\Omega)$ with respect the H^s-norm. It can be shown that $H^s(\Omega)$ and $H_0^s(\Omega)$ are the same for $0 \leq s < 1/2$, but differ for $1/2 < s \leq 1$; cf. Lions and Magenes [45].

For completeness, we will also define Sobolev spaces on the boundary $\partial\Omega$. These spaces can be defined on manifolds which are parts of the boundary of a Lipschitz domain in R^n; see Grisvard [40]. We only give the definition of $H^s(\partial\Omega)$ for $0 < s < 1$. First, we define

$$L^2(\partial\Omega) = \{u | \int_{\partial\Omega} |u|^2 ds < \infty\},$$

with the natural norm

$$\|u\|_{L^2(\partial\Omega)} = (\int_{\partial\Omega} |u|^2 ds)^{1/2}.$$

Then we can define

$$H^s(\partial\Omega) = \{u \in L^2(\partial\Omega) | \ \ |u|_{H^s(\partial\Omega)} < \infty\},$$

where

$$|u|^2_{H^s(\partial\Omega)} = \int_{\partial\Omega}\int_{\partial\Omega} \frac{|u(x) - u(y)|^2}{|x - y|^{n-1+2s}} ds_x ds_y.$$

The norm for this space is given by

$$\|u\|^2_{H^s(\partial\Omega)} = \|u\|^2_{L^2(\partial\Omega)} + |u|^2_{H^s(\partial\Omega)}.$$

We now introduce the concept of trace maps. We have an obvious definition of the boundary value, or trace on $\partial\Omega$, for smooth functions on $\overline{\Omega}$. This can be generalized to functions in $H^s(\Omega)$ under specific hypotheses on $\partial\Omega$; see Nečas [61] or Lions and Magenes [45]. We consider only $1/2 < s \leq 1$ for simplicity.

Theorem 1.2.1 (Trace theorem) *Let $1/2 < s \leq 1$. The trace map $\gamma : u \to u|_{\partial\Omega}$ defined from $C^\infty(\overline{\Omega})$ to $L^2(\partial\Omega)$ can be extended by density to a linear, continuous operator from $H^s(\Omega)$ to $L^2(\partial\Omega)$. Moreover, the range of γ is $H^{s-1/2}(\partial\Omega)$ and*

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

We remark that $H_0^s(\Omega)$ can be regarded as the subspace of $H^s(\Omega)$ with zero trace on $\partial\Omega$ for $1/2 < s \leq 1$.

Sobolev spaces of negative order can also be defined as follows. $H^{-s}(\Omega)$, with $s > 0$, is defined to be the dual space of $H_0^s(\Omega)$. Therefore its norm is given by

$$\|f\|_{H^{-s}(\Omega)} \equiv \sup_{u \in H_0^s(\Omega)} \frac{|f(u)|}{\|u\|_{H^s(\Omega)}}, \quad \forall f \in H^{-s}(\Omega).$$

By using the fact that each Hilbert space is reflexive, we also get

$$\|u\|_{H^s(\Omega)} = \sup_{f \in H^{-s}(\Omega)} \frac{|f(u)|}{\|f\|_{H^{-s}(\Omega)}}, \quad \forall u \in H_0^s(\Omega).$$

Each element $v \in L^2(\Omega)$ determines an element L_v of $H^{-s}(\Omega)$ by means of $L_v(u) = (u, v)_{L^2(\Omega)}$. By using the reflexivity of $H_0^s(\Omega)$, it is easy to see that $L^2(\Omega)$ is in fact dense in $H^{-s}(\Omega)$. Then the previous formula can be written as

$$\|u\|_{H^s(\Omega)} = \sup_{v \in L^2(\Omega)} \frac{|(v, u)_{L^2(\Omega)}|}{\|v\|_{H^{-s}(\Omega)}}, \quad \forall u \in H_0^s(\Omega).$$

1.2.2 Some Technical Lemmas

We now introduce some technical lemmas, which will often be used in this dissertation. Throughout, we make the same assumption on Ω as in last subsection.

The following three lemmas establish equivalences of certain norms in some subspaces of $H^1(\Omega)$.

Lemma 1.2.1 (Friedrichs' inequality) *There exists a positive constant $C(\Omega)$, which depends only on the Lipschitz constant of $\partial\Omega$, such that, for all $u \in H_0^1(\Omega)$,*

$$\|u\|_{L^2(\Omega)} \leq C(\Omega) H_\Omega |u|_{H^1(\Omega)}.$$

Here H_Ω is the diameter of Ω.

Lemma 1.2.2 (Poincaré's inequality) *Let*

$$\{u\}_\Omega = \frac{1}{|\Omega|} \int_\Omega u.$$

Then there exists a constant $C(\Omega)$, which depends only on the Lipschitz constant of $\partial\Omega$, such that for all $u \in H^1(\Omega)$ we have

$$\|u - \{u\}_\Omega\|_{L^2(\Omega)} \leq C(\Omega) H_\Omega |u|_{H^1(\Omega)}.$$

Here H_Ω is the diameter of Ω.

Lemma 1.2.3 (Poincaré-Friedrichs' inequality) *Let $\Gamma_D \subset \partial\Omega$ be of positive measure. Then there exists a constant $C(\Omega, \Gamma_D)$, which depends on Ω and Γ_D, such that, for all $u \in H^1(\Omega)$,*

$$\|u\|^2_{L^2(\Omega)} \leq C(\Omega, \Gamma_D)(|u|^2_{H^1(\Omega)} + \int_{\Gamma_D} (\gamma u)^2).$$

Here γ is the trace operator.

Proofs of Poincaré's inequality may be found in [61] and [62], a proof of Friedrichs' inequality in [62] and of Poincaré-Friedrichs' inequality in [61].

We also need to evaluate the constant $C(\Omega, \Gamma_D)$ of Poincaré-Friedrichs' inequality in some special cases. The proof is based on using smooth functions to approximate elements in $H^1(\Omega)$ and applying the fundamental theorem of calculus to the region.

Lemma 1.2.4 *Let Ω be a domain in R^2, which has the following special form : $\{(x_1, x_2) | a < x_1 < b, g(x_1) < x_2 < f(x_1)\}$. Here $f(x)$ and $g(x)$ are piecewise C^1, continuous functions on $[a, b]$ such that $g(x) < f(x)$ on (a, b). Then for all $u \in H^1(\Omega)$ vanishing on $\{(x_1, x_2) | a \leq x_1 \leq b, x_2 = g(x_1)\}$, we have*

$$\|u\|_{L^2(\Omega)} \leq \max_{a \leq x \leq b} |f(x) - g(x)| \cdot |u|_{H^1(\Omega)}.$$

We can also get similar inequalities in R^n for $n > 2$.

There is a variation of Lemma 1.2.2 in fractional Sobolev spaces. Its proof can be established from the definition of Sobolev norms.

Lemma 1.2.5 *Let*

$$\{u\}_\Omega = \frac{1}{|\Omega|} \int_\Omega u.$$

Then there exists a constant $C(\Omega)$, which depends only on the domain Ω, such that for all $u \in H^s(\Omega)$, with $0 < s < 1$, we have

$$\|u - \{u\}_\Omega\|_{L^2(\Omega)} \leq C(\Omega)|u|_{H^s(\Omega)}.$$

In order to prove extension theorems for finite element functions, which are used to establish the optimality of iterative refinement method, we need the following theorem from Stein [75].

Lemma 1.2.6 (Stein's extension theorem) *There exists a linear operator $\mathcal{E}$ extending functions on Ω to functions on R^n with the properties*

i) $\mathcal{E}u|_\Omega = u$,

ii) $\|\mathcal{E}u\|_{H^1(R^n)} \leq C(\Omega)\|u\|_{H^1(\Omega)}$.

Here the constant $C(\Omega)$ depends only on the Lipschitz constant of $\partial\Omega$.

1.2.3 Hilbert Scale and Interpolation Theorems

In this subsection, we describe a procedure of constructing intermediate spaces between two Hilbert spaces and state the interpolation theorems which are needed to establish the upper bound of the iteration operator for the iterative refinement methods.

Let H be a Hilbert space and $T : \mathcal{D}(T) \subset H \to H$ be a linear operator, where $\mathcal{D}(T)$ is a dense linear subspace of H. We wish to associate an adjoint T^* to T. Its domain $\mathcal{D}(T^*)$ consists of all $v \in H$ for which the linear functional

$$u \to (Tu, v)$$

is continuous on $\mathcal{D}(T)$. If $v \in \mathcal{D}(T^*)$, we can apply the Hahn-Banach theorem to extend the functional to a continuous linear functional on H. Therefore there exists an element $T^*v \in H$ that satisfies

$$(Tu, v) = (u, T^*v) \quad \forall u \in \mathcal{D}(T)$$

by Riesz representation theorem. If $T = T^*$, then T is said to be self-adjoint.

We now recall the procedure of defining Hilbert scales. Let X and Y be two separable Hilbert spaces with $X \subset Y$ and let X be dense in Y with a continuous injection.

Let $(\ ,\ )_X$ and $(\ ,\ )_Y$ be the scalar products in X and Y, respectively. We denote by $\mathcal{D}(T)$ the set of $u's$ such that the antilinear form

$$v \to (u, v)_X, \quad v \in X$$

is continuous in the topology induced by Y . Then

$$(u, v)_X = (Tu, v)_Y$$

defines T as a generally unbounded operator in Y, with domain $\mathcal{D}(T)$.

It can be shown that $\mathcal{D}(T)$ is dense in Y and that T is a strictly positive self-adjoint operator. Using the resolution of the identity for self-adjoint operators, cf. Dunford and Schwartz [35] , we can define $f(T)$ as a self-adjoint operator in H for all Borel-measurable functions f on R. In particular, the powers T^θ of T, $\theta \in R$, can be defined.

Let

$$\Lambda = T^{1/2}.$$

The operator Λ is self-adjoint and positive in Y, with domain X, and satisfies

$$(u, v)_X = (\Lambda u, \Lambda v)_Y, \quad \forall u, v \in X.$$

We give the following definition of the intermediate spaces $[X, Y]_\theta$, cf. Lions and Magenes [45].

Definition 1.2.1 *Let*

$$[X, Y]_\theta = \mathcal{D}(\Lambda^{1-\theta}), \qquad 0 \le \theta \le 1,$$

with norm on $[X, Y]_\theta = \|\Lambda^{1-\theta} u\|_Y$.

With the above definition, we now state a useful interpolation theorem. Its proof can be found in Lions and Magenes [45], and Butzer and Berens [19] .

Theorem 1.2.2 *Let* $[X_1, Y_1]_\theta$ *and* $[X_2, Y_2]_\theta$ *be two Hilbert scales. If* $A \in L(X_1, X_2)$ *and* $A \in L(Y_1, Y_2)$, *then* $A \in L([X_1, Y_1]_\theta, [X_2, Y_2]_\theta)$ *for all* $\theta \in [0, 1]$ *and*

$$\|A\|_{[X_1,Y_1]_\theta \to [X_2,Y_2]_\theta} \le \|A\|^{1-\theta}_{X_1 \to X_2} \cdot \|A\|^{\theta}_{Y_1 \to Y_2}.$$

Let $H^s(\Omega)$ be the Sobolev spaces for $0 \le s \le 1$. It can be shown that these spaces almost form a Hilbert scale. "Almost" means the following. If we set $X = H^1(\Omega)$ and $Y = H^0(\Omega) = L^2(\Omega)$, then $[X, Y]_{1-s} = H^s(\Omega)$. Furthermore, there exist constants $C_1(\Omega, s)$ and $C_2(\Omega, s)$, depending only on Ω and s, such that

$$C_1(\Omega, s)\| \cdot \|_{H^s(\Omega)} \le \| \cdot \|_{[X,Y]_{1-s}} \le C_2(\Omega, s)\| \cdot \|_{H^s(\Omega)}.$$

The proof can be found in Lions and Magenes [45]. By the above remark and Theorem 1.2.2, we have the following corollary.

Corollary 1.2.7 *Let A be a continuous linear functional over R on $L^2(\Omega)$ and $H^1(\Omega)$. Then there exists a constant $C(\Omega, s)$, depending only on Ω and s, such that*

$$\|A\|_{H^s(\Omega)\to R} \le C(\Omega, s) \cdot \|A\|^s_{H^1(\Omega)\to R} \cdot \|A\|^{1-s}_{L^2(\Omega)\to R}$$

We observe that if M is a finite dimensional subspace in $H^1(\Omega)$, then from the definition of Hilbert scales M, equipped with the $H^s(\Omega)$ norm for $0 \le s \le 1$, also almost forms a Hilbert scale. Furthermore, it can be shown that the new induced Hilbert scale norm is the same as the Hilbert scale norm obtained by applying the procedure to the whole space for all s. Therefore we can modify the corollary to get the following lemma .

Lemma 1.2.8 *Let M be a finite-dimensional subspace of $H^1(\Omega)$ and let A be a continuous linear functional over R on M both in $H^1(\Omega)$ and $L^2(\Omega)$. Then there exists a constant $C(\Omega, s)$, which is the same as in Corollary 1.2.7, such that*

$$\|A\|_{M\subset H^s(\Omega)\to R} \le C(\Omega, s) \cdot \|A\|^s_{M\subset H^1(\Omega)\to R} \cdot \|A\|^{1-s}_{M\subset L^2(\Omega)\to R}.$$

1.3 Finite Element Approximation for Elliptic Problems

We are interested in partial differential equation of the form

$$-\sum_i \sum_j \frac{\partial}{\partial x_i} a_{ij}(x) \frac{\partial u}{\partial x_j} = f \qquad \text{in } \Omega, \tag{1.1}$$

$$u = u_0 \qquad \text{on } \Gamma_D \subset \partial\Omega, \tag{1.2}$$

$$\sum_j \sum_i a_{ij}(x) \frac{\partial u}{\partial x_j} n_i = g \qquad \text{on } \Gamma_N = \partial\Omega \backslash \Gamma_D. \tag{1.3}$$

Here the matrix $\{a_{ij}(x)\}$ is positive definite with a uniform lower bound c for almost everywhere x in Ω. Each $a_{ij}(x)$ is a bounded measurable function in Ω and $\vec{n}$ is the unit outward normal to $\partial\Omega$. We assume that the measure of Γ_D is strictly greater than zero. This insures a unique solution to problem (1.1-1.3).

We will assume, without loss of generality, that $u_0 = 0$. If not, we can always substract an arbitrary function w that equals u_0 on Γ_D from u. Let

$$H^1_D(\Omega) = \{u \in H^1(\Omega) | \gamma u = 0 \quad \text{on } \Gamma_D\}.$$

The standard weak formulation of problems of (1.1-1.3) consists of finding $u^* \in V = H^1_D(\Omega)$ such that

$$a(u^*, v) = f(v) \qquad \forall v \in V, \tag{1.4}$$

where

$$a(u,v) = \int_\Omega \sum_{i,j} a_{ij}(x) \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} dx \qquad \text{and} \qquad f(v) = \int_\Omega f v dx + \int_{\Gamma_N} g v ds.$$

For any symmetric, positive definite problem a related minimization problem is given by

$$\min_{u \in V} J(u) = \min_{u \in V} (\frac{1}{2} a(u,u) - f(u)). \tag{1.5}$$

Using Poincaré-Friedrichs' inequality, we can prove that $a(\cdot,\cdot)$ is strongly elliptic and equivalent to the H^1-norm in the space $H^1_D(\Omega)$. By the Lax-Milgram Theorem, the boundness and strong ellipticity of the bilinear form $a(u,v)$ imply existence and uniqueness of the solution; cf. Lax and Milgram [44] or Ciarlet [24].

The finite element formulation is obtained by replacing the infinite dimensional space V with a finite dimensional subspace $V^h \subset V$. We triangulate the domain Ω into non-overlapping regions called elements, generally triangles, rectangles, tetrahedra, prisms or bricks. We always assume that the intersection of any two different elements in the triangulation is empty, a point or a whole edge of a triangle in two dimension. A similar assumption is made in three dimensional spaces. V^h is then defined to be the subspace of continuous functions on Ω which are piecewise polynomials with respect to the triangulation. We usually use Lagrange elements for second order elliptic problems. The Lagrange element of type p is the space of piecewise polynomials with degree less than or equal to p for which there are no degrees of freedom involving derivatives. The finite element solution $u_h \in V^h$ satisfies

$$a(u_h, \phi_h) = f(\phi_h), \qquad \forall \phi_h \in V^h. \tag{1.6}$$

We note that the finite element solution u_h is called the Galerkin approximation and is the $a(\cdot,\cdot)$- projection of the solution u^* onto the finite element space V^h. For a fuller discussion of the finite element method and its analysis ; cf. Ciarlet [24] or Strang and Fix [77].

Let $\{\phi_i\}$ be the nodal basis for V^h. Then u_h can be represented as

$$u_h = \sum_i x_i \phi_i.$$

We obtain a linear system for x, the degrees of freedom of u_h,

$$Kx = b$$

with the stiffness matrix given by

$$K_{i,j} = a(\phi_i, \phi_j)$$

and the load vector by

$$b_i = f(\phi_i).$$

The stiffness matrix K is symmetric, positive definite. We can use the conjugate gradient algorithm to solve such linear systems. However, the rate of convergence deteriorates very quickly as the diameter h of elements tends to zero. Therefore we need find good preconditioners for K to apply the preconditioned conjugate gradient algorithm. It is the main topic in this dissertation.

1.4 Substructuring Methods

The domain decomposition methods provide preconditioners of linear systems of equations representing the discretization of elliptic problems. Preconditioners can be associated with matrix splittings. Some of them can best be understood in terms of the substructuring ideas, which have provided a very successful framework for the development of very large programming systems for structural mechanics computing, cf. Przemieniecki [66].

We consider the same elliptic problems as in last section and also use the notations of the previous section. The triangulation of Ω is introduced in the following way. The region is divided into non-overlapping substructures Ω_i, $i = 1, \cdots, N$. To simplify the description, our study is confined to triangular (simplicial) substructures. All the substructures Ω_i are further divided into elements. The common assumption in finite element theory that all elements are shape regular is adopted and the same assumption is made concerning the substructures. On the element level this means that there is a bound on h_E/ρ_E, which is

independent of the number of degrees of freedom and of E. Here h_E is the diameter of the element E and ρ_E the diameter of the largest sphere that can be inscribed in E.

By using ideas of structural engineering, the stiffness matrix K can be constructed in the following way. The elements of K are given by

$$K_{lm} = a(\phi_l, \phi_m),$$

where ϕ_l and ϕ_m are standard finite element basis functions. Since an integral over Ω can be written as a sum of integrals over the substructures, the stiffness matrix can be assembled from the stiffness matrices $K^{(i)}$ whose entries are defined by

$$K^{(i)}_{lm} = a_{\Omega_i}(\phi_l, \phi_m),$$

with l and m corresponding to the degrees of freedom of the substructure Ω_i. The form $a_{\Omega_i}(u_h, v_h)$ represents the contribution to the integral from that substructure .

This so called subassembly process can be summarized in the formula

$$x^T K y = \sum_{i=1}^{N} {x^{(i)}}^T K^{(i)} y^{(i)},$$

where $x^{(i)}$ is the subvector of parameter values associated with the substructure Ω_i and its boundary $\partial\Omega_i$.

If we divide the subvectors $x^{(i)}$ associated with the i-th substructure into two, $x_I^{(i)}$ and $x_B^{(i)}$, corresponding to the variables shared with other substructures and those which are interior to the substructure, then the matrix $K^{(i)}$ can be written as

$$\begin{pmatrix} K_{II}^{(i)} & K_{IB}^{(i)} \\ {K_{IB}^{(i)}}^T & K_{BB}^{(i)} \end{pmatrix}.$$

Since the interior variables are associated with only one of the substructures, they can be eliminated locally and in parallel. The reduced matrix is a so-called Schur complement and has the form

$$S^{(i)} = K_{BB}^{(i)} - {K_{IB}^{(i)}}^T {K_{II}^{(i)}}^{-1} K_{IB}^{(i)}.$$

It is now easy to show that if the corresponding Schur complement of the global stiffness matrix K is denoted by S, then

$$x_B^T S y_B = \sum_{i=1}^{N} {x_B^{(i)}}^T S^{(i)} y_B^{(i)}. \tag{1.7}$$

Therefore we can obtain the Schur complement S after the elimination of the interior variables from the substructures if we carry out a certain block Gaussian elimination of the stiffness matrix K. An iterative substructuring method is obtained by selecting a preconditioner for the matrix S. Once an approximation of the solution has been found on the boundaries of the substructures, the solution can be found everywhere by solving local problems on each substructure separately.

In the case of two substructures, there is only one set of interface variables and S is the sum of $S^{(1)}$ and $S^{(2)}$, the Schur complements originating from the two subregions, respectively. The basic Neumann-Dirichlet algorithm amounts to using $S^{(1)}$ or $S^{(2)}$ as a preconditioner for S; cf. Bjørstad and Widlund [6]. This algorithm can be extended directly to the case of many substructures if a red-black ordering of the substructures exists; cf. Dryja, Proskurowski and Widlund [27] and Widlund [79]. It is shown in Widlund [79] that the condition number of the resulting preconditioner grows in proportional to $(1 + \log(H/h))^2$. The algorithm is less attractive in the three dimensional case, since the number of variables, which have to be treated in a special way, are relatively more numerous than for a problem in the plane. However, the same kind of bound can be established on the rate of convergence. There is a "face based" iterative substructuring method in Dryja, Smith and Widlund [28] which has the condition number growing in proportional to $(1 + \log(H/h))$. For other iterative substructuring algorithms in three dimension which are all wire-basket methods, we refer to [26], [14] and [72].

In Chapter 5 of this dissertation, we consider the effect of using inexact solvers. In fact, we can summarize which kind of inexact solvers we actually need. In the case of two substructures and using the Neumann-Dirichlet algorithm, we observe that

$$x_B^{(1)^T} S^{(1)} x_B^{(1)} = \begin{pmatrix} -{K_{II}^{(1)}}^{-1} K_{IB}^{(1)} x_B^{(1)} \\ x_B^{(1)} \end{pmatrix}^T K^{(1)} \begin{pmatrix} -{K_{II}^{(1)}}^{-1} K_{IB}^{(1)} x_B^{(1)} \\ x_B^{(1)} \end{pmatrix}.$$

If we find a matrix $B_{II}^{(1)}$ such that the replacement of ${K_{II}^{(1)}}^{-1}$ by ${B_{II}^{(1)}}^{-1}$ results in an operator $R^{(1)}$ which is spectrally equivalent to $S^{(1)}$, then $R^{(1)}$ would also be a good preconditioner of S. A good choice of $B_{II}^{(1)}$ is just a good approximate solver for an inhomogeneous Dirichlet subproblem of Ω_1 in the sense that the solution has a uniformly bounded energy relative to the energy of exact solution.

1.5 The Conjugate Gradient Method and Preconditioning

The conjugate gradient (CG) method is of fundamental importance for the domain decomposition methods. All algorithms studied in this dissertation are variants of the preconditioned conjugate gradient method (PCG). Since CG and PCG are well known algorithms, we will only give a brief description. For a more complete understanding, see Golub and Van Loan [39]. An essential feature of the conjugate gradient method is that an explicit representation of A is not needed: it is enough to be able to apply it to a given vector. We have the following basic form of the conjugate gradient method.

The Conjugate Gradient Algorithm

Set $k = 0; x_0 = 0; r_0 = b$.
while $|r_k| \geq \epsilon |r_0|$
    $k = k + 1$
    if $k = 1$
        $p_1 = r_0$
    else
        $\beta_k = (r_{k-1}, r_{k-1})/(r_{k-2}, r_{k-2})$
        $p_k = r_{k-1} + \beta_k p_{k-1}$
    end
    $\alpha_k = (r_{k-1}, r_{k-1})/(p_k, Ap_k)$
    $x_k = x_{k-1} + \alpha_k p_k$
    $r_k = r_{k-1} - \alpha_k A p_k$
end

Here (x, y) is the ordinary inner product $x^T y$ in R^N. We also define $(x, y)_A$ by the A-inner product $x^T A y$ in R^N. It is a well known result that in the absence of round-off errors, the conjugate gradient method gives the exact solution after at most N steps because the search directions p_k are A-conjugate:

$$(p_i, Ap_j) = 0, \quad \forall i \neq j.$$

The subspace

$$W_k(A, r_0) = \text{span}\{r_0, Ar_0, \cdots, A^{k-1} r_0\} = \text{span}\{p_0, p_1, \cdots, p_{k-1}\}$$

is called the Krylov subspace related to the matrix A and r_0. An important observation is that $\|x - x_k\|_A$ is the norm of the difference between the initial error $x - x_0$ and its

A-orthogonal projection on W_k. The reduction in A-norm of the error, after n steps of conjugate gradient iteration, is given by

$$\|x - x_n\|_A \leq 2(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1})^n \|x - x_0\|_A,$$

for proof see e.g. Luenberger [47]. Here $\kappa = \kappa(A)$ is the condition number of A given by

$$\kappa = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}.$$

The conjugate gradient method is therefore often an effective iterative algorithm to solve symmetric, positive definite systems

$$Ax = b.$$

The algorithm converges to a good approximate solution in relatively few iterations for a well conditioned matrix A. When A is not well conditioned, which is generally the case for discretizations of elliptic problems, we can introduce a preconditioner B and solve the preconditioned linear system

$$B^{-1}Ax = B^{-1}b$$

by using the conjugate gradient method with the A-inner product. We should choose B such that $\kappa(B^{-1}A) << \kappa(A)$ and the system $Bz = d$ is computationally easier to solve than $Az = d$. Standard preconditioners used in practice are diagonal scaling and incomplete factorization, see e.g. [57] for detail.

We can compute the condition number of A approximately during the conjugate gradient iteration by using a variant of an eigenvalues algorithm due to Lanczos [43]. Let R_K be the $N \times k$ matrix with columns chosen as the normalized residual vectors. We can prove that

$$R_k^T A R_k = J_k = \begin{pmatrix} 1/\alpha_1 & -\sqrt{\beta_2}/\alpha_1 & & & \\ -\sqrt{\beta_2}/\alpha_1 & 1/\alpha_2 + \beta_2/\alpha_1 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -\sqrt{\beta_k}/\alpha_{k-1} \\ & & & -\sqrt{\beta_k}/\alpha_{k-1} & 1/\alpha_k + \beta_k/\alpha_{k-1} \end{pmatrix}$$

is tridiagonal. J_k is the matrix representation of the operator A restricted to the subspace spanned by the residual vectors $r_0, \cdots, r_{k-1}$. Approximations of the eigenvalues of A can

be obtained from the eigenvalues of J_k at a small extra cost because J_k is symmetric and tridiagonal. See Parlett [64] for a complete discussion of the Lanczos method.

Chapter 2

Domain Decomposition Methods

2.1 Multiplicative Schwarz Methods

The earliest domain decomposition method is the Schwarz alternating method, which was proposed by H. A. Schwarz in 1869, [68]. The convergence of this method was established by using the *maximum principle*. In the 1930's, Sobolev extended the result to the partial differential equations of linear elasticity [74]. We consider the domain, shown in Figure 2.1, with $\Omega = \Omega_1 \cup \Omega_2$ on which we wish to solve

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

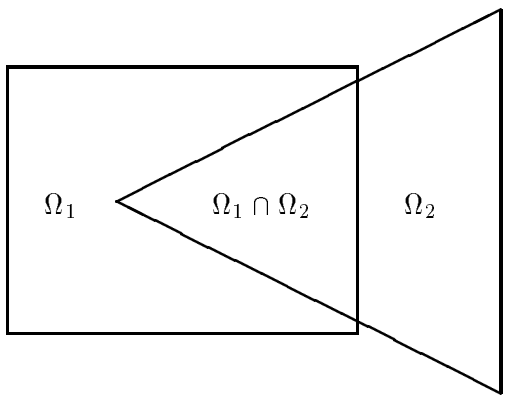

Figure 2.1: Schwarz alternating method

Given an initial guess u_0 of solution u, the Schwarz alternating method computes the

solution u iteratively by solving problems on the each subdomain Ω_i.

$$\begin{cases} -\Delta u^{n+1/2} = f & \text{in } \Omega_1, \\ u^{n+1/2} = u^n & \text{on } \partial\Omega_1, \end{cases}$$

and

$$\begin{cases} -\Delta u^{n+1} = f & \text{in } \Omega_2, \\ u^{n+1} = u^{n+1/2} & \text{on } \partial\Omega_2. \end{cases}$$

The algorithm has been interpreted in a variational framework by Lions [46]. If we let

$$a(u,v) = \int_\Omega \nabla u \cdot \nabla v dx \quad \text{and} \quad f(v) = \int_\Omega f v dx,$$

then the algorithm can be written as

$$a(u^{n+1/2} - u^n, v) = f(v) - a(u^n, v), \qquad u^{n+1/2} - u^n \in H_0^1(\Omega_1), \forall v \in H_0^1(\Omega_1),$$

and

$$a(u^{n+1} - u^{n+1/2}, v) = f(v) - a(u^{n+1/2}, v), \qquad u^{n+1} - u^{n+1/2} \in H_0^1(\Omega_2), \forall v \in H_0^1(\Omega_2).$$

Let us define the orthogonal projection $P_i : H_0^1(\Omega) \to H_0^1(\Omega_i)$, by

$$a(P_i u, v) = a(u, v), \quad \forall v \in H_0^1(\Omega_i).$$

In each half-step, the correction is thus the orthogonal projection of the error onto the subspace $H_0^1(\Omega_1)$ or $H_0^1(\Omega_2)$. Therefore the error propagation equation is

$$e_{n+1} = (I - P_2)(I - P_1)e_n$$

where $e_n = u - u_n$.

We can regard the alternating Schwarz method as a particular case of an abstract multiplicative Schwarz method. We consider more general variational problem on finite dimensional space V: Find $u \in V$ such that

$$a(u,v) = f(v), \quad \forall v \in V. \tag{2.1}$$

Let V_i be subspaces of V so that $V = V_1 + \cdots + V_N$. We can associate each subspace V_i with the corresponding orthogonal projection $P_i : V \to V_i$. The basic abstract multiplicative Schwarz method proceeds by sequentially projecting the error onto the subspaces V_i. The

error propagation operator for one complete step of the abstract multiplicative Schwarz method is given by

$$(I - P_N) \cdots (I - P_1).$$

New tools for the analysis of multiplicative Schwarz methods were developed by Lions [46], Nepomnyaschikh [60], Mandel and McCormick [49] and Bjørstad and Mandel [3]. The most useful results were obtained by Bramble, Pasciak, Wang and Xu [16]. We will review some of their abstract results and then use similar ideas to analyze the multiplicative variants of the algorithms of this dissertation.

We can replace the above P_i by an approximate projection T_i on V_i which is often taken to be spectrally equivalent to P_i. Then the corresponding error propagation operator is

$$E_N = (I - T_N) \cdots (I - T_1).$$

The convergence rate of the multiplicative algorithm is therefore determined by the reduction γ of the error in each iteration :

$$\|E_N u\|_a \leq \gamma \|u\|_a.$$

Therefore this algorithm can be viewed as a simple iterative method for solving

$$(I - (I - T_N) \cdots (I - T_1)) u_h = g_h,$$

with an appropriate right-hand side g_h.

We can consider another way of constructing the above T_i. Let $b_i(u, v)$ be an inner product defined on $V_i \times V_i$. We assume that there exists a constant ω such that

$$a(u_i, u_i) \leq \omega b_i(u_i, u_i), \quad \forall u_i \in V_i. \tag{2.2}$$

We can define the operator T_i, which replaces P_i, by

$$b_i(T_i u, v_i) = a(u, v_i), \quad \forall v_i \in V_i. \tag{2.3}$$

It is easy to show that the operator T_i defined above is positive semidefinite and symmetric with respect to $a(\cdot, \cdot)$ and that the minimal constant ω in equation (2.2) is $\|T_i\|_a$. In multiplicative algorithms, we scale T_i and $b_i(\cdot, \cdot)$ such that $0 < \omega < 2$.

In order to estimate the rate of convergence γ of abstract multiplicative Schwarz methods and additive Schwarz methods, which will be introduced in next section, we need the following lemma. For completeness, we also include the proof. We note that the operator T introduced in this lemma plays a central role in the additive Schwarz method; cf. Section 2.2.

Lemma 2.1.1 *Let $T = T_1 + T_2 + \cdots + T_N$. Then T is invertible and*

$$a(T^{-1}u, u) = \min_{u=\sum_i u_i} \sum_{i=1}^{N} b_i(u_i, u_i),$$

with $u_i \in V_i$. The minimum is achieved at $u_i = T_i T^{-1} u$.

Proof. Let us first prove that T is invertible. Suppose that $Tu = 0$. Then

$$0 = a(Tu, u) = \sum_{i=1}^{N} a(T_i u, u) = \sum_{i=1}^{N} b_i(T_i u, T_i u).$$

It implies that $b_i(T_i u, T_i u) = 0$ for all i. Therefore $T_i u = 0$ for all i. By equation (2.3), we have

$$a(u, v_i) = 0, \qquad \forall v_i \in V_i$$

for $i = 1, 2, \cdots, N$. But since $V = V_1 + V_2 + \cdots + V_N$, we finally obtain $a(u, u) = 0$. Then u must be equal to 0 and therefore T is invertible.

If we use equation (2.3) and the Cauchy-Schwarz inequality, we will have

$$\begin{aligned} a(T^{-1}u, u) &= \sum_i a(T^{-1}u, u_i) = \sum_i b_i(T_i T^{-1} u, u_i) \\ &\le (\sum_i b_i(T_i T^{-1} u, T_i T^{-1} u))^{1/2} (\sum_i b_i(u_i, u_i))^{1/2} \\ &= (\sum_i a(T^{-1}u, T_i T^{-1} u))^{1/2} (\sum_i b_i(u_i, u_i))^{1/2} \\ &= (a(T^{-1}u, u))^{1/2} (\sum_i b_i(u_i, u_i))^{1/2}. \end{aligned}$$

Therefore

$$a(T^{-1}u, u) \le \sum_i b_i(u_i, u_i).$$

Equality holds if and only if $u_i = T_i T^{-1} u$. □

As a direct consequence of this lemma, we get the following lemma.

Lemma 2.1.2 *Let $T = T_1 + T_2 + \cdots + T_N$. If a decomposition of u, $u = \sum_i u_i$, can be found such that*

$$\sum_i b_i(u_i, u_i) \leq C_1 a(u, u), \quad \forall u \in V^h,$$

then

$$\lambda_{\min}(T) \geq C_1^{-1}.$$

Conversely if for all representations of u, $u = \sum_i u_i$, we have

$$a(u, u) \leq C_2 \sum_i b_i(u_i, u_i), \quad \forall u \in V^h,$$

then

$$\lambda_{\max}(T) \leq C_2.$$

An estimate of γ from Bramble, Pasciak, Wang and Xu [16] is given in the following theorem.

Theorem 2.1.1

$$\|E_N u\|_a \leq \gamma \|u\|_a,$$

with

$$\gamma^2 = 1 - \max\left(\frac{(2-\omega)\lambda_{\min}(T)}{(N + \omega^2 N(N-1)/2)}, \frac{(2-\omega)\lambda_{\min}(T)}{2(1 + \omega^2 N(N-1)/2)}\right).$$

In the above theorem, $\lambda_{\min}(T)$ can be estimated by using Lemma 2.1.2.

To obtain a refined result, we make assumptions on the subspace V_i. Let $\mathcal{E}$ be a matrix such that

$$|a(u_i, u_j)| \leq \epsilon_{ij} a(u_i, u_i)^{1/2} a(u_j, u_j)^{1/2}, \quad \forall u_i \in V_i, \forall u_j \in V_j, i, j = 1, \cdots, N. \tag{2.4}$$

The inequalities (2.4) are called strengthened Cauchy-Schwarz inequalities . Let $\rho(\mathcal{E})$ be the spectral radius of $\mathcal{E}$, then we can obtain an improved result, cf. Xu [83].

Theorem 2.1.2 *Assume that (2.4) holds. Then*

$$\|E_N u\|_a \leq \gamma \|u\|_a,$$

with

$$\gamma^2 = 1 - \frac{(2-\omega)\lambda_{\min}(T)}{2(1 + \omega^2 \rho(\mathcal{E})^2)}.$$

The classical multiplicative Schwarz algorithm can be accelerated by different iteration methods. Since the polynomial $I-E_N$ does not contain any constant term, we can compute the right hand side g of

$$(I - E_N)u = g$$

without knowing the exact solution u. Then we can use GMRES method, which is a conjugate gradient type method of solving unsymmetric and positive definite linear systems, to compute the solution u of the above equation. For details of GMRES, we refer to Saad and Schultz [67]. Similarly, we can consider the symmetrized multiplicative Schwarz algorithm which corresponds to the linear system

$$(I - E_N^T E_N)u = g$$

and use the conjugate gradient method to accelerate it. However the number of fractional steps in each iteration is then doubled.

2.2 Additive Schwarz Methods

The main problem with the multiplicative Schwarz Methods is the sequential nature of the fractional steps of each iteration. If we use the symmetrized variants in order to apply the standard conjugate gradient method to accelerate the convergence, then the sequential nature will be even further pronounced. In order to remove the inherent sequential behavior of the fractional steps, Dryja and Widlund introduced the additive Schwarz method, see [29] and [31]. Independent work on additive Schwarz methods can also be found in Matsokin and Nepomnyaschikh [53] and Nepomnyaschikh [60]. In recent years, the method has been generalized in several directions, see e.g. Cai and Widlund [20], [21] and [23] for nonsymmetric and indefinite elliptic problems, Zhang [85] for biharmonic Dirichlet problems and Pavarino [65] for p-version finite element methods.

Let us use the same symbols as in the last section for convenience. Additive Schwarz methods can be viewed as iterative methods providing good preconditioners for the stiffness matrix of the original finite element problem. The basic idea of the method is to work with lowest degree polynomial T of the operators T_i's, namely

$$T = T_1 + \cdots T_N.$$

We call T the additive Schwarz operator. The additive Schwarz method for solving equation (2.1) is replaced by an alternative problem: Find $u \in V$ by solving

$$Tu = \sum_i T_i = g. \tag{2.5}$$

Because T is a polynomial in T_i, which does not contain any constant term, we can compute $g = Tu = \sum_i g_i = \sum_i T_i u$ without knowing the solution of (2.1) by solving

$$b_i(g_i, v) = a(u, v) = f(v), \quad \forall v \in V_i.$$

Once the right hand side g is known, we can use the conjugate gradient method to solve equation (2.5). Therefore the main issue is to estimate the condition number of T.

The basis for estimating $\lambda_{\min}(T)$ and $\lambda_{\max}(T)$ is Lemma 2.1.2. However, for domain decomposition algorithms, the following lemma gives a convenient way of estimating $\lambda_{\max}(T)$.

Lemma 2.2.1 *Consider the undirected graph with a node for each subspace V_i, and an edge between node i and node j if and only if $V_i \cap V_j \neq \{0\}$. Let N_c be the number of colors needed to color the graph so that no two nodes connected by an edge have the same color. Then*

$$\lambda_{\max}(T) \leq \omega \cdot N_c.$$

Proof. Because $\|T_i\|_a \leq \omega$, then

$$a(T_i v_i, v_i) \leq \omega a(v_i, v_i), \quad \forall v_i \in V_i.$$

Now we take an arbitrary u and write

$$u = P_i u + (I - P_i)u \equiv u_i + u_i^{\perp}.$$

Here P_i is the orthogonal projection operator of V_i. We note that

$$\begin{aligned} a(T_i u_i, u_i^{\perp}) &= 0, \\ a(T_i u_i^{\perp}, u_i) = a(u_i^{\perp}, T_i u_i) &= 0, \\ \text{and} \qquad a(T_i u_i^{\perp}, u_i^{\perp}) &= 0. \end{aligned}$$

Therefore

$$a(T_i u, u) = a(T_i u_i, u_i) \leq \omega a(u_i, u_i) = \omega a(P_i u, u).$$

This implies that

$$T_i \leq \omega P_i, \quad \forall i.$$

Therefore $T \leq \omega P$ where $P = \sum P_i$. All the subspaces of a particular color are disjoint; hence their corresponding orthogonal projection operators are mutually orthogonal. Therefore the sum of the orthogonal projection operators of a particular color is itself an orthogonal projection operator. P is then the sum of N_c orthogonal projection operators, which have norm one. □

Another way of estimating $\lambda_{\max}(T)$, as in the multiplicative case, is to use a strengthened Cauchy-Schwarz inequality. The proof of the following lemma is elementary.

Lemma 2.2.2 *Assume equation (2.4) holds, then*

$$\lambda_{\max}(T) \leq \omega \cdot \rho(\mathcal{E}).$$

From the above results, we see that the same parameters $\lambda_{\min}$, ω and $\rho(\mathcal{E})$ determine the bounds for the convergence rate of both the additive and multiplicative Schwarz methods. In the case of two subspaces, the relation of the convergence rates of the additive and multiplicative versions for the same subspaces is well understood, see Bjørstad [2] and Bjørstad and Mandel [3]. However little can be said in the case of many subspaces. For a numerical comparison of the two versions in the case of many subspaces using massively parallel computers, see e.g. Bjørstad and Skogen [5]. The architectures of the parallel computers available and the nature of the problems will probably determine which variant to prefer rather than a strict mathematical analysis.

2.3 Multigrid Methods

In this section, we describe the basic multigrid methods and point out that they can be regarded as multiplicative Schwarz methods. Finally we remark that the corresponding error propagation operators have spectral radii less than 1 which are independent of the number of level under quite general assumptions.

We consider the elliptic problem (2.1) on the finite-dimensional space V. Let us assume that we are given a nested sequence of subspaces of V

$$V_1 \subset V_2 \subset \cdots \subset V_N \equiv V.$$

For each i, we define Q_i, $P_i : V \to V_i$ and $A_i : V_i \to V_i$ by

$$(Q_i u, v_i)_{L^2(\Omega)} = (u, v_i)_{L^2(\Omega)}, \qquad a(P_i u, v_i) = a(u, v_i), \qquad u \in V, v_i \in V_i,$$

and

$$(A_i u_i, v_i) = (A u_i, v_i) = a(u_i, v_i), \qquad u_i, v_i \in V_i. \tag{2.6}$$

It follows from the definition that

$$A_i P_i = Q_i A. \tag{2.7}$$

We note that any simple iterative method for solving $Au = f$ is of the form:

$$u_{k+1} = u_k + B(f - A u_k), \qquad k = 0, 1, 2, \cdots. \tag{2.8}$$

The convergence rate of scheme (2.8) is determined by $\rho = \|I - BA\|_a$. We now describe a basic multigrid algorithm as follows.

Algorithm 2.3.1 *Let $B_1 = A_1^{-1}$ and assume that $B_{i-1} : V_{i-1} \to V_{i-1}$ has been defined; then for $g \in V_i$, $B_i : V_i \to V_i$ is defined by*

Step 1. $v^1 = B_{i-1} Q_{i-1} g$.

Step 2. $B_i g = v^1 + R_i(g - A_i v^1)$.

In the definition of this algorithm, we call step 1 the correction (to the coarse space) and step 2 the smoothing step. The operator R_i is called a smoother. This operator should have the property that $R_i(g - A_i v^1)$ has small components of high frequencies, which correspond to the eigenvectors of the large eigenvalues of A_i. The main idea of multigrid methods is to use the correction to damp out the low frequencies, which correspond to the eigenvectors of the smaller eigenvalues, and to use smoothing to damp out the high frequencies.

We remark that if we replace the L^2 inner product by an equivalent discrete inner product on the finite element space V and the L^2 projection Q_i by the projection $\tilde{Q}_i$ with

respect to the discrete inner product, the prolongation is just the inclusion $I_i : V_i \to V$ and the restriction is just the projection $\tilde{Q}_i : V \to V_i$ in classical terminology of multigrid methods.

It is straightforward to show that

$$I - B_i A_i P_i = (I - R_i A_i P_i)(I - B_{i-1} A_{i-1} P_{i-1}).$$

Since $P_N = I$, a successive application of the above identity yields the error propagation operator

$$E_N \equiv I - B_N A_N = (I - T_N)(I - T_{N-1}) \cdots (I - T_1), \tag{2.9}$$

where

$$T_1 = P_1, \quad T_i = R_i A_i P_i, \quad i = 2, 3, \cdots, N.$$

The above form of E_N shows that we can regard the basic multigrid methods as abstract multiplicative Schwarz methods. We can also symmetrize the above algorithm to obtain multigrid V-cycle algorithms.

Algorithm 2.3.2 *Let $B_1 = A_1^{-1}$ and assume that $B_{i-1} : V_{i-1} \to V_{i-1}$ has been defined; then for $g \in V_i$, $B_i : V_i \to V_i$ is defined by*

Step 1. $v^1 = R_i g$.

Step 2. $v^2 = v^1 + B_{i-1} Q_{i-1}(g - A_i v^1)$.

Step 3. $B_i g = R_i Q_i v^2$.

The above algorithm has the error propagation operator $E_N^T E_N$ where E_N^T is the adjoint of E_N with respect to $a(\cdot, \cdot)$. Therefore it is sufficient to analyze Algorithm 2.3.1. Suppose that V_i is the Lagrange finite element space with quasi-uniform mesh size h_i. We assume that h_i/h_{i+1} is uniformly bounded. Technically, the assumption on R_i is

$$C_1 \lambda_i^{-2} \|A_i v\|^2_{L^2(\Omega)} \leq \|R_i A_i v\|^2_{L^2(\Omega)} \leq \omega_1 \lambda_i^{-1} (A_i v, v)_{L^2(\Omega)}, \quad \forall v \in V_i. \tag{2.10}$$

Here λ_i is the spectral radius of the operator A_i. As was demonstrated in Xu [83], All damped Jacobi and symmetric Gauss-Seidel iterations satisfy the above assumption. The following theorem establishes the optimality of multigrid V-cycle algorithm. A proof can be found in Xu [83].

Theorem 2.3.1 *Assume that R_i's satisfy (2.10) with $\omega_1 < 2$; then*

$$\|E_N\|_a \leq \eta < 1$$

where η is a constant independent of the number of level N and the mesh size of coarsest triangulation.

For more convergence theory of multigrid methods, we refer to Bramble, Pasciak and Xu [15], Bramble and Pasciak [10], Zhang [86], and Hackbusch [41].

Chapter 3

Iterative Refinement Methods

3.1 Introduction

In this chapter, we consider the solution of the large linear systems of algebraic equations which arise when working with elliptic finite element approximations on composite meshes.

Finite element models on composite meshes can be constructed systematically in the framework of conforming finite elements as developed in Ciarlet [24]. We use conforming spaces in order to be able to use a number of technical tools which are then available. We begin by introducing a basic finite element approximation on the entire region and we then repeatedly select subregions, and subregions of subregions, where the finite element model is further refined. The resulting linear system is solved by a conjugate gradient method while solving only so-called standard problems. They correspond to the finite element models on the original region and those on the subregions, prior to further refinement. We can thus use solvers for problems with uniform or relatively uniform mesh sizes, while the composite mesh can be strongly graded.

This approach offers a number of advantages. Existing codes can be upgraded, in order to increase the accuracy locally, without a radical redesign of the data structures etc., since we can use the old programs to solve one or several of the standard problems. Issues of data structures, geometry and vectorization are generally easier if we design programs for composite mesh problems in terms of such standard problems. We also note that if each of the standard problems has appreciatively fewer degrees of freedom than the composite model, then we might benefit from solving a number of smaller problems rather than one large one. In other words, this approach can also be viewed as a divide-and-conquer

strategy.

The systematic study of iterative refinement methods goes back at least to 1983, when work on the Fast Adaptive Composite (FAC) method was begun by Thomas and McCormick; cf. [54]. Issues related to the implementation of this method on parallel computers led to the introduction of Asynchronous FAC (AFAC) methods a few years later; cf. Hart and McCormick [42]. In our terminology, these are multiplicative and additive algorithms, respectively. Numerical experiments have been reported for model cases and Ewing, McCormick and others have also tested the algorithms for more difficult problems arising in industry. The convergence of the FAC method is established in McCormick et al. [54], [55] under a certain additional regularity assumption. An important contribution to the theory is given by Bramble, Ewing, Pasciak, and Schatz [9], for a two level FAC algorithm. Their proof of optimality is regularity free. In Bjørstad [2] and Mandel and McCormick [49], a theory for two-level multiplicative and additive algorithms was developed. In particular, they obtained interesting results concerning the relationship between the spectra of the two methods. In another paper, Mandel and McCormick [48] established optimality for a multi-level AFAC algorithm for a special model problem. Then Ewing, Lazarov and Vassilevski [36] obtained interesting bounds for finite volume schemes for the multi-level multiplicative algorithm corresponding to FAC. Bjørstad, Moe and Skogen [4], and Moe [59] have also carried out experiments that illustrate the convergence behavior and they have also described parallel implementations on both shared memory computers and on local memory systems. The use of FAC and AFAC methods for mixed finite element problems has been studied by Mathew in [50] [51]. The results of this chapter generalize the work by Dryja and Widlund and provide a correct proof of one of the results given in [81].

3.2 FAC and AFAC Formulations

We consider linear, self adjoint, elliptic problems (1.1)-(1.3) discretized by finite element methods on a bounded Lipschitz polyhedral region Ω in R^n. To simplify the presentation, we use continuous, Lagrange finite element of type p, p being fixed. However, almost all of our results can be extended immediately to general conforming finite element approx-

imations of any self adjoint elliptic problem which can be formulated as a minimization problem. The continuous and discrete problems are given in (1.4) and (1.6) on the spaces V and V^h respectively. The space V and V^h are defined in the next few paragraphs.

To simplify the presentation, we assume that we have a homogeneous Dirichlet boundary condition on $\partial\Omega$, the boundary of Ω. The space V is thus $H_0^1(\Omega)$. We note that we equally well could have considered an inhomogeneous Dirichlet problem. Later we will also make some remarks about the case of mixed type boundary value condition. The space V^h is defined on a composite triangulation, which is possibly the result of a large number of successive refinements. The triangulation of Ω is given in the following way.

We first introduce a relatively coarse triangulation of Ω, also denoted by Ω_1, and denote the corresponding space of finite element functions by V^{h_1}. We can think of this space as having a relatively uniform mesh size h_1. Let Ω_2 be a subregion where we wish to increase the resolution. We do so by subdividing the elements and introducing an additional finite element space V^{h_2}. We assure that the resulting composite space $V^{h_1} + V^{h_2}$ is conforming by having the functions of V^{h_2} vanish on $\partial\Omega_2$. We repeat this process by selecting a subregion Ω_3 of Ω_2 and introducing a further refinement of the mesh and finite element space, etc.. We denote the resulting nested subregions and subspaces by Ω_i and V^{h_i} respectively. Throughout, we have $\Omega_i \subset \Omega_{i-1}$ and $V^{h_i-1} \cap H_0^1(\Omega_i) \subset V^{h_i} \subset H_0^1(\Omega_i)$, $i = 2, \cdots, k$. The composite finite element space on the repeatedly refined mesh, is

$$V^h = V^{h_1} + V^{h_2} + \cdots + V^{h_k}.$$

We assume that all the elements are shape regular in the sense that there is a uniform bound on h_K/ρ_K. Here h_K and ρ_K are the diameter and the radius of the largest inscribed sphere of any element K, respectively. Our theoretical bounds, developed in this chapter, also depend on the shapes of the subregions Ω_i.

The finite element problem is defined by equation (1.6) and the corresponding stiffness matrix can conveniently be computed by using a process of subassembly. Introducing subscripts to indicate the domain of integration, we write

$$a(u, v) = a_{\Omega_1 \setminus \Omega_2}(u, v) + a_{\Omega_2 \setminus \Omega_3}(u, v) + \cdots + a_{\Omega_k}(u, v).$$

The stiffness matrices corresponding to the regions $\Omega_i \setminus \Omega_{i+1}, i \leq k-1$, and Ω_k are computed by working with basis functions related to the mesh size h_i. The quadratic

form corresponding to the composite stiffness matrix is the sum of the quadratic forms corresponding to $\Omega_i \setminus \Omega_{i+1}, i \leq k-1$, and Ω_k. When we refine a finite element model locally, the modified stiffness matrix is obtained by replacing the quadratic form associated with the subregion in question by the one corresponding to the refined model on the same subregion. It is therefore relatively easy to design a method which systematically generates the stiffness matrices for all the standard problems necessary while, at the same time, the stiffness matrix of the composite model is computed.

We now can describe FAC as follows.

Algorithm 3.2.1 (FAC) *Let $u_h^n \in V^h$ be the nth approximation to the solution of (1.6). Then an iteration step of computing u_h^{n+1} consists of k fractional steps: compute $w_i \in V^{h_i}$ from*

$$a(u_h^{n+(i-1)/k} + w_i, v_i) = f(v_i), \qquad \forall v_i \in V^{h_i} \tag{3.1}$$

and set

$$u_h^{n+i/k} = u_h^{n+(i-1)/k} + w_i \qquad \text{for } i = 1, 2, \cdots, k.$$

We remark that the above w_i can be regarded as the solution of a inhomogeneous Dirichlet boundary value problem in Ω_i with the boundary data $u_h^{n+(i-1)/k}$ on $\partial\Omega_i$.

Let us define P_j^i, $i \leq j$, as the projections onto the spaces $V^{h_i} \cap H_0^1(\Omega_j)$. We note that if $j > i$, then we solve a problem on Ω_j with a coarser mesh than if V^{h_j} were used. The projection $P_j^i, i \leq j$, is defined, in terms of the unique element of $V^{h_i} \cap H_0^1(\Omega_j)$, which satisfies

$$a(P_j^i v_h, \phi_h) = a(v_h, \phi_h), \qquad \forall\ \phi_h \in V^{h_i} \cap H_0^1(\Omega_j). \tag{3.2}$$

In each fractional step of FAC, we essentially find a correction $w_i \in V^{h_i}$ of the current approximation u_h^n by solving

$$a(w_i, v_i) = f(v_i) - a(u_h^n + \sum_{j=1}^{i-1} w_j, v_i) = a(u_h^* - (u_h^n + \sum_{j=1}^{i-1} w_j), v_i), \qquad \forall v_i \in V^{h_i},$$

where $u_h^* \in V^h$ is the solution of the discrete problem (1.6). Therefore (3.2) gives us

$$w_i = P_i^i(u_h^* - (u_h^n + \sum_{j=1}^{i-1} w_j)). \tag{3.3}$$

Equation (3.3) shows that, in the case of k levels, one iteration of FAC yields the following partial solutions

$$\begin{aligned}
w_1 &= P_1^1(u_h^* - u_h^n), \\
w_2 &= P_2^2(u_h^* - u_h^n - w_1) = P_2^2(I - P_1^1)(u_h^* - u_h^n), \\
w_3 &= P_3^3(u_h^* - u_h^n - w_1 - w_2) = P_3^3(I - P_2^2)(I - P_1^1)(u_h^* - u_h^n), \\
&\vdots \\
w_k &= P_k^k(u_h^* - u_h^n - \sum_{j=1}^{k-1} w_j) = P_k^k(I - P_{k-1}^{k-1}) \cdots (I - P_1^1)(u_h^* - u_h^n).
\end{aligned}$$

Based on the above equations, we observe that the error $e_h^n = u_h^n - u_h^*$ propagates as

$$\begin{aligned}
e_h^{n+1} &= u_h^{n+1} - u_h^* \\
&= u_h^n - u_h^* + \sum_{j=1}^{k} w_j \\
&= u_h^n - u_h^* - \sum_{j=2}^{k} P_j^j(I - P_{j-1}^{j-1}) \cdots (I - P_1^1)(u_h^n - u_h^*) - P_1^1(u_h^n - u_h^*) \\
&= (I - P_1^1)(I - P_2^2) \cdots (I - P_k^k) e_h^n.
\end{aligned}$$

Therefore we can regard FAC as a multiplicative Schwarz method based on the subspaces V^{h_i}, for $i = 1, 2, \cdots, k$.

The FAC multiplicative operator is a nonsymmetric operator that cannot be accelerated by standard conjugate gradient methods. In order to get a symmetric operators, we have to use additional steps. We can formulate the symmetrized FAC in the following algorithm.

Algorithm 3.2.2 (Symmetrized FAC) *Let $u_h^n \in V^h$ be the nth approximation to the solution of (1.6). Then one iteration step can be expressed as computing u_h^{n+1} in $2k$ fractional steps: first compute $w_i \in V^{h_i}$ from*

$$a(u_h^{n+(i-1)/2k} + w_i, v_i) = f(v_i), \quad \forall v_i \in V^{h_i}$$

and set

$$u_h^{n+i/2k} = u_h^{n+(i-1)/2k} + w_i \qquad \text{for } i = 1, 2, \cdots, k.$$

Next compute $w'_{k+1-i} \in V^{h_{k+1-i}}$ *from*

$$a(u_h^{n+1/2+(i-1)/2k} + w'_{k+1-i}, v_{k+1-i}) = f(v_{k+1-i}), \quad \forall v_{k+1-i} \in V^{h_{k+1-i}}$$

and set

$$u_h^{n+1/2+i/2k} = u_h^{n+1/2+(i-1)/2k} + w'_{k+1-i} \qquad \text{for } i = 1, 2, \cdots, k.$$

In order to adapt the FAC method to multiprocessor systems, several different asynchronous FAC algorithms have been introduced; cf. [42] and [81]. These variants of FAC belong to the additive Schwarz methods which is more easily decomposed into independent processes, at the cost of somewhat slower convergence.

Algorithm 3.2.3 (AFAC1) *Apply an iterative method, e.g. the conjugate gradient method, to the symmetric and positive definite system*

$$P_a^{(1)} u_h \equiv (P_1^1 + P_2^2 + \cdots + P_k^k) u_h = g_h$$

for an appropriate g_h *such that the solution* u_h *is the same as that of (1.6).*

The condition number of the AFAC1 operator grows linearly with k as shown by Widlund in [81]. The eigenvalues of $P_a^{(1)}$ are always bounded from above by k. This bound is attained if $V_{h_1} \cap H_0^1(\Omega_k)$ is not empty, i.e. the coarsest mesh size h_1 is fine enough.

Improvements of the basic AFAC method were described by Mandel and McCormick in [49], [48] and by Widlund [81]. However, the original idea was known even earlier. In terms of orthogonal projections, the operator can be represented as

$$P_a^{(2)} = P_1^1 + (P_2^2 - P_2^1) + \cdots + (P_k^k - P_k^{k-1}).$$

All the terms, except the first, are similar to multigrid corrections since they each represent the difference between two solutions on the same subregion, using two different mesh sizes.

Algorithm 3.2.4 (AFAC2) *Apply an iterative method, e.g. the conjugate gradient method, to the symmetric and positive definite system*

$$P_a^{(2)} u_h \equiv (P_1^1 + (P_2^2 - P_2^1) + \cdots + (P_k^k - P_k^{k-1})) u_h = g_h$$

for an appropriate g_h *such that the solution* u_h *is the same as that of (1.6).*

The modified version of AFAC2 developed by Widlund, which we call AFAC3, is given next. This was the first additive algorithm for which an optimal lower bound was established; cf. Widlund [81].

Algorithm 3.2.5 (AFAC3) *Apply an iterative method, e.g. the conjugate gradient method, to the symmetric and positive definite system*

$$P_a^{(3)} u_h \equiv (P_1^1 + P_2^2 + (P_3^3 - P_3^1) \cdots + (P_k^k - P_k^{k-2}))u_h = g_h$$

for an appropriate g_h such that the solution u_h is the same as that of (1.6).

We end this section by discussing several variants of these algorithms. We can use other energy-symmetric operators T_i^i to replace the projections P_i^i. Practically we should choose the T_i^i such that the $T_i^i u_h$ are much easier to compute than the $P_i^i u_h$ and T_i^i is spectrally equivalent to P_i^i. The corresponding operator for symmetrized FAC with k subspaces involves $2k$ fractional steps and has the form

$$I - E_k E_k^*, \quad \text{where} \quad E_k = (I - T_1^1)(I - T_2^2) \cdots (I - T_k^k). \tag{3.4}$$

The error propagation operators for the basic FAC algorithm is E_k^*, while the convergence rate of the symmetrized FAC algorithm can be bounded in terms of the condition number of $I - E_k E_k^*$. We can also consider a variation, which we call the modified FAC, with an iteration operator

$$(I - E_k) + (I - E_k)^*, \tag{3.5}$$

cf. Dryja, Smith, and Widlund [28]. We observe that

$$I - E_k E_k^* \leq (I - E_k) + (I - E_k^*) \leq (2 + 2\|E_k\|_a)I.$$

Therefore the optimality of the operator (3.5) follows from that of the operator (3.4). The benefit of using this modified algorithm is that we can use two, or two groups of, processors of a parallel computer to do the computation. However, we will see later that in the cases we consider the modified algorithm has a larger condition number than the standard symmetrized multiplicative algorithm.

3.3 Optimal Convergence Rate Estimates for FAC and AFAC

In this section, we will prove optimal bounds for FAC and AFAC including the case of inexact solvers by using an extension theorem for finite element functions and an interpolation theorem of Hilbert scales. We have obtained an estimate of the condition number for AFAC1 in the last section. We observe that

$$P_a^{(3)} = P_a^{(2)} + P_2^1 + (P_3^2 - P_3^1) + \cdots + (P_k^{k-1} - P_k^{k-2}) \equiv P_a^{(2)} + \tilde{P}_a^{(2)}.$$

Because $\tilde{P}_a^{(2)}$ is a positive semi-definite operator on V^h, $P_a^{(3)}$ will have an optimal lower bound if $P_a^{(2)}$ has an optimal lower bound. The forms of $P_a^{(2)}$ and $\tilde{P}_a^{(2)}$ are quite similar. We can also apply our estimate of the upper bound of $P_a^{(2)}$ to the case of $P_a^{(3)}$. Therefore it is sufficient to analyze the convergence rate for AFAC2.

In order for our proofs to work, we often make the following assumptions.

Assumption 3.3.1 *For each j, there exists a bounded Lipschitz polyhedral region $\tilde{\Omega}_j$ such that $\Omega_j \subset \tilde{\Omega}_j$, $(\tilde{\Omega}_j \setminus \Omega_j) \cap \Omega = \emptyset$, $\partial\tilde{\Omega}_j \cap \partial\Omega_{j+1} = \emptyset$ and the Lipschitz constants of $\tilde{\Omega}_j \setminus \Omega_{j+1}$ are uniformly bounded.*

Assumption 3.3.2 *The mesh sizes h_i are bounded from above and below by const.q^i uniformly for all i. Here q is some constant less than 1.*

Assumption 3.3.1 makes sure that the following type of extension theorem holds: Given $u_h \in V^{h_1} + V^{h_2} + \cdots + V^{h_j}$, there exists $\tilde{u}_h \in V^{h_1} + V^{h_2} + \cdots + V^{h_j}$ which satisfies $\tilde{u}_h = u_h$ except in the subregion Ω_{j+1} and

$$a_{\Omega_{j+1}}(\tilde{u}_h, \tilde{u}_h) \leq C a_{\Omega_j \setminus \Omega_{j+1}}(u_h, u_h)$$

where the constant C only depends on the Lipschitz constant of $\tilde{\Omega}_j \setminus \Omega_{j+1}$ and the shape regularity. In the case of $\Omega_j \subset\subset \Omega$, we can only choose Ω_{j+1} such that $\Omega_{j+1} \subset\subset \Omega_j$ and in this case the only choice of $\tilde{\Omega}_{j+1}$ is Ω_{j+1}. To avoid arbitrarily large C, we therefore cannot allow the sets $\tilde{\Omega}_j \setminus \Omega_{j+1}$ to have arbitrarily large Lipschitz constants. However, our proof can be extended to more general cases with only a slight modification. We can replace Assumption 3.3.1 by the following assumption.

Assumption 3.3.3 *For each j, there exists a bounded Lipschitz polyhedral region $\tilde{\Omega}_{N(j)}$, with $1 \leq N(j) \leq j$ and $j - N(j)$ uniformly bounded with respect to j, such that $\Omega_{N(j)} \subset \tilde{\Omega}_{N(j)}$, $(\tilde{\Omega}_{N(j)} \setminus \Omega_{N(j)}) \cap \Omega = \emptyset$, $\partial\tilde{\Omega}_{N(j)} \cap \partial\Omega_{j+1} = \emptyset$ and the Lipschitz constants of $\tilde{\Omega}_{N(j)} \setminus \Omega_{j+1}$ are uniformly bounded.*

Using this assumption, we can change the extension theorem given above to the following:

$$a_{\Omega_{j+1}}(\tilde{u}_h, \tilde{u}_h) \leq C a_{\Omega_{N(j)} \setminus \Omega_{j+1}}(u_h, u_h)$$

where u_h and $\tilde{u}_h$ are the same as above and the constant C only depends on the Lipschitz constant of $\tilde{\Omega}_{N(j)} \setminus \Omega_{j+1}$ and the shape regularity.

3.3.1 An Extension Theorem for Finite Element Functions

In this subsection, we discuss an extension theorem for finite element functions given in Widlund [78]. For completeness, we also include a proof.

In the following lemma, we assume that we are given a triangulation on Ω with quasi-uniform mesh size h and that $R^n \setminus \Omega$ has been triangulated in an equally benign way as Ω. We assume that we are dealing with Lagrange type finite element spaces.

Lemma 3.3.1 (Widlund) *Let Ω be an open, bounded, and Lipschitz polyhedral region in R^n, $n \geq 2$. There exists a constant $C(\Omega)$, which depends only on the Lipschitz constant of $\partial\Omega$ and the shape regularity of the triangulation such that for all $u_h \in V^h(\Omega)$ we can extend u_h to $\tilde{u}_h \in V^h(R^n)$ and $\tilde{u}_h$ satisfies*

$$\|\tilde{u}_h\|_{H^1(R^n)} \leq C(\Omega)\|u_h\|_{H^1(\Omega)}.$$

The result is also true if we only extend u_h to any domain $\tilde{\Omega}$ which contains Ω and do not introduce any extra constraint on the extended function $\tilde{u}_h$.

Proof. Let $\{\phi_i\}$ be the usual nodal basis functions and let $\triangle_i$ denote the support of ϕ_i. It is easy to see that each basis function is uniform of order 1, i.e. there exist constants c_0 and c_1, depending only on the shape regularity, such that

$$\max_{x \in \triangle_i} |\phi_i(x)| \leq c_0 \tag{3.6}$$

and

$$\max_{x\in\triangle_i}\sum_{j=1}^{n}|\frac{\partial\phi_i(x)}{\partial x_j}|\le c_1h^{-1}. \tag{3.7}$$

We are now ready to describe how to find an extension $\mathcal{E}^hu_h\in V^h(R^n)$ of a given element $u_h\in V^h(\Omega)$ such that

$$\|\mathcal{E}^hu_h\|_{H^1(R^n)}\le C(\Omega)\|u_h\|_{H^1(\Omega)}.$$

We first use Stein's extension theorem in Lemma 1.2.6 and the fact the finite element space is conforming to find $\mathcal{E}u_h\in H^1(R^n)$ such that

$$\|\mathcal{E}u_h\|_{H^1(R^n)}\le C(\Omega)\|u_h\|_{H^1(\Omega)}. \tag{3.8}$$

Here $C(\Omega)$ depends only on the Lipschitz constant of Ω. However $\mathcal{E}u_h$ does not belong to $V^h(R^n)$.

Next we follow a procedure in Strang [76] and use a special mollifier, which preserves all polynomials of degree less than or equal to p, to smooth $\mathcal{E}u_h$ and interpolate the smooth function to get an element $w_h\in V^h(R^n)$ such that

$$\|w_h\|_{H^1(R^n)}\le C\|\mathcal{E}u_h\|_{H^1(R^n)}\le C(\Omega)C\|u_h\|_{H^1(\Omega)} \tag{3.9}$$

and

$$\|w_h-\mathcal{E}u_h\|^2_{L^2(K)}\le Ch^2\sum_{K'}\|w_h-\mathcal{E}u_h\|^2_{H^1(K')}. \tag{3.10}$$

Here the sum is over a fixed number N_0 of neighboring elements. This is possible because of shape regularity. The constants C in (3.9) and (3.10) depend only on the shape regularity.

Finally we patch w_h and u_h together to create $\mathcal{E}^hu_h$. We do the interpolation by using all the original parameters obtained from u_h in $\overline{\Omega}$ and those associated with w_h in $R^n\setminus\Omega$. For any element $K\subset\Omega$ the original values of u_h are recovered. In $R^n\setminus\Omega$, $\mathcal{E}^hu_h$ differs from w_h only on elements which have at least one vertex on $\partial\Omega$.

By the triangle inequality

$$\|\mathcal{E}^hu_h\|_{H^1(R^n\setminus\Omega)}\le\|\mathcal{E}^hu_h-w_h\|_{H^1(R^n\setminus\Omega)}+\|w_h\|_{H^1(R^n\setminus\Omega)}.$$

By (3.9), $\|w_h\|_{H^1(R^n\setminus\Omega)}$ can be estimated by $\|u_h\|_{H^1(\Omega)}$.

Consider an element $K' \subset R^n \setminus \Omega$ with at least one vertex on $\partial\Omega$ such that $\mathcal{E}^h u_h - w_h$ is not identically zero on K'. By using (3.6) and (3.7), we can establish the following inverse inequality :

$$\|\mathcal{E}^h u_h - w_h\|^2_{H^1(K')} \leq ch^{-2}\|\mathcal{E}^h u_h - w_h\|^2_{L^2(K')}. \tag{3.11}$$

Under the shape regularity assumptions, we can replace the L^2 norm by an equivalent discrete norm on the space $V^h(R^n)$. When considering the discrete norm of $\mathcal{E}^h u_h - w_h$ on a neighboring element $K \subset \Omega$ which shares the nodes of K' which lie on $\partial\Omega$, we see that

$$\|\mathcal{E}^h u_h - w_h\|_{L^2(K')} \leq C\|\mathcal{E}^h u_h - w_h\|_{L^2(K)} = C\|\mathcal{E} u_h - w_h\|_{L^2(K)}. \tag{3.12}$$

If we combine (3.10), (3.11) and (3.12), the proof of the lemma will follow easily. □.

3.3.2 Convergence Rate Estimate for AFAC

In this subsection, we give a new proof of uniform lower bound for AFAC2. The old proof is given in Dryja and Widlund [30] using Lemma 2.1.2 and a variation of the extension theorem mentioned in last subsection. We use two different approaches to establish the upper bound under Assumption 3.3.1 or Assumption 3.3.2. We first need use the following theorem which is stated and proved in Widlund [81].

Theorem 3.3.1 *Under Assumption 3.3.1, the symmetrized FAC algorithm based on the projections P_i^i, $i \leq k$, has a condition number which is independent of k and the number of degrees of freedom. The spectral radius of the basic FAC algorithm is bounded by a constant which is uniformly less than one.*

We note that the proof makes use of a variation of the extension theorem mentioned in last subsection, which we will state later, and therefore no additional assumptions on the mesh sizes h_i are needed.

Now let us give a new proof of the following theorem.

Theorem 3.3.2 *Under Assumption 3.3.1, the eigenvalues of $P_a^{(2)}$ are bounded from below by a constant which is independent of k and the number of degrees of freedom.*

Proof. We first observe that

$$(I - P_k^k) \cdots (I - P_2^2)(I - P_1^1) = (I - P_k^k + P_k^{k-1}) \cdots (I - P_2^2 + P_2^1)(I - P_1^1)$$

since $P_i^{i-1}(I - P_{i-1}^{i-1}) = 0$ for $2 \leq i \leq k$.

Let $S_i = (I - P_i^i + P_i^{i-1}) \cdots (I - P_2^2 + P_2^1)(I - P_1^1)$, $1 \leq i \leq k$, and set $S_k = S$. Using Theorem 3.3.1 and our observation, we can conclude that there exists a constant η, uniformly less than 1, such that $\|Su\|_a \leq \eta \|u\|_a$. Next if we define $P_1^0 = 0$ for convenience and use the identity

$$u = \sum_{i=1}^{k} (P_i^i - P_i^{i-1}) S_{i-1} u + Su$$

and the Cauchy-Schwarz inequality, we will obtain

$$\begin{aligned}
\|u\|_a^2 = a(u,u) &= a(u, Su + \sum_{i=1}^{k} (P_i^i - P_i^{i-1}) S_{i-1} u) \\
&\leq \|u\|_a \|Su\|_a + \sum_{i=1}^{k} a((P_i^i - P_i^{i-1})u, u)^{1/2} a((P_i^i - P_i^{i-1}) S_{i-1} u, S_{i-1} u)^{1/2} \\
&\leq \|u\|_a \|Su\|_a + a(P_a^{(2)} u, u)^{1/2} [\sum_{i=1}^{k} a((P_i^i - P_i^{i-1}) S_{i-1} u, S_{i-1} u)]^{1/2} \\
&= \|u\|_a \|Su\|_a + a(P_a^{(2)} u, u)^{1/2} [\|u\|_a^2 - \|Su\|_a^2]^{1/2}.
\end{aligned}$$

The last step follows from the observation that

$$\begin{aligned}
&(P_1^1)^* P_1^1 + ((P_2^2 - P_2^1) S_1)^* ((P_2^2 - P_2^1) S_1) + \cdots \\
&+ ((P_k^k - P_k^{k-1}) S_{k-1})^* ((P_k^k - P_k^{k-1}) S_{k-1}) + S_k^* S_k \\
&= (P_1^1)^* P_1^1 + S_1^* (P_2^2 - P_2^1) S_1 + \cdots + S_{k-1}^* (P_k^k - P_k^{k-1}) S_{k-1} + S_k^* S_k \\
&= (P_1^1)^* P_1^1 + S_1^* (P_2^2 - P_2^1) S_1 + \cdots + S_{k-1}^* (P_k^k - P_k^{k-1}) S_{k-1} + S_{k-1}^* (I - P_k^k + P_k^{k-1}) S_{k-1} \\
&= (P_1^1)^* P_1^1 + S_1^* (P_2^2 - P_2^1) S_1 + \cdots S_{k-2}^* (P_{k-1}^{k-1} - P_{k-1}^{k-2}) S_{k-2} + S_{k-1}^* S_{k-1} \\
&= \cdots = (P_1^1)^* P_1^1 + S_1^* S_1 = P_1^1 + I - P_1^1 = I.
\end{aligned}$$

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

$$1 \leq \frac{\|Su\|_a}{\|u\|_a} + [\frac{a(P_a^{(2)} u, u)}{a(u,u)}]^{1/2} [1 - \frac{\|Su\|_a^2}{\|u\|_a^2}]^{1/2}.$$

A simple calculation gives

$$1 - \|S\|_a \leq 1 - \frac{\|Su\|_a}{\|u\|_a} \leq \frac{a(P_a^{(2)} u, u)}{a(u,u)},$$

which shows that

$$\lambda_{\min}(P_a^{(2)}) \geq 1 - \|S\|_a \geq 1 - \eta. \quad \square$$

Next we will prove the following theorem using an extension theorem only.

Theorem 3.3.3 *Under Assumption 3.3.1, the operator $P_a^{(2)}$ has a condition number which is bounded by const.$\sqrt{k}$, but is independent of the number of degrees of freedom.*

In order to prove this theorem, we need the following lemma, which is a consequence of the extension theorem given in last subsection. Before stating the lemma, we define a family of operators, H_{j+1}^j, by

$$\begin{aligned} H_{j+1}^j v_h(x) &\in V^{h_1} + V^{h_2} + \cdots + V^{h_j}, \\ H_{j+1}^j v_h(x) &= v_h(x), \quad x \in \Omega_j \setminus \Omega_{j+1}, \\ a(H_{j+1}^j v_h, w_h) &= 0, \quad \forall w_h \in V^{h_j} \cap H_0^1(\Omega_{j+1}). \end{aligned}$$

It is natural to call $H_{j+1}^j v_h$ the h_j-harmonic extension of v_h to Ω_{j+1}, since it is the solution in V^{h_j} of a discrete Dirichlet problem with zero right hand side and with boundary data on $\partial\Omega_{j+1}$ given by v_h.

Lemma 3.3.2 *There exists a constant $C(\tilde{\Omega}_j \backslash \Omega_{j+1})$, which only depends on the Lipschitz constant of $\tilde{\Omega}_j \backslash \Omega_{j+1}$, where $\tilde{\Omega}_j$ is given in Assumption 3.3.1, such that*

$$a_{\Omega_{j+1}}(H_{j+1}^j u_h, H_{j+1}^j u_h) \leq C(\tilde{\Omega}_j \backslash \Omega_{j+1}) a_{\Omega_j \backslash \Omega_{j+1}}(u_h, u_h), \quad \forall u_h \in V^h.$$

Proof. To prove this Lemma, we need to use Lemma 3.3.1 and Poincaré's inequality, cf. Lemma 1.2.2. Using the condition $(\tilde{\Omega}_j \backslash \Omega_j) \cap \Omega = \emptyset$, we can make a constant extension of any function in V^h, which has a constant value on $\partial\Omega$, to the larger region $\Omega \cup \tilde{\Omega}_j$. The extended function is still a composite finite element function. We now apply Lemma 3.3.1 to the region $\tilde{\Omega}_j \backslash \Omega_{j+1}$. Given any function $u_h \in V^h$, we now define $c = \{u_h\}_{\tilde{\Omega}_j \backslash \Omega_{j+1}}$ as in Lemma 1.2.2. Let $\tilde{u}_{h_j}$ be the finite element function extension of $u_h - c$ from $\tilde{\Omega}_j \backslash \Omega_{j+1}$ to $\tilde{\Omega}_j$, as in Lemma 3.3.1, such that $\tilde{u}_{h_j}$ is in $V^{h_j}(\Omega_j)$. Since $\partial\tilde{\Omega}_j \cap \partial\Omega_{j+1} = \emptyset$, we can

conclude that $H^j_{j+1}(u_h - c)$ and $\tilde{u}_{h_j}$ have the same boundary values on $\partial\Omega_{j+1}$. However $H^j_{j+1}(u_h - c)$ is the minimal energy extension of $u_h - c$ from $\tilde{\Omega}_j \backslash \Omega_{j+1}$ to $\tilde{\Omega}_j$. Therefore

$$\begin{aligned}
&a_{\Omega_{j+1}}(H^j_{j+1}u_h, H^j_{j+1}u_h) = a_{\Omega_{j+1}}(H^j_{j+1}(u_h - c), H^j_{j+1}(u_h - c)) \\
\leq\ & a_{\Omega_{j+1}}(\tilde{u}_{h_j}, \tilde{u}_{h_j}) \leq C\|\tilde{u}_{h_j}\|^2_{\Omega_{j+1}} \leq CC_1^2(\tilde{\Omega}_j \backslash \Omega_{j+1})\|u_h - c\|^2_{H^1(\tilde{\Omega}_j \backslash \Omega_{j+1})} \\
&\leq CC_1^2(\tilde{\Omega}_j \backslash \Omega_{j+1})(1 + C_2^2(\tilde{\Omega}_j \backslash \Omega_{j+1})H^2_{\tilde{\Omega}_j \backslash \Omega_{j+1}})a_{\tilde{\Omega}_j \backslash \Omega_{j+1}}(u_h, u_h) \\
&= CC_1^2(\tilde{\Omega}_j \backslash \Omega_{j+1})(1 + C_2^2(\tilde{\Omega}_j \backslash \Omega_{j+1})H^2_{\tilde{\Omega}_j \backslash \Omega_{j+1}})a_{\Omega_j \backslash \Omega_{j+1}}(u_h, u_h)
\end{aligned}$$

by using Lemma 3.3.1 and Lemma 1.2.2. Here C_1 and C_2 are the constants which appear in Lemma 3.3.1 and Lemma 1.2.2.

Finally we can use a simple dilation argument to remove the dependence of the constant upon $H_{\tilde{\Omega}_j \backslash \Omega_{j+1}}$. □

Proof of Theorem 3.3.3: Let $u_i = (P^i_i - P^{i-1}_i)u_h$ for $i = 1, 2, \cdots, k$. Then $P^{(2)}u_h = u_1 + u_2 + \cdots + u_k$. Also note that if $j > i$, then $a(u_j, u_i) = a(u_j, u_i + v)$ for all $v \in V^{h_{j-1}} \cap H^1_0(\Omega_j)$. Therefore we can choose v such that

$$a(u_j, u_i + v) = a(u_j, H^{j-1}_j u_i) = a(u_j, u_i).$$

Hence if we use this equation, Lemma 3.3.2 and Cauchy-Schwarz inequality, we get

$$\begin{aligned}
a(P^{(2)}u_h, P^{(2)}u_h) &= \sum_{i=1}^k a(u_i, u_i) + 2\sum_{j>i} a(u_i, u_j) = \sum_{i=1}^k a(u_i, u_i) + 2\sum_{j>i} a(u_j, H^{j-1}_j u_i) \\
&\leq \sum_{i=1}^k a(u_i, u_i) + 2\sum_{i=1}^k \sum_{j=i+1}^k a(u_j, u_j)^{1/2} \cdot a_{\Omega_j}(H^{j-1}_j u_i, H^{j-1}_j u_i)^{1/2} \\
&\leq \sum_{i=1}^k a(u_i, u_i) + C\sum_{i=1}^k \sum_{j=i+1}^k a(u_j, u_j)^{1/2} \cdot a_{\Omega_{j-1} \backslash \Omega_j}(u_i, u_i)^{1/2} \\
&\leq \sum_{i=1}^k a(u_i, u_i) + C\sum_{i=1}^k [\sum_{j=i+1}^k a(u_j, u_j)]^{1/2} \cdot [\sum_{j=i+1}^k a_{\Omega_{j-1} \backslash \Omega_j}(u_i, u_i)]^{1/2} \\
&= \sum_{i=1}^k a(u_i, u_i) + C\sum_{i=1}^k [\sum_{j=i+1}^k a(u_j, u_j)]^{1/2} \cdot a(u_i, u_i)^{1/2} \\
&\leq \sum_{i=1}^k a(u_i, u_i) + C[\sum_{i=1}^k \sum_{j=i+1}^k a(u_j, u_j)]^{1/2} \cdot [\sum_{i=1}^k a(u_i, u_i)]^{1/2} \\
&\leq \sum_{i=1}^k a(u_i, u_i)(1 + C\sqrt{k}) \leq C'\sqrt{k}a(P^{(2)}u_h, u_h).
\end{aligned}$$

From this result and Theorem 3.3.2, Theorem 3.3.3 follows easily. □

Next we will prove that the above additive Schwarz operator has an optimal condition number under the additional Assumption 3.3.2.

We will prove some technical lemmas to establish the uniform upper bound of the operator $P_a^{(2)}$. We consider the Dirichlet boundary value problem

$$-\sum_i \sum_j \frac{\partial}{\partial x_i} a_{ij}(x) \frac{\partial U}{\partial x_j} = F \quad \text{in } \Omega, \quad U = 0 \quad \text{on } \partial\Omega$$

in a weak sense, i.e. $U \in H_0^1(\Omega)$ is the solution of

$$a(U, \chi) = (F, \chi)_{L^2(\Omega)} \quad \forall \chi \in H_0^1(\Omega).$$

We make the following elliptic regularity assumption. There exists a constant $s \in (0, 1]$ such that

$$\|U\|_{H^{1+s}(\Omega)} \leq C \|F\|_{H^{s-1}(\Omega)} \tag{3.13}$$

for the solution U, where C is a constant depending only on the domain Ω. In Grisvard [40], it is established that the above assumption holds for some $s \in (1/2, 1]$ if Ω is a Lipschitz polyhedral bounded domain and the coefficients $\{a_{ij}(x)\}$ are smooth enough in two or three dimensional space.

For a given Lipschitz polyhedral bounded domain $\Omega_0 \subset R^n$, we introduce a sequence of triangulations $\mathcal{T}_k$ of Ω_0. The triangulation $\mathcal{T}_k$ has a quasi-uniform mesh size h_k. For these triangulations, the finite element spaces $\tilde{V}^{h_k}$ are defined as the Lagrange finite element space of type p. The nodes associated with the degrees of freedom in $\tilde{V}^{h_k}$ are denoted by $\mathcal{N}_k$. We assume that

$$\tilde{V}^{h_1} \subset \tilde{V}^{h_2} \subset \tilde{V}^{h_3} \subset \cdots \subset \tilde{V}^{h_k} \subset \cdots \cdot.$$

and define the Galerkin projection $P_k : H_0^1(\Omega_0) \to \tilde{V}^{h_k} \cap H_0^1(\Omega_0)$ by

$$a(P_k u, v) = a(u, v), \quad \forall v \in \tilde{V}^{h_k} \cap H_0^1(\Omega_0).$$

We can use the regularity assumption to prove the following Lemma.

Lemma 3.3.3 *There exists a $s \in (1/2, 1]$ and a constant C such that*

$$\|(I - P_k)u\|_{H^{1-s}(\Omega_0)} \leq C h_k^s \|(I - P_k)u\|_{H^1(\Omega_0)}, \quad \forall u \in H_0^1(\Omega_0).$$

Proof. Given an arbitrary $F \in L^2(\Omega_0)$, let U be the solution of the following auxiliary dual problem

$$-\sum_i \sum_j \frac{\partial}{\partial x_i} a_{ij}(x) \frac{\partial U}{\partial x_j} = F \quad \text{in } \Omega_0, \qquad U = 0 \quad \text{on } \partial\Omega_0$$

Then there exists a $s \in (1/2, 1]$ such that

$$\|U\|_{H^{1+s}(\Omega_0)} \leq C_0 \|F\|_{H^{s-1}(\Omega_0)}.$$

We now apply a standard duality argument and obtain

$$\begin{aligned}\|(I - P_k)u\|_{H^{1-s}(\Omega_0)} &= \sup_{F \in L^2(\Omega_0)} \frac{(F, (I - P_k)u)_{L^2(\Omega_0)}}{\|F\|_{H^{s-1}(\Omega_0)}} = \sup_{F \in L^2(\Omega_0)} \frac{a(U, (I - P_k)u)}{\|F\|_{H^{s-1}(\Omega_0)}} \\ &= \sup_{F \in L^2(\Omega_0)} \frac{a(U - \chi, (I - P_k)u)}{\|F\|_{H^{s-1}(\Omega_0)}}, \qquad \forall \chi \in \tilde{V}^{h_k} \cap H_0^1(\Omega_0).\end{aligned}$$

We choose χ such that

$$|U - \chi|_{H^1(\Omega_0)} \leq C' h_k^s |U|_{H^{1+s}(\Omega_0)}.$$

Then we obtain

$$\begin{aligned}\|(I - P_k)u\|_{H^{1-s}(\Omega_0)} &\leq \sup_{F \in L^2(\Omega_0)} \frac{|(I - P_k)u|_{H^1(\Omega_0)} \cdot |U - \chi|_{H^1(\Omega_0)}}{\|F\|_{H^{s-1}(\Omega_0)}} \\ \leq \sup_{F \in L^2(\Omega_0)} \frac{|(I - P_k)u|_{H^1(\Omega_0)} C' h_k^s |U|_{H^{1+s}(\Omega_0)}}{\|F\|_{H^{s-1}(\Omega_0)}} &\leq C h_k^s \|(I - P_k)u\|_{H^1(\Omega_0)}. \qquad \square\end{aligned}$$

The proof of the following lemma is a generalization of a result in Yserentant [84]; cf. also Xu [83].

Lemma 3.3.4 *Assume that the coefficient matrix $\{a_{ij}(x)\}$ is in $W^{1,\infty}(\Omega_0)$ and that Assumption 3.3.2 holds. Then*

$$a(u, v) \leq C' q^{(j-i)/2} h_j^{-1} |u|_{H^1(\Omega_0)} \|v\|_{L^2(\Omega_0)}, \qquad \forall u \in \tilde{V}^{h_i}, v \in \tilde{V}^{h_j}$$

if $i \leq j$. Here C' depends only upon the shape regularity and the bound of $\{a_{ij}(x)\}$ in $W^{1,\infty}(\Omega_0)$.

Before we proceed with the proof of the above lemma, we need the following lemma whose proof is quite elementary.

Lemma 3.3.5 *There exists a constant C, which depends only on p, such that*

$$\max_{x\in T}|w(x)| \le C\cdot\frac{1}{|T|}\int_T |w(x)|dx$$

for any triangle T in R^n and any polynomial $w(x)$ of degree less than or equal to $2p-2$.

Proof. We first map the triangle T to a fixed triangle T_0 by an affine mapping. We now use the fact that any two norms on a finite dimensional space are equivalent. Then this lemma follows easily.

Proof of Lemma 3.3.4. It suffices to show that for any element $K\in\mathcal{T}_i$,

$$\int_K \sum_{l,m} a_{lm}(x)\frac{\partial u}{\partial x_l}\frac{\partial v}{\partial x_m}dx \le Cq^{(j-i)/2}h_j^{-1}|u|_{H^1(K)}\|v\|_{L^2(K)},$$

since, by the ordinary Cauchy-Schwarz inequality, this would imply

$$a(u,v)=\sum_{K\in\mathcal{T}_i}\int_K \sum_{l,m} a_{lm}(x)\frac{\partial u}{\partial x_l}\frac{\partial v}{\partial x_m}dx \le Cq^{(j-i)/2}h_j^{-1}|u|_{H^1(K)}\|v\|_{L^2(K)}$$
$$\le Cq^{(j-i)/2}h_j^{-1}\Big(\sum_{K\in\mathcal{T}_i}|u|^2_{H^1(K)}\Big)^{\frac12}\Big(\sum_{K\in\mathcal{T}_i}\|v\|^2_{L^2(K)}\Big)^{\frac12}=Cq^{(j-i)/2}h_j^{-1}|u|_{H^1(\Omega_0)}\|v\|_{L^2(\Omega_0)}.$$

Now, for a given $v\in\tilde{V}^{h_j}$, we define a $v_1\in\tilde{V}^{h_j}$ which vanishes on ∂K and agrees with v at the interior nodes of K. Let $v_0=v-v_1$. By Green's formula

$$\int_K\sum_{l,m}a_{l,m}(x)\frac{\partial u}{\partial x_l}\frac{\partial v_1}{\partial x_m}dx=-\int_K\sum_{l,m}\frac{\partial}{\partial x_m}(a_{lm}(x))\frac{\partial u}{\partial x_l}v_1dx-\int_K\sum_{l,m}a_{lm}(x)\frac{\partial^2 u}{\partial x_l\partial x_m}v_1dx$$
$$\le C|u|_{H^1(K)}\|v_1\|_{L^2(K)}+Cq^{-i}|u|_{H^1(K)}\|v_1\|_{L^2(K)}\le Cq^{(j-i)/2}h_j^{-1}|u|_{H^1(K)}\|v_1\|_{L^2(K)}.$$

The last step follows from the inverse inequality.

Let T be the union of the $\tau\in\mathcal{T}_j$, which satisfy $\overline{\tau}\cap\partial K=\emptyset$; then $supp\{v_0\}\subset\overline{T}$ and

$$|T|=O((\frac{h_i}{h_j})^{n-1}h_j^n)=O(h_i^{n-1}h_j).$$

By Lemma 3.3.5, we have

$$\|\nabla u\|_{L^2(T)}\le|T|^{\frac12}\max_{x\in T}|\nabla u(x)|\le C\frac{|T|^{1/2}}{|K|^{1/2}}\|\nabla u\|_{L^2(K)}$$
$$\le(\frac{h_i^{n-1}h_j}{h_i^n})^{1/2}\|\nabla u\|_{L^2(K)}\le Cq^{(j-i)/2}|u|_{H^1(K)}.$$

Additionally,

$$\|\nabla v_0\|^2_{L^2(T)}\le Ch_j^{n-2}\sum_{x\in\mathcal{N}_j\cap\partial K}v_0^2(x)=h_j^{n-2}\sum_{x\in\mathcal{N}_j\cap\partial K}v^2(x)\le Ch_j^{-2}\|v\|^2_{L^2(K)}.$$

Since $a(u,v) = a(u,v_1) + a(u,v_0)$, the lemma follows. $\Box$

We can also replace the L^2-norm in the above lemma by the H^s-norm by using Lemma 1.2.8.

Lemma 3.3.6 *Assume that the coefficient matrix $\{a_{ij}(x)\}$ is in $W^{1,\infty}(\Omega_0)$. Then*

$$a(u,v) \leq C'(\Omega_0, s)(q^{(j-i)/2} h_j^{-1})^s |u|_{H^1(\Omega_0)} \|v\|_{H^{1-s}(\Omega_0)}, \qquad \forall u \in \tilde{V}^{h_i}, \forall v \in \tilde{V}^{h_j}.$$

Here $C'(\Omega_0, s)$ is a generic constant which only depends upon the domain Ω_0, s, the shape regularity and the bound of $\{a_{ij}(x)\}$ in $W^{1,\infty}(\Omega_0)$.

Proof. For each fixed $u \in \tilde{V}^{h_i}$, there is a bounded linear functional $A_u : \tilde{V}^{h_j} \to R$ such that

$$A_u(v) = a(u,v).$$

We observe that

$$|A_u(v)| \leq C|u|_{H^1(\Omega_0)} \|v\|_{H^1(\Omega_0)}.$$

Therefore

$$\|A_u\|_{\tilde{V}^{h_j} \subset H^1(\Omega_0) \to R} \leq C|u|_{H^1(\Omega_0)}.$$

By Lemma 3.3.4, we also have

$$\|A_u\|_{\tilde{V}^{h_j} \subset L_2(\Omega_0) \to R} \leq C' q^{(j-i)/2} h_j^{-1} |u|_{H^1(\Omega_0)}.$$

We now can apply Lemma 1.2.8 to get a good bound for $\|A_u\|_{\tilde{V}^{h_j} \subset H^{1-s}(\Omega_0) \to R}$. We have

$$\begin{aligned} \|A_u\|_{\tilde{V}^{h_j} \subset H^{1-s}(\Omega_0) \to R} &\leq C(\Omega_0, s) \|A_u\|^{1-s}_{\tilde{V}^{h_j} \subset H^1(\Omega_0) \to R} \cdot \|A_u\|^{s}_{\tilde{V}^{h_j} \subset L_2(\Omega_0) \to R} \\ &\leq C'(\Omega_0, s)(q^{(j-i)/2} h_j^{-1})^s \cdot |u|_{H^1(\Omega_0)}. \end{aligned}$$

From this result, the lemma follows easily. $\Box$

We can now return to the original problem. The following lemma is of importance.

Lemma 3.3.7 *Assume that the coefficient matrix $\{a_{ij}(x)\}$ is in $W^{1,\infty}(\Omega)$ and that Assumption 3.3.2 holds. Then*

$$a(u,v) \leq C(\Omega_j, s) q^{(j-i-2)s/2} |u|_{H^1(\Omega)} |v|_{H^1(\Omega)}, \quad \forall u \in Range(P_i^i), v \in Range(P_j^j - P_j^{j-1}).$$

Here $C(\Omega_j, s)$ is a constant which only depends upon the subdomain Ω_j, s, the shape regularity and the bound of $\{a_{ij}(x)\}$ in $W^{1,\infty}(\Omega)$.

Proof. If we combine the results of Lemma 3.3.6 and Lemma 3.3.3, we will obtain

$$\begin{aligned} a(u,v) = a_{\Omega_j}(u,v) &\le C'(\Omega_j,s)(q^{(j-i)/2}h_j^{-1})^s \cdot |u|_{H^1(\Omega_j)} \cdot \|v\|_{H^{1-s}(\Omega_j)} \\ &\le C'(\Omega_j,s)(q^{(j-i)/2}h_j^{-1})^s \cdot |u|_{H^1(\Omega_j)} \cdot C \cdot h_{j-1}^s \cdot \|v\|_{H^1(\Omega_j)} \\ &\le C(\Omega_j,s)(q^{(j-i)/2}h_j^{-1})^s |u|_{H^1(\Omega_j)} \cdot C \cdot h_{j-1}^s |v|_{H^1(\Omega_j)} \\ &\le C(\Omega_j,s) \cdot q^{(j-i-2)s/2} \cdot |u|_{H^1(\Omega_j)} \cdot |v|_{H^1(\Omega_j)} \\ &\le C(\Omega_j,s) \cdot q^{(j-i-2)s/2} \cdot |u|_{H^1(\Omega)} \cdot |v|_{H^1(\Omega)}. \square \end{aligned}$$

We can now prove the main Theorem in this subsection.

Theorem 3.3.4 *Under Assumption 3.3.1 and 3.3.2, the operator $P_a^{(2)}$ has a condition number which is independent of k and the number of degrees of freedom.*

Proof. By Theorem 3.3.2, it is sufficient to prove that the eigenvalues of $P_a^{(2)}$ are uniformly bounded from above by a constant. Let $u_i = (P_i^i - P_i^{i-1})u_h$ for $i = 1, 2, \cdots, k$; then $P_a^{(2)}u_h = u_1 + u_2 + \cdots + u_k$. Lemma 3.3.7 shows that these terms are increasingly orthogonal:

$$a(u_l, u_m) \le const.q^{|l-m-2|s/2}(a(u_l,u_l))^{1/2}(a(u_m,u_m))^{1/2}.$$

We now note that

$$a(P_a^{(2)}u_h, P_a^{(2)}u_h) = \sum_{i,j=1}^{k} a(u_i,u_j) \le U^T A U.$$

Here

$$A = const. * q^{-s} \begin{pmatrix} 1 & q^{1s/2} & q^{2s/2} & \cdots & q^{(k-1)s/2} \\ q^{1s/2} & 1 & q^{1s/2} & \cdots & q^{(k-2)s/2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ q^{(k-1)s/2} & q^{(k-2)s/2} & q^{(k-3)s/2} & \cdots & 1 \end{pmatrix}$$

and

$$U^T = (a(u_1,u_1)^{1/2}, a(u_2,u_2)^{1/2}, \cdots, a(u_k,u_k)^{1/2}).$$

It is easy to show that the Euclidean norm of A is uniformly bounded. Thus

$$a(P_a^{(2)}u_h, P_a^{(2)}u_h) \le |A|_{l_2} \sum_{i=1}^{k} a(u_i,u_i) \le C \sum_{i=1}^{k} a(u_i,u_i).$$

To complete the proof, we note that

$$a(u_i,u_i) = a(P_i^i u_h - P_i^{i-1}u_h, P_i^i u_h - P_i^{i-1}u_h) = a(P_i^i u_h - P_i^{i-1}u_h, P_i^i u_h),$$

since $P_i^i u_h - P_i^{i-1} u_h$ is h_{i-1}-harmonic. By using elementary properties of the projections, we find that $a(u_i, u_i) = a(u_i, u_h)$. Therefore

$$\sum_{i=1}^{k} a(u_i, u_i) = \sum_{i=1}^{k} a(u_i, u_h) = a(P_a^{(2)} u_h, u_h).$$

Thus

$$a(P_a^{(2)} u_h, P_a^{(2)} u_h) \leq C \cdot a(P_a^{(2)} u_h, u_h),$$

from which the upper bound on $P_a^{(2)}$ follows. □

However, we can replace Assumption 3.3.2 by a much weaker assumption such that Theorem 3.3.4 holds.

Assumption 3.3.4 *The mesh sizes h_i are uniformly bounded from above and below by const.$q^{N(i)}$. Here $N(i)$ is a strictly increasing sequence of positive integers and q is some constant less than 1.*

Theorem 3.3.5 *Under Assumptions 3.3.1 and 3.3.4, the operator $P_a^{(2)}$ has a condition number which is independent of k and the number of degrees of freedom.*

Proof. It is sufficient to prove that $P_a^{(2)}$ has a uniform upper bound. It is obvious that we can introduce intermediate subdomains and mesh sizes such that the resulting new AFAC2 operator $P_{a,n}^{(2)}$ is the same as $P_a^{(2)}$. Now $P_{a,n}^{(2)}$ satisfies Assumption 3.3.2, so it follows that $P_{a,n}^{(2)}$ has a uniformly upper bound. Therefore $P_a^{(2)}$ has a uniformly upper bound. □

3.3.3 Convergence Rate Estimate for FAC with Inexact Solvers

We consider T_i^i which are spectrally equivalent to P_i^i for $i = 1, 2, ..., k$, in the sense that there exist constants ω_1 and ω_2 with $0 < \omega_1 < \omega_2 < 2$ such that $\omega_1 P_i^i \leq T_i^i \leq \omega_2 P_i^i$, for all $i = 1, 2, ..., k$. In this subsection, we will establish the optimality for FAC based on such T_i^i. We make the same assumptions about the domains $\Omega_1, \cdots, \Omega_k$ as in the last subsection.

Theorem 3.3.6 *Under Assumptions 3.3.1 and 3.3.2, the energy norm of the basic FAC algorithm based on the $T_i^i, 1 \leq i \leq k$, is bounded by a constant less than one with the constant independent of k and the number of degrees of freedom, but depending upon ω_1, ω_2 and the shape regularity. Therefore the corresponding symmetrized FAC algorithm has a condition number which depends only upon ω_1, ω_2 and the shape regularity.*

Proof. Let $E_i = (I - T_i^i) \cdots (I - T_1^1)$. Then $E_{i-1} - E_i = T_i^i E_{i-1}$, for $i = 1, 2, \cdots, k$. and we obtain

$$\|E_{i-1}u\|_a^2 - \|E_i u\|_a^2 = \|T_i^i E_{i-1}u\|_a^2 + 2a(T_i^i E_{i-1}u, E_i u)$$
$$= a(T_i^i(2I - T_i^i)E_{i-1}u, E_{i-1}u) \geq (2 - \omega_2)a(T_i^i E_{i-1}u, E_{i-1}u).$$

Summing these terms for $i = 1, 2, \cdots, k$, we get

$$\|u\|_a^2 - \|E_k u\|_a^2 \geq (2 - \omega_2)\sum_{i=1}^{k} a(T_i^i E_{i-1}u, E_{i-1}u).$$

We can write u as

$$u = T_1^1 u + (I - T_1^1)u = T_1^1 u + E_1 u = T_1^1 E_0 u + T_2^2 E_1 u + (I - T_2^2)E_1 u$$
$$= T_1^1 E_0 u + T_2^2 E_1 u + E_2 u = \cdots = \sum_{j=1}^{i-1} T_j^j E_{j-1}u + E_{i-1}u.$$

If we define θ_{ij} by

$$\theta_{ij} = \begin{cases} 1/\sqrt{\omega_1} & \text{if } i = j \\ \max_{v,w \in V^h} \frac{a((P_i^i - P_i^{i-1})v, T_j^j w)}{a((P_i^i - P_i^{i-1})v, v)^{\frac{1}{2}} \cdot a(T_j^j w, w)^{\frac{1}{2}}} & \text{if } i \geq j \\ \theta_{ji} & \text{if } i \leq j, \end{cases}$$

we will obtain

$$a((P_i^i - P_i^{i-1})u, u) = a((P_i^i - P_i^{i-1})u, E_{i-1}u) + \sum_{j=1}^{i-1} a((P_i^i - P_i^{i-1})u, T_j^j E_{j-1}u)$$
$$\leq a((P_i^i - P_i^{i-1})u, u)^{\frac{1}{2}} \cdot a((P_i^i - P_i^{i-1})E_{i-1}u, E_{i-1}u)^{\frac{1}{2}}$$
$$+ \sum_{j=1}^{i-1} \theta_{ij} a((P_i^i - P_i^{i-1})u, u)^{\frac{1}{2}} \cdot a(T_j^j E_{j-1}u, E_{j-1}u)^{\frac{1}{2}}$$
$$\leq a((P_i^i - P_i^{i-1})u, u)^{\frac{1}{2}} \cdot a(P_i^i E_{i-1}u, E_{i-1}u)^{\frac{1}{2}}$$
$$+ \sum_{j=1}^{i-1} \theta_{ij} a((P_i^i - P_i^{i-1})u, u)^{\frac{1}{2}} \cdot a(T_j^j E_{j-1}u, E_{j-1}u)^{\frac{1}{2}}$$
$$\leq \frac{1}{\sqrt{\omega_1}} a((P_i^i - P_i^{i-1})u, u)^{\frac{1}{2}} \cdot a(T_i^i E_{i-1}u, E_{i-1}u)^{\frac{1}{2}}$$
$$+ \sum_{j=1}^{i-1} \theta_{ij} a((P_i^i - P_i^{i-1})u, u)^{\frac{1}{2}} \cdot a(T_j^j E_{j-1}u, E_{j-1}u)^{\frac{1}{2}}$$
$$= a((P_i^i - P_i^{i-1})u, u)^{\frac{1}{2}} \cdot (\sum_{j=1}^{i} \theta_{ij} a(T_j^j E_{j-1}u, E_{j-1}u)^{\frac{1}{2}}).$$

Thus, for $1 \leq i \leq k$, we get

$$a((P_i^i - P_i^{i-1})u, u) \leq (\sum_{j=1}^{k} \theta_{ij} a(T_j^j E_{j-1} u, E_{j-1} u)^{\frac{1}{2}})^2.$$

Finally we conclude that

$$a(P_a^{(2)} u, u) \leq \|\theta\|_2^2 \sum_{i=1}^{k} a(T_i^i E_{i-1} u, E_{i-1} u)$$

where $P_a^{(2)}$ is the AFAC2 operator and $\|\theta\|_2$ is the l^2-norm of the matrix θ.

We now need to estimate the θ_{ij}. Observe that $Null(T_j^j) = Null(P_j^j)$. We also have $Range(T_j^j) = Null(T_j^j)^{\perp}$ and $Range(P_j^j) = Null(P_j^j)^{\perp}$ because T_j^j and P_j^j are energy-symmetric. Now we can conclude that $Range(T_j^j) = Range(P_j^j)$. It is also easy to see that $Range((T_j^j)^{\frac{1}{2}}) = Range(T_j^j)$. For $i \geq j$,

$$\begin{aligned}
\theta_{ij} &= \max_{v,w \in V^h} \frac{a((P_i^i - P_i^{i-1})v, T_j^j w)}{a((P_i^i - P_i^{i-1})v, v)^{\frac{1}{2}} \cdot a(P_j^j (T_j^j)^{\frac{1}{2}} w, (T_j^j)^{\frac{1}{2}} w)^{\frac{1}{2}}} \\
&\leq \sqrt{\omega_2} \max_{v,w \in V^h} \frac{a((P_i^i - P_i^{i-1})v, T_j^j w)}{a((P_i^i - P_i^{i-1})v, v)^{\frac{1}{2}} \cdot a(T_j^j (T_j^j)^{\frac{1}{2}} w, (T_j^j)^{\frac{1}{2}} w)^{\frac{1}{2}}} \\
&= \sqrt{\omega_2} \max_{v,w \in V^h} \frac{a((P_i^i - P_i^{i-1})v, T_j^j w)}{a((P_i^i - P_i^{i-1})v, (P_i^i - P_i^{i-1})v)^{\frac{1}{2}} \cdot a(T_j^j w, T_j^j w)^{\frac{1}{2}}} \\
&\leq \sqrt{\omega_2} \max_{u_1 \in Range(P_i^i - P_i^{i-1}), u_2 \in Range(P_j^j)} \frac{a(u_1, u_2)}{a(u_1, u_1)^{\frac{1}{2}} \cdot a(u_2, u_2)^{\frac{1}{2}}} \\
&\leq \sqrt{\omega_2} C q^{(i-j-2)s/2}.
\end{aligned}$$

The last step is a result of Lemma 3.3.7. So we have $\|\theta\|_2^2 \leq \|\theta\|_1^2 \leq C'$. Therefore

$$\begin{aligned}
a(P_a^{(2)} u, u) = C' \sum_{i=1}^{k} a(T_i^i E_{i-1} u, E_{i-1} u) &\leq \frac{C'}{2 - \omega_2} (\|u\|_a^2 - \|E_k u\|_a^2) \\
&\equiv C_0 (\|u\|_a^2 - \|E_k u\|_a^2)
\end{aligned}$$

where C_0 is a constant which only depends on ω_1, ω_2, q, the subdomains and the shape regularity. By Theorem 3.3.2, we have $P_a^{(2)} \geq cI$, where $c \geq 0$ is a constant which is independent of k and the number of degrees of freedom and it follows that $\|E_k u\|_a^2 \leq (C_0 - c)/C_0 \cdot \|u\|_a^2$. $\square$

When $T_i^i = P_i^i$ for all i, we can prove this theorem under Assumptions 3.3.1 and 3.3.4 by using the same trick as in Theorem 3.3.5 and tracing the proof of above theorem. Consequently, we have obtained another proof of Theorem 3.3.1 given in Widlund [81].

We conclude this subsection by describing how to generalize our theoretical results to more general elliptic problems with mixed type boundary condition. We can prove that Theorem 3.3.2 and 3.3.3 hold if we replace Assumption 3.3.1 by the following assumption. The proofs are quite similar.

Assumption 3.3.5 *For each j, there exists a bounded Lipschitz polyhedral region $\tilde{\Omega}_j$ such that $\Omega_j \subset \tilde{\Omega}_j$, $(\tilde{\Omega}_j \setminus \Omega_j) \cap \Omega = \emptyset$, $(\partial\tilde{\Omega}_j \cap \partial\Omega_{j+1}) \setminus \Gamma_N = \emptyset$, $\tilde{\Omega}_j \cap \Gamma_N = \emptyset$ and the Lipschitz constants of $\tilde{\Omega}_j \setminus \Omega_{j+1}$ are uniformly bounded.*

We also need one additional assumption to obtain the analogs of Theorem 3.3.4 and 3.3.6. We consider the homogeneous mixed type boundary value problem

$$-\sum_i \sum_j \frac{\partial}{\partial x_i} a_{ij}(x) \frac{\partial U}{\partial x_j} = F \qquad \text{in } \Omega,$$

$$U = 0 \qquad \text{on } \Gamma_D, \qquad \sum_j \sum_i a_{ij}(x) \frac{\partial U}{\partial x_j} n_i = 0 \qquad \text{on } \Gamma_N$$

in a weak sense. Let $U \in H^1_{\Gamma_D}(\Omega) \equiv \{u | u \in H^1(\Omega) \text{ and } \gamma u = 0 \quad \text{on } \Gamma_D\}$ be the solution of

$$a(U, \chi) = (F, \chi)_{L^2(\Omega)} \quad \forall \chi \in H^1_{\Gamma_D}(\Omega). \tag{3.14}$$

Let us define a norm which is similar to $H^{s-1}(\Omega)$ for $0 \le s \le 1$ by

$$\|F\|_{H^{s-1}_{\Gamma_D}(\Omega)} \equiv \sup_{u \in H^{1-s}_{\Gamma_D}(\Omega)} \frac{(F, u)_{L^2(\Omega)}}{\|u\|_{H^{1-s}(\Omega)}}$$

where $H^{1-s}_{\Gamma_D}(\Omega)$ is the closure of the subspace of $C^\infty(\overline{\Omega})$ which is zero on Γ_D with respect to the norm H^{1-s}-norm and make the following elliptic regularity assumption.

Assumption 3.3.6 *There exists a constant $s \in (0, 1]$ such that*

$$\|U\|_{H^{1+s}(\Omega)} \le C \|F\|_{H^{s-1}_{\Gamma_D}(\Omega)}$$

for the solution U of (3.14), where C is a constant depending only on the domain Ω.

If the additional Assumption 3.3.6 is true for the subdomains Ω_j with Neumann boundary condition on $\Gamma_N \cap \partial\Omega_j$ for $j \ge 2$, then we can prove Theorem 3.3.4 and 3.3.6 by using similar techniques as before.

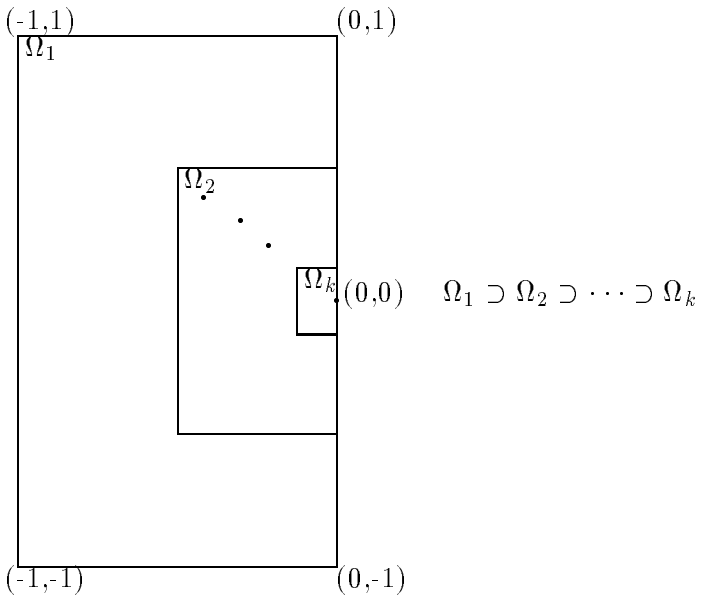

Figure 3.1: Model Problem 1

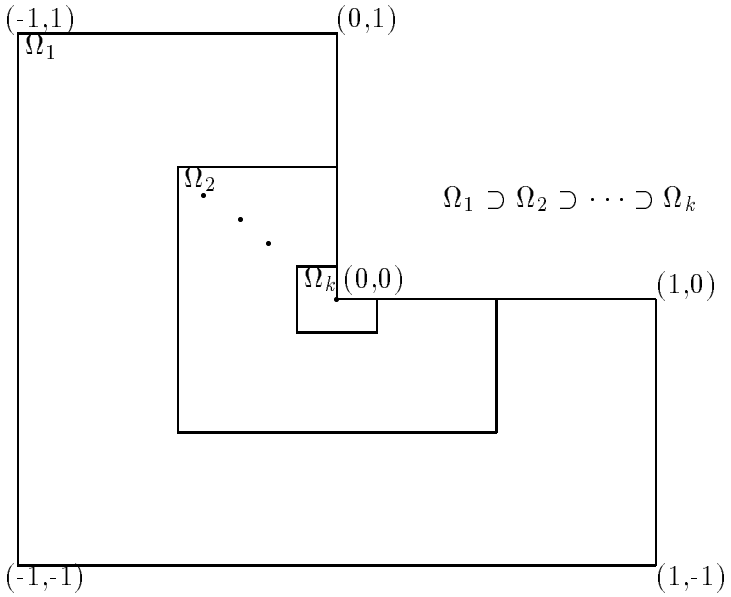

Figure 3.2: Model Problem 2

$h_1 = 1/8$									
k	2	3	4	5	6	7	8	9	10
min. ev.	0.5801	0.5770	0.5765	0.5762	0.5761	0.5760	0.5759	0.5759	0.5759
max. ev.	1.421	1.420	1.420	1.420	1.421	1.421	1.421	1.422	1.422
cond. no.	2.450	2.461	2.463	2.464	2.467	2.467	2.467	2.469	2.469
no. of iter.	9	10	10	10	10	10	10	10	10
$h_1 = 1/16$									
k	2	3	4	5	6	7	8	9	10
min. ev.	0.5811	0.5771	0.5768	0.5767	0.5766	0.5766	0.5766	0.5765	0.5765
max. ev.	1.420	1.422	1.422	1.422	1.421	1.421	1.421	1.421	1.421
cond. no.	2.444	2.464	2.465	2.466	2.464	2.464	2.464	2.465	2.465
no. of iter.	9	10	10	10	10	10	10	10	10

Table 3.1: AFAC2 for Model problem 1

3.4 Numerical Results

In this section, we report on some numerical experiments, which verify our theoretical results. We first consider Model Problem 1 in which the Ω_i, for $i = 1, 2, \cdots, k$, are rectangles and similar to each other; cf. Fig. 3.1. We then consider Model Problem 2 in which the Ω_i are L-shaped and similar to each other; cf. Fig. 3.2. In all experiments, the differential operator is the Laplacian, the subdomains are triangulated using right triangles of equal size, the finite element space is the Lagrange element of type 1 and we consider only the Dirichlet boundary value problem.

We first want to see whether the elliptic regularity assumption affects the condition number very much or not. In these experiments, we choose arbitrary right hand sides and the iterations are stopped when the residual size with respect to the energy norm has been reduced by a factor 10^{-6}. In all experiments, we choose $h_{i+1} = h_i/2$ and the ratio of the diameters of Ω_{i+1} and Ω_i to be $1/2$ for $i = 1, 2, \cdots, k-1$. These results are summarized in Tables 3.1-3.8.

From these data, we can conclude that the elliptic regularity does not affect the values of the condition numbers very much.

In Table 3.9, we consider the FAC algorithm for model problem 1, replacing the projections P_i^i by the approximation T_i^i obtained by using one multigrid V-cycle with zero

$h_1 = 1/8$									
k	2	3	4	5	6	7	8	9	10
min. ev.	0.8226	0.8227	0.8227	0.8228	0.8228	0.8228	0.8228	0.8228	0.8228
max. ev.	1.000	1.000	1.000	0.9999	0.9999	0.9999	0.9998	0.9998	0.9998
cond. no.	1.216	1.216	1.216	1.215	1.215	1.215	1.215	1.215	1.215
no. of iter.	5	5	5	5	5	5	5	5	5
$h_1 = 1/16$									
k	2	3	4	5	6	7	8	9	10
min. ev.	0.8213	0.8214	0.8214	0.8214	0.8214	0.8214	0.8214	0.8215	0.8215
max. ev.	1.000	1.000	1.000	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
cond. no.	1.218	1.217	1.217	1.217	1.217	1.217	1.217	1.217	1.217
no. of iter.	5	5	5	5	5	5	5	5	5

Table 3.2: FAC for Model problem 1 with $T_i^i = P_i^i$

$h_1 = 1/8$									
k	2	3	4	5	6	7	8	9	10
min. ev.	0.5367	0.5358	0.5354	0.5352	0.5351	0.5350	0.5350	0.5350	0.5349
max. ev.	0.9375	0.9826	0.9839	0.9846	0.9850	0.9853	0.9854	0.9854	0.9854
cond. no.	1.747	1.834	1.838	1.840	1.841	1.842	1.842	1.842	1.842
no. of iter.	7	8	8	8	8	8	8	8	8
$h_1 = 1/16$									
k	2	3	4	5	6	7	8	9	10
min. ev.	0.5381	0.5355	0.5352	0.5349	0.5346	0.5345	0.5344	0.5343	0.5342
max. ev.	0.9375	0.9837	0.9893	0.9908	0.9914	0.9918	0.9920	0.9922	0.9923
cond. no.	1.742	1.837	1.848	1.852	1.854	1.856	1.856	1.857	1.858
no. of iter.	7	8	8	8	8	8	8	8	8

Table 3.3: FAC for Model problem 1 with $T_i^i = 0.5 * P_i^i$

$h_1 = 1/8$									
k	2	3	4	5	6	7	8	9	10
min. ev.	0.5366	0.5374	0.5379	0.5382	0.5383	0.5384	0.5384	0.5385	0.5385
max. ev.	0.9375	0.9820	0.9818	0.9818	0.9818	0.9819	0.9819	0.9819	0.9819
cond. no.	1.747	1.827	1.825	1.824	1.824	1.824	1.824	1.823	1.823
no. of iter.	7	8	8	8	8	8	8	8	8
$h_1 = 1/16$									
k	2	3	4	5	6	7	8	9	10
min. ev.	0.5379	0.5357	0.5357	0.5355	0.5354	0.5353	0.5352	0.5351	0.5350
max. ev.	0.9375	0.9835	0.9894	0.9908	0.9914	0.9917	0.9920	0.9922	0.9925
cond. no.	1.743	1.836	1.847	1.850	1.852	1.853	1.854	1.854	1.855
no. of iter.	7	8	8	8	8	8	8	8	8

Table 3.4: FAC for Model problem 1 with $T_i^i = 1.5 * P_i^i$

$h_1 = 1/8$									
k	2	3	4	5	6	7	8	9	10
min. ev.	0.5810	0.5786	0.5767	0.5763	0.5761	0.5759	0.5758	0.5757	0.5757
max. ev.	1.421	1.419	1.421	1.422	1.423	1.424	1.425	1.426	1.427
cond. no.	2.446	2.452	2.464	2.467	2.470	2.473	2.475	2.477	2.479
no. of iter.	9	9	10	10	10	10	10	10	10
$h_1 = 1/16$									
k	2	3	4	5	6	7	8	9	10
min. ev.	0.5840	0.5819	0.5812	0.5765	0.5763	0.5761	0.5760	0.5759	0.5758
max. ev.	1.421	1.419	1.419	1.422	1.422	1.423	1.424	1.425	1.426
cond. no.	2.433	2.439	2.442	2.467	2.467	2.470	2.472	2.474	2.477
no. of iter.	9	9	9	10	10	10	10	10	10

Table 3.5: AFAC2 for Model problem 2

$h_1 = 1/8$									
k	2	3	4	5	6	7	8	9	10
min. ev.	0.8229	0.8231	0.8233	0.8234	0.8234	0.8234	0.8235	0.8235	0.8235
max. ev.	0.9999	0.9999	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
cond. no.	1.215	1.215	1.214	1.214	1.214	1.214	1.214	1.214	1.214
no. of iter.	5	5	5	5	5	5	5	5	5
$h_1 = 1/16$									
k	2	3	4	5	6	7	8	9	10
min. ev.	0.8296	0.8222	0.8223	0.8223	0.8224	0.8224	0.8225	0.8227	0.8228
max. ev.	1.000	0.9999	0.9999	0.9999	0.9999	0.9998	0.9998	0.9998	0.9998
cond. no.	1.205	1.216	1.216	1.216	1.216	1.216	1.216	1.215	1.215
no. of iter.	4	5	5	5	5	5	5	5	5

Table 3.6: FAC for Model problem 2 with $T_i^i = P_i^i$

$h_1 = 1/8$									
k	2	3	4	5	6	7	8	9	10
min. ev.	0.5392	0.5365	0.5358	0.5355	0.5353	0.5351	0.5350	0.5349	0.5349
max. ev.	0.9375	0.9805	0.9860	0.9900	0.9927	0.9944	0.9955	0.9962	0.9967
cond. no.	1.739	1.828	1.840	1.849	1.854	1.858	1.861	1.862	1.863
no. of iter.	7	8	8	8	8	8	8	8	8
$h_1 = 1/16$									
k	2	3	4	5	6	7	8	9	10
min. ev.	0.5404	0.5396	0.5345	0.5340	0.5336	0.5333	0.5331	0.5330	0.5330
max. ev.	0.9375	0.9825	0.9889	0.9907	0.9931	0.9949	0.9958	0.9969	0.9972
cond. no.	1.735	1.821	1.850	1.855	1.861	1.866	1.868	1.870	1.871
no. of iter.	7	7	8	8	8	8	8	8	8

Table 3.7: FAC for Model problem 2 with $T_i^i = 0.5 * P_i^i$

$h_1 = 1/8$									
k	2	3	4	5	6	7	8	9	10
min. ev.	0.5367	0.5373	0.5379	0.5381	0.5382	0.5383	0.5384	0.5384	0.5384
max. ev.	0.9375	0.9806	0.9843	0.9883	0.9884	0.9895	0.9901	0.9905	0.9907
cond. no.	1.747	1.825	1.830	1.837	1.836	1.838	1.839	1.840	1.840
no. of iter.	7	8	8	8	8	8	8	8	8
$h_1 = 1/16$									
k	2	3	4	5	6	7	8	9	10
min. ev.	0.5376	0.5353	0.5354	0.5353	0.5352	0.5351	0.5350	0.5350	0.5349
max. ev.	0.9375	0.9831	0.9878	0.9905	0.9912	0.9916	0.9920	0.9922	0.9927
cond. no.	1.744	1.837	1.845	1.850	1.852	1.853	1.854	1.855	1.856
no. of iter.	7	8	8	8	8	8	8	8	8

Table 3.8: FAC for Model problem 2 with $T_i^i = 1.5 * P_i^i$

initial guess. We will give a brief description of these cycles here. For the transfer of corrections from coarse to fine grids, we use piecewise bilinear interpolation. For the transfer of residuals from fine to coarse grids, we use full weighting, which, up to a scalar factor, is the Euclidean adjoint of piecewise bilinear interpolation. For the smoothing of the errors, we use red-black Gauss-Seidel iteration, with three half steps, red-black-red, before and after each coarse grid correction step. By the standard theory of multigrid V-cycle method, P_i^i and T_i^i are spectrally equivalent. From our experiments, we can conclude that the condition number of FAC, with the T_i^i, is almost the same as that of the original FAC. However it is much cheaper to compute $T_i^i u_h$ than $P_i^i u_h$.

In Table 3.10, we consider the modified FAC algorithm mentioned in Section 3.2 for Model Problem 1 with the above multigrid V-cycle T_i^i. We see that the upper and lower bounds are both larger than those of Table 3.9, but that the upper bounds increase more than the lower bounds. Consequently the condition numbers are bigger than those of Table 3.9. However it is easier to parallelize the algorithm.

In Table 3.11, we consider the AFAC2 algorithm for Model Problem 1 but replace the P_i^i and P_i^{i-1} by the previously described multigrid V-cycles. We see that the resulting new algorithm appears to have a uniformly bounded condition number. So far we do not know how to prove its optimality.

$h_1 = 1/8$									
k	2	3	4	5	6	7	8	9	10
min. ev.	0.8227	0.8230	0.8230	0.8230	0.8230	0.8230	0.8230	0.8230	0.8230
max. ev.	0.9983	0.9985	0.9986	0.9987	0.9988	0.9988	0.9988	0.9989	0.9989
cond. no.	1.213	1.213	1.213	1.213	1.214	1.214	1.214	1.214	1.214
no. of iter.	5	5	5	5	5	5	5	5	5
$h_1 = 1/16$									
k	2	3	4	5	6	7	8	9	10
min. ev.	0.8230	0.8230	0.8229	0.8228	0.8228	0.8227	0.8227	0.8226	0.8226
max. ev.	0.9983	0.9985	0.9986	0.9986	0.9986	0.9987	0.9987	0.9987	0.9988
cond. no.	1.213	1.213	1.214	1.214	1.214	1.214	1.214	1.214	1.214
no. of iter.	5	5	5	5	5	5	5	5	5

Table 3.9: FAC for Model problem 1 with T_i^i being one multigrid V-cycle

$h_1 = 1/8$									
k	2	3	4	5	6	7	8	9	10
min. ev.	1.391	1.390	1.389	1.389	1.389	1.389	1.389	1.388	1.388
max. ev.	2.212	2.203	2.202	2.202	2.202	2.202	2.203	2.203	2.204
cond. no.	1.590	1.585	1.585	1.585	1.585	1.585	1.586	1.587	1.588
no. of iter.	7	7	7	7	7	7	7	7	7
$h_1 = 1/16$									
k	2	3	4	5	6	7	8	9	10
min. ev.	1.396	1.395	1.394	1.393	1.393	1.393	1.392	1.392	1.391
max. ev.	2.207	2.202	2.200	2.200	2.199	2.199	2.199	2.199	2.198
cond. no.	1.581	1.578	1.578	1.579	1.579	1.579	1.580	1.580	1.580
no. of iter.	7	7	7	7	7	7	7	7	7

Table 3.10: Modified FAC for Model problem 1 with T_i^i being one multigrid V-cycle

$h_1 = 1/8$									
k	2	3	4	5	6	7	8	9	10
min. ev.	0.5747	0.5728	0.5722	0.5719	0.5718	0.5717	0.5716	0.5716	0.5715
max. ev.	1.415	1.413	1.412	1.412	1.412	1.412	1.412	1.412	1.412
cond. no.	2.462	2.467	2.468	2.469	2.469	2.470	2.470	2.470	2.471
no. of iter.	10	10	10	10	10	10	10	10	10
$h_1 = 1/16$									
k	2	3	4	5	6	7	8	9	10
min. ev.	0.5752	0.5742	0.5739	0.5739	0.5737	0.5735	0.5735	0.5735	0.5734
max. ev.	1.418	1.417	1.417	1.417	1.417	1.417	1.417	1.416	1.416
cond. no.	2.465	2.468	2.469	2.469	2.470	2.471	2.471	2.469	2.469
no. of iter.	10	10	10	10	10	10	10	10	10

Table 3.11: FAC for Model problem 1 with T_i^i and T_i^{i-1} being one multigrid V-cycles

Chapter 4

Multilevel Additive Schwarz Methods with Partial Refinement

4.1 Introduction

In this chapter, we consider the same elliptic problems with composite meshes as in Chapter 3. To simplify the presentation, we only consider second order elliptic problems (1.1)-(1.3) with homogeneous Dirichlet boundary condition and Lagrange elements of type 1. However, our analysis can be generalized to more general mixed type boundary condition and more general Lagrange elements without any difficulty.

We use the framework of Multilevel additive Schwarz methods which is described in Dryja and Widlund [33], and Zhang [86] to develop a new kind of algorithm for composite finite element problems. If we compare this kind of algorithm with the AFAC methods of Chapter 3, we can see that they are both additive Schwarz methods. In the new algorithms we decompose the problems corresponding to the refined subregions with uniform mesh size, used in AFAC, into many much smaller problems which are much easier to solve. However, this is at the expense of slower convergence of the algorithms.

We apply quasi-interpolants to prove that the iteration operators of these methods have a uniform lower bound under quite general assumptions. We remark that our proof can be applied to the case of refinement everywhere and is different from Zhang's which was obtained by considering a decomposition based on the Galerkin projection on a larger convex domain. Bornemann and Yserentant [8] have also obtained another proof based on the use of K-functionals.

4.2 General Multilevel Additive Schwarz Methods

4.2.1 Basic Two-Level Methods

In Chapter 2, we described a framework of general additive Schwarz methods. Therefore it is now sufficient to define the subspace decomposition; different subspace decompositions give different algorithms.

Following Dryja and Widlund [29], [31], we define two levels of triangulations. We start with a coarse triangulation $\mathcal{T}^H = \{\Omega_i\}_{i=1}^N$. Each Ω_i is then further divided into smaller elements to obtain a fine triangulation $\mathcal{T}^h$. We assume that the Ω_i in the triangulation $\mathcal{T}^H$ have a quasi-uniform mesh size H and that the elements in the triangulation $\mathcal{T}^h$ a quasi-uniform mesh size h.

To get an overlapping decomposition of Ω, we extend each Ω_i to a larger region $\tilde{\Omega}_i$, such that $cH \leq \text{dist}\{\partial\Omega_i, \partial\tilde{\Omega}_i\} \leq CH$, and the $\partial\tilde{\Omega}_i$ align with the boundaries of elements in $\mathcal{T}^h$. We cut off the part of each $\tilde{\Omega}_i$ that is outside of Ω.

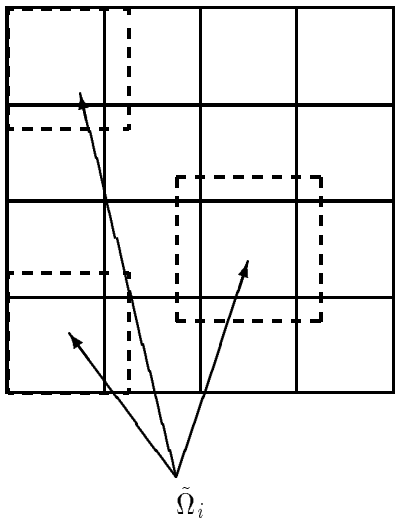

Figure 4.1: Two levels of triangulation

It is easy to see that $\{\tilde{\Omega}_i\}$ forms a *finite covering* of domain Ω. We see that we can color $\{\tilde{\Omega}_i\}$, using at most N_c colors in such a way that the subdomains of the same color are disjoint. Here N_c is a constant which depends only on the shape regularity of $\mathcal{T}^H$.

Due to the generous overlap of $\{\tilde{\Omega}_i\}$, we have a *partition of unity* $\{\theta_i\}$ satisfying

$$\sum_{i=1}^{N} \theta_i = 1 \text{ with } \theta_i \in W^{1,\infty}(\tilde{\Omega}_i), 0 \leq \theta_i \leq 1, \text{ and } |\theta_i|_{W^{1,\infty}} \leq C/H.$$

Let V^h and V^H be the Lagrange finite elements of type 1 associated with the triangulations $\mathcal{T}^h$ and $\mathcal{T}^H$, respectively. Let $V_0 = V^H$ and $V_i = V^h(\tilde{\Omega}_i) = V^h \cap H_0^1(\tilde{\Omega}_i)$. We obtain a decomposition of the finite element space V^h

$$V^h = \sum_{i=0}^{N} V_i.$$

Let $P_{V_i} : V^h \to V_i$, be the Galerkin projection defined by

$$a(P_{V_i} u, v) = a(u, v), \qquad \forall v \in V_i.$$

The matrix form of P_{V_i}, after a permutation, is

$$P_{V_i} = \begin{pmatrix} K_i^{-1} & 0 \\ 0 & 0 \end{pmatrix} K = \begin{pmatrix} K_i & 0 \\ 0 & 0 \end{pmatrix}^{\dagger} K,$$

where K is the stiffness matrix associated with the domain Ω and K_i is the stiffness matrix associated with the Dirichlet problem on the subdomain $\tilde{\Omega}_i$. The additive Schwarz operator, P is given by

$$P = \sum_{i=0}^{N} P_{V_i}.$$

The additive Schwarz algorithm for equation (1.6) is to solve an auxilary linear system which is equivalent to (1.6).

Algorithm 4.2.1 *Apply conjugate gradient method to the symmetric, positive definite system*

$$P u_h = g_h,$$

with respect to the inner product $a(\cdot, \cdot)$ for an appropriate g_h such that the solution u_h is the same as that of (1.6).

In this algorithm, we only need to compute Pv_h for a given $v_h \in V^h$ in each iteration, so the explicit representation of P is not needed. For the rate of convergence, we refer to the following theorem.

Theorem 4.2.1 (Dryja and Widlund) *For the operator P defined above, there exists a constant C_0 such that*

$$C_0 a(u,u) \leq a(Pu,u) \leq (N_c+1)a(u,u), \qquad \forall u \in V^h.$$

Thus $\kappa(P) \leq (N_c+1)/C_0$, and the rate of convergence of Algorithm 4.2.1 is independent of h and H.

The proof can be found in Dryja and Widlund [31]. We remark that if we do not use the coarse space, the condition number of the operator will grow like H^{-2}. This fact was pointed out and proved by Widlund in [80].

4.2.2 General Multilevel Additive Schwarz Methods

In basic two-level methods, we need to solve a coarse problem of size $O(1/H^2)$ and some local problems of size $(H/h)^2$. If h is small, we cannot both have $1/H$ and H/h small. Thus at least one of the subproblems is large. The computation can be made cheaper by recursively using the additive Schwarz method to solve the coarser problems.

We now give a description of general multilevel additive Schwarz methods which is developed in Dryja and Widlund [33]. We define a sequence of nested triangulations $\{\mathcal{T}_{l=1}^k\}$. We start with a coarse triangulation $\mathcal{T}^1 = \{\tau_i^1\}_{i=1}^{N_1}$ with quasi-uniform mesh size h_1, where τ_i^1 represent an individual triangle. The successively finer triangulations $\mathcal{T}^l = \{\tau_i^l\}(l = 2, \cdots, L)$ are defined by dividing each triangle in the triangulation $\mathcal{T}^{l-1}$ into several triangles, i.e.

$$\mathcal{T}^1 = \{\tau_i^1\}_1^{N_1} \overset{\text{refinement}}{\Longrightarrow} \mathcal{T}^2 = \{\tau_i^1\}_1^{N_2} \overset{\text{refinement}}{\Longrightarrow} \cdots \overset{\text{refinement}}{\Longrightarrow} \mathcal{T}^k = \{\tau_i^k\}_1^{N_k}.$$

We assume that the triangulations $\mathcal{T}^l$ have quasi-uniform mesh size h_{l-1} for each l.

Let $V^{h_l}, l = 1, \cdots, k$, be the space of continuous piecewise linear element associated with the triangulation $\mathcal{T}^l$. The finite element solution, $u_h \in V^h = V^{h_k}$, satisfies

$$a(u_h, \phi_h) = f(\phi_h), \quad \forall \phi_h \in V^h = V^{h_k}. \tag{4.1}$$

We assume that there are $k-1$ sets of overlapping subdomains $\{\tilde{\Omega}_i^l\}_{i=1}^{N_l}, l = 2, 3, \cdots, k$. On each level, we have an overlapping decomposition

$$\Omega = \cup_{i=1}^{N_l} \tilde{\Omega}_i^l.$$

We assume that the sets $\{\tilde{\Omega}_i^l\}$ satisfy the following assumption.

Assumption 4.2.1 *The decomposition* $\Omega = \cup_{i=1}^{N_l} \tilde{\Omega}_i^l$ *satisfies*

- $\partial\tilde{\Omega}_i^l$ *aligns with the boundaries of level* l *triangles, i.e.* $\tilde{\Omega}_i^l$ *is the union of level* l *triangles. Diameter*$(\tilde{\Omega}_i^l) = O(h_{l-1})$.
- *On each level, the subdomains* $\{\tilde{\Omega}_i^l\}_{i=1}^{N_l}$ *form a finite covering of* Ω*, with a covering constant* N_c*, i.e. we can color* $\{\tilde{\Omega}_i^l\}_{i=1}^{N_l}$*, using at most* N_c *colors in such a way that subdomains of the same color are disjoint.*
- *On each level, associated with* $\{\tilde{\Omega}_i^l\}_{i=1}^{N_l}$*, there exists a* partition of unity $\{\theta_i^l\}$ *satisfying*

$$\sum_i \theta_i^l = 1, \text{ with } \theta_i^l \in H_0^1(\tilde{\Omega}_i^l) \cap C^0(\tilde{\Omega}_i^l), 0 \le \theta_i^l \le 1 \text{ and } |\nabla\theta_i^l| \le C/h_{l-1}.$$

One way of constructing subdomains $\{\tilde{\Omega}_i^l\}_{i=1}^{N_l}, l = 2, \cdots, k$, with the above properties is described in Dryja and Widlund [29], [31]. Each triangle $\tau_i^{l-1}, i = 1, \cdots, N_l, l = 2, \cdots, k$, is extended to a larger region $\tilde{\tau}_i^{l-1}$ so that $ch_{l-1} \le \text{dist}(\partial\tilde{\tau}_i^{l-1}, \partial\tau_i^{l-1}) \le Ch_{l-1}$, aligning $\partial\tilde{\tau}_i^{l-1}$ with the boundaries of level l triangles. We cut off the part of $\tilde{\tau}_i^{l-1}$ that is outside Ω. We use $\tilde{\tau}_i^{l-1}$ as the subdomains $\tilde{\Omega}_i^l$. Another way of constructing $\{\tilde{\Omega}_i^l\}$ is given in the next section.

Let $N_1 = 1, V_1^{h_1} = V^{h_1}$ and $V_i^{h_l} = V^{h_l} \cap H_0^1(\tilde{\Omega}_i^l)$ for $i = 1, \cdots, N_l, l = 2, \cdots, k$. The finite element space $V^h = V^{h_k}$ is represented by

$$V^h = \sum_{l=1}^{k} V^{h_l} = \sum_{l=1}^{k} \sum_{i=1}^{N_l} V_i^{h_l}.$$

Let $P_i^l : V^h \to V_i^{h_l}$, be the projection defined by

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

The k-level additive Schwarz operator P is defined by

$$P = \sum_{l=0}^{k} \sum_{i=1}^{N_l} P_i^l. \tag{4.2}$$

Instead of solving the original finite element equation (4.1), we use the following algorithm.

Algorithm 4.2.2 *Let P be the operator defined by (4.2). Apply the conjugate gradient method to the following symmetric and positive definite system*

$$Pu_h = g_h,$$

with respect to the inner product $a(\cdot,\cdot)$ for an appropriate g_h such that the solution u_h is the same as that of (4.1).

The following theorem which is given in Zhang [86] proves that this multilevel additive Schwarz method has an optimal rate of convergence.

Theorem 4.2.2 *For P defined above, the following inequalities hold*

$$C_1 a(u_h, u_h) \leq a(Pu_h, u_h) \leq C_2 a(u_h, u_h) \qquad \forall u_h \in V^h.$$

Thus $\kappa(P) \leq C_2 C_1^{-1}$. Here the constants C_1 and C_2 are independent of the mesh sizes $\{h_l\}$ and k.

4.3 Description of Multilevel Additive Schwarz Methods with Partial Refinement and the Condition Number Estimate

We can modify the general multilevel additive Schwarz methods such that they can handle the finite element problems (1.6) with composite mesh sizes.

We now give a description of multilevel additive Schwarz methods with partial refinement. Like the procedure in last section, we define a sequence of nested triangulations $\{\mathcal{T}_{l=1}^k\}$. We start with a coarse triangulation $\mathcal{T}^1 = \{\tau_i^1\}_{i=1}^{N_1}$ with quasi-uniform mesh size h_1, where τ_i^1 represent an individual triangle. The successively finer triangulations $\mathcal{T}^l = \{\tau_i^l\}(l = 2, \cdots, k)$ are defined by dividing each triangle in the triangulation $\mathcal{T}^{l-1}$ into several triangles, i.e.

$$\mathcal{T}^1 = \{\tau_i^1\}_1^{N_1} \overset{\text{refinement}}{\Longrightarrow} \mathcal{T}^2 = \{\tau_i^1\}_1^{N_2} \overset{\text{refinement}}{\Longrightarrow} \cdots \overset{\text{refinement}}{\Longrightarrow} \mathcal{T}^k = \{\tau_i^k\}_1^{N_k}.$$

We assume that the triangulations $\mathcal{T}^l$ have quasi-uniform mesh sizes h_{l-1} for each l.

Let us define $\Omega_1 = \Omega$. Then for each $2 \leq l \leq k$, we choose Ω_l, which is a subregion of Ω_{l-1}, such that $\partial\Omega_l$ aligns with boundaries of level $l-1$ triangles. Let $\tilde{V}^{h_k}, l = 1, \cdots, k$,

be the subspace of continuous piecewise linear element associated with the triangulation $\mathcal{T}^l$ of $H_0^1(\Omega)$. We also set V^{h_l} to be $\tilde{V}^{h_l} \cap H_0^1(\Omega_i)$. The finite element problem is to find $u_h \in V^h = V^{h_1} + \cdots + V^{h_k}$ satisfying

$$a(u_h, \phi_h) = f(\phi_h), \qquad \forall \phi_h \in V^h = V^{h_1} + \cdots + V^{h_k}. \tag{4.3}$$

We assume that there are $k-1$ sets of overlapping subdomain $\{\tilde{\Omega}_i^l\}_{i=1}^{N_l}, l = 2, 3, \cdots, k$. On each level, we have an overlapping decomposition

$$\Omega_l = \cup_{i=1}^{N_l} \tilde{\Omega}_i^l.$$

We also assume that there are another $k-1$ sets of overlapping subdomains $\{\tilde{\Omega}_i^l\}_{i=N_l+1}^{N_l+M_l}$ such that we have

$$\Omega = \cup_{i=1}^{N_l+M_l} \tilde{\Omega}_i^l.$$

We can now make the following assumptions similar to Assumption 4.2.1.

Assumption 4.3.1 *Let us assume that*

- *These numbers h_l are bounded from above and below by const.q^l uniformly for all l.*
- *$(\partial\Omega_{l-1} \cap \partial\Omega_l) \setminus \partial\Omega = \emptyset$ for $l = 2, 3, \cdots, k$.*
- *$\partial\tilde{\Omega}_i^l$ aligns with boundaries of level l triangles, i.e. $\tilde{\Omega}_i^l$ is the union of level l triangles. Diameter$(\tilde{\Omega}_i^l) = O(h_{l-1})$.*
- *On each level, the subdomains $\{\tilde{\Omega}_i^l\}_{i=1}^{N_l+M_l}$ form a finite covering of Ω, with a covering constant N_c, i.e. we can color $\{\tilde{\Omega}_i^l\}_{i=1}^{N_l+M_l}$, using at most N_c colors in such a way that subdomains of the same color are disjoint.*
- *On each level, associated with $\{\tilde{\Omega}_i^l\}_{i=1}^{N_l}$, there exists a* partition of unity *$\{\theta_i^l\}$ satisfying*

$$\sum_i \theta_i^l = 1, \text{ with } \theta_i^l \in H_0^1(\tilde{\Omega}_i^l) \cap C^0(\tilde{\Omega}_i^l), 0 \le \theta_i^l \le 1 \text{ and } |\nabla\theta_i^l| \le C/h_{l-1}.$$

One way of constructing subdomains $\{\tilde{\Omega}_{i=1}^{N_l+M_l}\}, l = 2, \cdots, k$, with the above properties is mentioned in last section. Let $N_1 = 1$, $V_1^{h_1} = V^{h_1}$ and $V_i^{h_l} = V^{h_l} \cap H_0^1(\tilde{\Omega}_i^l)$ for $i = 1, \cdots, N_l + M_l$, $l = 2, \cdots, k$. The finite element space V^h is represented by

$$V^h = \sum_{l=1}^{k} V^{h_l} = \sum_{l=1}^{k} \sum_{i=1}^{N_l} V_i^{h_l}.$$

Let us define P_i^l as the orthogonal projection from V^h onto $V_i^{h_l}$ with respect to $a(\cdot,\cdot)$ which is the same as those in last section. The k-level additive Schwarz operator P is defined by

$$P = \sum_{l=1}^{k} \sum_{i=1}^{N_l} P_i^l. \tag{4.4}$$

Then we have the following algorithm.

Algorithm 4.3.1 (MAS with partial refinement) *Let P be the operator defined by (4.4). Apply the conjugate gradient method to the following symmetric and positive definite system*

$$Pu_h = g_h,$$

with respect to the inner product $a(\cdot,\cdot)$ for appropriate g_h such that the solution u_h is the same as that of (4.3).

Our main object of this section is to prove the following theorem.

Theorem 4.3.1 *For P defined by (4.4), the following inequalities hold*

$$C_1 a(u_h, u_h) \leq a(Pu_h, u_h) \leq C_2 a(u_h, u_h) \qquad \forall u_h \in V^h.$$

Thus $\kappa(P) \leq C_2 C_1^{-1}$. Here the constants C_1 and C_2 are independent of the mesh sizes $\{h_l\}$ and k.

In order to prove the above theorem, we need to introduce the concept of quasi-interpolants.

Definition 4.3.1 *Given a triangulation $\mathcal{T}$ of Ω, we associate with $\mathcal{T}$ a finite element subspace $V(\mathcal{T})$ of $L^2(\Omega)$ which consists of piecewise polynomials of degree less than or equal to m. A linear mapping*

$$Q : L^2(\Omega) \to V(\mathcal{T})$$

is called a quasi-interpolant of order m if it satisfies the properties

$$Qu = u \qquad \forall u \in V(\mathcal{T}),$$

and for a constant C depending only on the shape regularity such that

$$\|Qu\|_{L^2(K)} \leq C \cdot \|u\|_{L^2(\overline{K})}, \qquad \forall K \in \mathcal{T}, \qquad \forall u \in L^2(\Omega).$$

Here $\overline{K}$ denotes the union of the neighbouring elements of K.

The following example is given in Oswald [63].

Example. We construct a quasi-interpolant for linear elements in two dimension. The procedure can be generalized to the cases of more general Lagrange elements and higher-dimensional spaces. Consider an arbitrary nodal point P_i(=vertex) of $\mathcal{T}$ and its adjacent triangles and define on the region a piecewise linear and continuous function with value 3 at P_i and -1 at the other vertices of this region. Extend this function by zero to Ω and scale it by the factor $3/|A_i|$, where A_i denotes the support of this function. Denote the nodal basis function corresponding to P_i by ϕ_i and this L^∞ function by ϕ_i^*. Now we may take

$$Qu = \sum_i \int_\Omega (u, \phi_i^*)_{L^2(\Omega)} \cdot \phi_i.$$

If we look at the quadrature rule for triangles which uses the side midpoints as integration points and is exact for polynomials of degree 2, it is easy to see that Q satisfies the first condition in Definition 4.3.1. There remains to verify the second condition. For any element K with area $|K|$, we have $|A_i| \geq |K|$, for each of its three vertices P_i, and

$$|(u, \phi_i^*)_{L^2(\Omega)}| \leq \|u\|_{L^2(A_i)} \cdot \|\phi_i^*\|_{L^2(A_i)} \leq C \cdot |A_i|^{-1/2} \cdot \|u\|_{L^2(\overline{K})}$$

for the corresponding coefficients. Therefore

$$\|Qu\|_{L^2(K)} \leq |K|^{1/2} \cdot \max |(u, \phi_i^*)_{L^2(\Omega)}| \leq C \cdot \|u\|_{L^2(\overline{K})}.$$

Let $\tilde{Q}_l$ be a quasi-interpolant of order 1 from $L^2(\Omega)$ onto the space $\tilde{V}^{h_l}$, for $l = 1, 2, \cdots, k$. The proof of the following lemma is similar to that in Xu [83] after replacing the L^2 projection by quasi-interpolants. However, in our proof, we do not need to use the fact that the quasi-interpolants are bounded linear mappings in the space $H_0^1(\Omega)$.

Lemma 4.3.1 *There exists a constant C, which depends only on the shape regularity, such that*

$$\sum_{l=2}^{k} \|(\tilde{Q}_l - \tilde{Q}_{l-1})u\|^2_{L^2(\Omega)} \cdot \frac{1}{h_l^2} \leq C|u|^2_{H^1(\Omega)} \qquad \forall u \in H_0^1(\Omega).$$

Proof. Let us set

$$\hat{Q}_l = \tilde{Q}_l - \tilde{Q}_{l-1}$$

and

$$u_i = (\tilde{P}_i - \tilde{P}_{i-1})u.$$

Here $\tilde{P}_i$ is the orthogonal projection onto $\tilde{V}^{h_i}$ with respect to $a(\cdot,\cdot)$. We observe that

$$\|\hat{Q}_l u_i\|_{L^2(\Omega)} \leq C\|u_i\|_{L^2(\Omega)}$$

by the shape regularity assumption and that

$$\|\hat{Q}_l u_i\|_{L^2(\Omega)} = \|\hat{Q}_l(u_i - Q_{l-1}u_i)\|_{L^2(\Omega)} \leq C\|u_i - Q_{l-1}u_i\|_{L^2(\Omega)} \leq Ch_l\|u_i\|_{H^1(\Omega)}.$$

Here Q_l is the L^2 projection onto the space $\tilde{V}^{h_l}$. By the interpolation theorem of Chapter 1, we have

$$\|\hat{Q}_l u_i\|_{L^2(\Omega)} \leq Ch_l^{1-s} \cdot \|u_i\|_{H^{1-s}(\Omega)} \qquad \forall s \in (0,1).$$

We choose s as in Lemma 3.3.3. Then

$$\|\hat{Q}_l u_i\|_{L^2(\Omega)} \leq Ch_l^{1-s}\|u_i\|_{H^{1-s}(\Omega)} \leq Ch_l^{1-s}h_i^s\|u_i\|_{H^1(\Omega)}.$$

With $i \wedge j = \min(i,j)$, we have

$$\sum_{l=1}^{k}\|\hat{Q}_l u\|^2_{L^2(\Omega)} \cdot \frac{1}{h_l^2} = \sum_{l=1}^{k}\sum_{i,j=l}^{k}(\hat{Q}_l u_i, \hat{Q}_l u_j)_{L^2(\Omega)} \cdot \frac{1}{h_l^2} = \sum_{i,j=1}^{k}\sum_{l=1}^{i\wedge j}(\hat{Q}_l u_i, \hat{Q}_l u_j)_{L^2(\Omega)} \cdot \frac{1}{h_l^2}$$

$$\leq C\sum_{i,j=1}^{k}\sum_{l=1}^{i\wedge j}\frac{1}{h_l^2}h_l^{2(1-s)}\|u_i\|_{H^{1-s}(\Omega)}\|u_j\|_{H^{1-s}(\Omega)} \leq C\sum_{i,j=1}^{k}\sum_{l=1}^{i\wedge j}\frac{1}{h_l^2}h_l^{2(1-s)}h_i^s h_j^s\|u_i\|_{H^1(\Omega)}\|u_j\|_{H^1(\Omega)}$$

$$= C\sum_{i,j=1}^{k}h_i^s h_j^s\|u_i\|_{H^1(\Omega)}\|u_j\|_{H^1(\Omega)}\sum_{l=1}^{i\wedge j}h_l^{-2s} \leq C\sum_{i,j=1}^{k}h_{i\wedge j}^{-2s}h_i^s h_j^s\|u_i\|_{H^1(\Omega)}\|u_j\|_{H^1(\Omega)}$$

$$\leq C\sum_{i,j=1}^{k}q^{s|i-j|}\|u_i\|_{H^1(\Omega)}\|u_j\|_{H^1(\Omega)} \leq C\sum_{i=1}^{k}\|u_i\|^2_{H^1(\Omega)} \leq C(1+C'(d))|u|^2_{H^1(\Omega)}.$$

The last step follows from Friedrichs' inequality. $C'(d)$ is a constant which only depends upon the diameter d of the domain Ω. By using a simple dilation argument, we can completely remove the dependence of the above constant upon the diameter of Ω and complete the proof. $\Box$

Proofs of the boundness of the L^2 projection in $H_0^1(\Omega)$ are given in Scott and Zhang [69], and in Bramble and Xu [18]. However, the proof in [18] does not seem to apply to the elements which have nonempty intersection with $\partial\Omega$. We have the following similar result for quasi-interpolants.

Lemma 4.3.2 *There exists a constant C, which depends only on the shape regularity, such that*

$$|\tilde{Q}_l u|_{H^1(\Omega)} \leq C|u|_{H^1(\Omega)}.$$

Proof. Let us first consider the elements K satisfying $\overline{K} \cap \partial\Omega = \emptyset$. Then we have

$$|\tilde{Q}_l u|_{H^1(K)} = |\tilde{Q}_l u - c|_{H^1(K)} = |\tilde{Q}_l(u-c)|_{H^1(K)}$$
$$\leq Ch_l^{-1}\|\tilde{Q}_l(u-c)\|_{L^2(K)} \leq Ch_l^{-1}\|u-c\|_{L^2(\overline{K})}$$

for any constant c. Here $\overline{K}$ is the union of the neighbouring elements of K.

Take c to achieve the infinum. By Lemma 1.2.2, we have

$$|\tilde{Q}_l u|_{H^1(K)} \leq Ch_l^{-1} \cdot \inf_c \|u-c\|_{L^2(\overline{K})} \leq Ch_l^{-1} \cdot C'h_l|u|_{H^1(\overline{K})} \equiv C|u|_{H^1(\overline{K})}.$$

Let Ω_0 denote the union of such elements K. By shape regularity, we obtain

$$|\tilde{Q}_l u|_{H^1(\Omega_0)} \leq C|u|_{H^1(\Omega)}$$

for some constant C. Now it is sufficient to prove that

$$|\tilde{Q}_l u|_{H^1(\Omega\setminus\Omega_0)} \leq C|u|_{H^1(\Omega)}.$$

It is obvious that we can write $\overline{\Omega} \setminus \Omega_0 = \cup_{i=1}^{N_0} \overline{\Omega}_{0i}$ as a nonoverlapping union. Let $\Omega'_{0i} = \cup_{K\in\Omega_{0i}} \overline{K}$ where $\overline{K}$ is the union of the neighbouring elements of K. It is obvious that each region Ω'_{0i} and each function $u \in H_0^1(\Omega)$ satisfy the conditions of Lemma 1.2.5 and that the constant of this lemma is $O(h_l)$. Then we have

$$|\tilde{Q}_l u|_{H^1(\Omega_{0i})} \leq Ch_l^{-1}\|\tilde{Q}_l u\|_{L^2(\Omega_{0i})}$$
$$\leq Ch_l^{-1}\|u\|_{L^2(\Omega'_{0i})} \leq Ch_l^{-1} \cdot h_l|u|_{H^1(\Omega'_{0i})} = C|u|_{H^1(\Omega'_{0i})}.$$

By combining the above results, the proof of the lemma follows easily. □

In order to prove the main theorem of this section, we need the following two lemmas. In these lemmas, $\tilde{I}_l$ denotes the usual nodal interpolants from $\tilde{V}^{h_k}$ onto $\tilde{V}^{h_l}$. The proof of the first lemma can be found in [11] and that of the second lemma can be established by replacing the L^2 norm with a discrete equivalent norm in the finite element space $\tilde{V}^{h_l}$.

Lemma 4.3.3 *Assume that $i \leq j$. Then there exists a constant C, which depends only on the shape regularity, such that*

$$|\tilde{I}_i u_j|^2_{H^1(\Omega)} \leq C\log(h_i/h_j) \cdot |u_j|^2_{H^1(\Omega)}, \qquad \forall u_j \in \tilde{V}^{h_j}$$

in two-dimensional space and

$$|\tilde{I}_i u_j|^2_{H^1(\Omega)} \leq C(h_i/h_j)^{n-2} \cdot |u_j|^2_{H^1(\Omega)}, \qquad \forall u_j \in \tilde{V}^{h_j}$$

in higher dimensional space.

Lemma 4.3.4 *Assume that $i \leq j$. Then there exists a constant C, which depends only on the shape regularity, such that*

$$\|\tilde{I}_i u_j\|^2_{L^2(\Omega)} \leq C(h_i/h_j)^{n} \cdot \|u_j\|^2_{L^2(\Omega)}, \qquad \forall u_j \in \tilde{V}^{h_j}.$$

Proof of Theorem 4.3.1. Let us first prove that the operator P has an uniform upper bound. We define $\tilde{P}$ by

$$\tilde{P} = \sum_{l=1}^{k} \sum_{i=1}^{N_l+M_l} P_i^l.$$

We observe that $\tilde{P}$ has a uniform upper bound by Theorem 4.2.2 and $P \leq \tilde{P}$. Therefore P has a uniform upper bound.

To establish the uniform lower bound, we will apply the first part of Lemma 2.1.2. We note that it is sufficient to find a good decomposition of $u \in V^h$ such that the constant is uniformly bounded from above. Let us first decompose u as

$$u = \tilde{Q}_1 + \sum_{l=2}^{k} (\tilde{Q}_l - \tilde{Q}_{l-1})u \equiv \sum_{l=1}^{k} v_l.$$

Although $(\tilde{Q}_l - \tilde{Q}_{l-1})u$ is not in the space $\tilde{V}^{h_l}$, since the nodal values of these functions on $\partial\Omega_l$ are not zero in general, these functions nevertheless have their support in Ω_{l-1} by the assumption $(\partial\Omega_l \cap \partial\Omega_{l-1}) \setminus \partial\Omega = \emptyset$. If we set $u_l = v_l - \tilde{I}_{l-1} v_l + \tilde{I}_l v_{l+1}$ for $l = 2, \cdots, k$ and $u_1 = v_1 + \tilde{I}_1 v_2$, then $u_l \in V^{h_l}$ for all l and $u = \sum_{l=1}^{k} u_i$. By Lemma 4.3.1 and Lemma 4.3.2, we have

$$|v_1|^2_{H^1(\Omega)} + \sum_{l=2}^{k} \|v_l\|^2_{L^2(\Omega)} \cdot \frac{1}{h_l^2} \leq C|u|^2_{H^1(\Omega)}.$$

If we apply Lemmas 4.3.3 and 4.3.4 to the above equation, we will find that

$$|u_1|^2_{H^1(\Omega)} + \sum_{l=2}^{k} \|u_l\|^2_{L^2(\Omega)} \cdot \frac{1}{h_l^2} \leq C|u|^2_{H^1(\Omega)}. \tag{4.5}$$

We need to further decompose u^l, for $l \geq 2$, as

$$u^l = \sum_{i=1}^{N_l} u_i^l, \text{ with } u_i^l \equiv \tilde{I}_l(\theta_i^l u^l) \in V_i^{h_l}.$$

Here $\{\theta_i^l\}$ is a partition of unity as in Assumption 4.3.1. It can be shown that

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

Summing over i and using the finite covering property of $\{\tilde{\Omega}_i^l\}$, we obtain

$$\begin{aligned} \sum_i |u_i^l|^2_{H^1(\Omega)} = \sum_i |u_i^l|^2_{H^1(\tilde{\Omega}_i^l)} &\leq C\sum_i(|u^l|^2_{H^1(\tilde{\Omega}_i^l)} + \frac{1}{h_{l-1}^2}\cdot\|u^l\|^2_{L^2(\tilde{\Omega}_i^l)}) \\ &\leq C(|u^l|^2_{H^1(\Omega)} + \frac{1}{h_l^2}\|u^l\|^2_{L^2(\Omega)}) \leq C\frac{1}{h_l^2}\|u^l\|^2_{L^2(\Omega)}. \end{aligned}$$

Summing over l, for $1 \leq l \leq k$, and using inequality (4.5), we get

$$\sum_{l=1}^k \sum_i |u_i^l|^2_{H^1(\Omega)} \leq C(|u^l|^2_{H^1(\Omega)} + \sum_{l=2}^k \frac{1}{h_l^2}\|u^l\|^2_{L^2(\Omega)}) \leq C|u|^2_{H^1(\Omega)}.$$

The lower bound of P now follows. □

We end this section by mentioning a special decomposition of the domain Ω in Zhang [86]. It is called the multilevel diagonal scaling. Let ϕ_i^l be a nodal basis function of V^{h_l}, and associate with each ϕ_i^l the subdomain $\tilde{\Omega}_i^l = supp\{\phi_i^l\}$. We may choose $V_i^l = span\{\phi_i^l\} = V^{h_l} \cap H_0^1(\tilde{\Omega}_i^l)$ and obtain the decomposition

$$V^{h_k} = \sum_{l=1}^k \sum_{i=1}^{N_l} V_i^{h_l}$$

and the Galerkin projection P_i^l corresponding to $V_i^{h_l}$. Let $P' = \sum_{l=1}^k \sum_{i=1}^{N_l} P_i^l$. It is easy to see that the above construction satisfies Assumption 4.3.1. Therefore we have another variant of Algorithm 4.3.1 whose optimality follows from Theorem 4.3.1.

Algorithm 4.3.2 (MDS with partial refinement) *Let P' be the operator defined above. Apply the conjugate gradient method to the following symmetric and positive definite system*

$$P'u_h = g_h,$$

with respect to the inner product $a(\cdot,\cdot)$ for an appropriate g_h such that the solution u_h is the same as that of (4.3).

Let K_l be the stiffness matrix associated with V^{h_l}, let K_h be the stiffness matrix associated with V^h and let $D_l = diag(K_l)$. Let $\mathcal{I}_l : V^{h_l} \to V^h$, $1 \leq l \leq k$, be the standard inclusion operator, and let $\mathcal{I}_l^t : V^h \to V^{h_l}$ be an operator related to $\mathcal{I}_l$ in the following way:

$$(\mathcal{I}_l^t u_h, v^l)_l = (u, \mathcal{I}_l v^l)_{L^2(\Omega)} \qquad \forall v^l \in V^{h_l}.$$

Here $(\cdot,\cdot)_l$ is the discrete inner product in V^{h_l}, which is equivalent to $L^2(\Omega)$, defined by

$$(u^l, v^l)_l = h_l^n \sum_{x \in \mathcal{N}_l} u^l(x) v^l(x) \qquad \forall u^l, v^l \in V^{h_l}.$$

Here $\mathcal{N}_l$ is the set of nodes of the degrees of freedom in V^{h_l}. Algorithm 4.3.2 can then be written as: Find the solution of $K_h x = b$ by solving the preconditioned system

$$B_h^{-1} K_h x = B_h^{-1} b,$$

where

$$B_h^{-1} = h_1^n \cdot \mathcal{I}_1 K_1^{-1} \mathcal{I}_1^t + \cdots + h_{k-1}^n \cdot \mathcal{I}_{k-1} D_{k-1}^{-1} \mathcal{I}_{k-1}^t + h_k^n \cdot D_k^{-1}.$$

We remark that if we replace the matrices D_l by identity matrices, we obtain the BPX algorithm with partial refinement.

4.4 Numerical Results

In this section, we report on some numerical experiments to verify our theoretical results. We take the differential operator to be the Laplacian and the subdomains are triangulated using right triangles of equal size. In these experiments, we choose arbitrary right hand sides and the iterations are stopped when the residual size with respect to the energy norm has been reduced by a factor 10^{-6}. In all experiments, we choose $h_{l+1} = h_l/2$ and the ratio of the diameters of Ω_{l+1} and Ω_l to be $1/2$ for $l = 1, 2, \cdots, k-1$ as shown in Fig. 3.1 and Fig 3.2. We will call the problem of Fig. 3.1 Model Problem 1 and that of Fig. 3.2 Model Problem 2.

In Table 4.1 and 4.2, we show results for the MAS method with partial refinement. To each node x_i in $\mathcal{N}_{l-1}$, for $l \geq 2$, we associate a subdomain $\tilde{\Omega}_i^l$ which is a rectangle with center at x_i and side length $2h_{l-1}$. It is easy to see that this construction of $\{\tilde{\Omega}_i^l\}$ satisfies Assumption 4.3.1.

$h_1 = 1/8$									
k	2	3	4	5	6	7	8	9	10
min. ev.	0.6660	0.6615	0.6589	0.6576	0.6570	0.6567	0.6564	0.6562	0.6561
max. ev.	4.462	6.329	6.623	6.717	6.843	6.983	7.100	7.194	7.270
cond. no.	6.700	9.568	10.05	10.21	10.42	10.63	10.82	10.96	11.08
no. of iter.	16	20	21	22	22	22	22	22	22
$h_1 = 1/16$									
k	2	3	4	5	6	7	8	9	10
min. ev.	0.6675	0.6645	0.6639	0.6635	0.6633	0.6632	0.6632	0.6631	0.6630
max. ev.	4.529	6.756	8.400	8.635	8.683	8.758	8.836	8.903	8.958
cond. no.	6.785	10.17	12.65	13.01	13.09	13.21	13.32	13.43	13.51
no. of iter.	16	21	23	24	24	24	24	25	25

Table 4.1: MAS with partial refinement for Model problem 1

In Table 4.3 and 4.4, we consider the method of MDS with partial refinement. We see that the condition numbers of Table 4.3 and 4.4 are usually larger than those of Table 4.1 and 4.2 because the overlap between the subdomains in MDS is smaller than those in MAS and that more subspaces are used in MDS than in MAS.

$h_1 = 1/8$									
k	2	3	4	5	6	7	8	9	10
min. ev.	0.6685	0.6618	0.6589	0.6577	0.6571	0.6567	0.6564	0.6562	0.6560
max. ev.	4.481	6.335	7.301	7.757	7.943	8.032	8.080	8.109	8.128
cond. no.	6.703	9.572	11.08	11.79	12.09	12.23	12.31	12.36	12.39
no. of iter.	16	20	21	21	22	22	23	23	23
$h_1 = 1/16$									
k	2	3	4	5	6	7	8	9	10
min. ev.	0.6703	0.6650	0.6642	0.6637	0.6635	0.6634	0.6633	0.6633	0.6633
max. ev.	4.525	6.818	8.412	9.128	9.459	9.599	9.667	9.725	9.754
cond. no.	6.751	10.25	12.66	13.75	14.26	14.47	14.57	14.66	14.71
no. of iter.	16	21	23	24	25	25	25	25	25

Table 4.2: MAS with partial refinement for Model problem 2

$h_1 = 1/8$									
k	2	3	4	5	6	7	8	9	10
min. ev.	0.7830	0.7836	0.7833	0.7835	0.7848	0.7834	0.7842	0.7844	0.7849
max. ev.	7.815	11.26	12.63	13.10	13.50	13.91	14.24	14.50	14.70
cond. no.	9.981	14.37	16.12	16.72	17.20	17.76	18.16	18.49	18.73
no. of iter.	21	25	27	28	28	29	29	29	29
$h_1 = 1/16$									
k	2	3	4	5	6	7	8	9	10
min. ev.	0.7880	0.7890	0.7891	0.7894	0.7914	0.7924	0.7933	0.7937	0.7940
max. ev.	9.903	11.78	13.84	14.92	15.26	15.55	15.87	16.13	16.35
cond. no.	10.03	14.93	17.54	18.90	19.28	19.62	20.01	20.32	20.59
no. of iter.	21	26	28	29	29	29	29	29	29

Table 4.3: MDS with partial refinement for Model problem 1

$h_1 = 1/8$									
k	2	3	4	5	6	7	8	9	10
min. ev.	0.7852	0.7846	0.7847	0.7851	0.7872	0.7885	0.7872	0.7880	0.7883
max. ev.	7.855	11.40	12.95	13.72	14.01	14.16	14.24	14.30	14.43
cond. no.	10.00	14.53	16.50	17.48	17.80	17.96	18.09	18.15	18.30
no. of iter.	21	25	27	28	28	28	29	29	29
$h_1 = 1/16$									
k	2	3	4	5	6	7	8	9	10
min. ev.	0.7884	0.7889	0.7898	0.7908	0.7923	0.7940	0.7949	0.7955	0.7959
max. ev.	7.949	11.84	13.96	15.20	15.82	16.05	16.17	16.24	16.30
cond. no.	10.08	15.01	17.68	19.22	19.97	20.21	20.34	20.41	20.48
no. of iter.	21	26	28	29	29	29	29	29	29

Table 4.4: MDS with partial refinement for Model problem 2

Chapter 5

Domain Decomposition Parallel Preconditioning Methods

5.1 Introduction

The substructuring technique has been developed into a standard method of solving finite element equations practically. Recently, this technique has been used to construct efficient parallel preconditioner for preconditioned conjugate gradient method which permits a high level of parellelization on multiprocessor computer systems. Here we consider the so called Dirichlet DD (Domain Decomposition) preconditioner. There are many papers devoted to this topic; cf. e.g. [6] [11,12,13,14] [7] [25] [80,32] and the references therein.

In this chapter, we concentrate on studying domain decomposition parallel preconditioning method (DD-PPM) for the Dirichlet elliptic problem (1.1)-(1.3) using iterative solvers for the subproblems in each subdomain. We also make some remarks on the general mixed-type boundary value problem. The subproblems used in this method are the finite element approximation of the elliptic problem on separate subdomains with homogeneous Dirichlet boundary conditions. In addition, we need a good Schur complement preconditioner such as the one proposed by Dryja [25] for a special case and by Bramble [11,12,13,14] and Smith [73,72] in more general situations. Börgers [7] first analyzed the convergence of the Neumann-Dirichlet and the Dirichlet DD-PPM method for solving plane second-order elliptic boundary value problems with inexact solvers on the subdomains and obtained the following estimate

$$\kappa(C^{-1}K) \leq c_1(1+\sqrt{\lambda-1})^2(1-\sqrt{\lambda-1})^{-2}$$

of the relative spectral condition number $\kappa(C^{-1}K)$ under the rather restrictive condition that the spectral radius $\lambda = \rho(S_B^{-1}R_B)$ of the matrix $S_B^{-1}R_B$, cf. Section 5.3, is less than 2. Then Hasse, Langer and Meyer in [38] improved Börger's estimate without making any additional assumption. We will establish a shaper estimate than the previous using the additive Schwarz framework under the same assumptions. A consequence of Börger's estimate is that an h-independent bound of $\kappa(C^{-1}K)$ can be obtained if

$$\eta^k \leq const.h^{0.5}.$$

Here η is the rate of convergence of iterative methods used for the subproblems, which we assume are energy-symmetric. k is the number of inner iterations applied per conjugate gradient iteration step, h the usual discretization parameter and C is the DD preconditioner. Our main idea is that we should use a uniformly bounded extension from the boundary of the subdomains to the interiors to obtain an h-independent bound for $\kappa(C^{-1}K)$. We remark that our results hold for general self-adjoint coercive elliptic operators with uniformly bounded coefficients.

5.2 Description of a General DD-PPM

Let us consider the system of linear algebraic equations

$$\begin{bmatrix} K_{II} & K_{IB} \\ K_{BI} & K_{BB} \end{bmatrix} \begin{bmatrix} u_I \\ u_B \end{bmatrix} = \begin{bmatrix} f_I \\ f_B \end{bmatrix} \tag{5.1}$$

resulting from the DD finite element discretization of the elliptic problem (1.1)-(1.3) with $\Gamma_D \neq \emptyset$. Here the indices I and B correspond to the interior and the boundary nodes of the subdomains, respectively. The system matrix K of (5.1) is symmetric and positive definite. If the domain Ω, on which the boundary value problem is given, is decomposed into p substructures, i.e. $\overline{\Omega} = \cup_{i=1}^{p} \overline{\Omega_i}$, $\Omega_i \cap \Omega_j = \emptyset$ for $i \neq j$, then the matrix K_{II} has block diagonal form

$$K_{II} = diag[K_{II}^{(i)}]_{i=1,\ldots,p} \tag{5.2}$$

where the block diagonal entries $K_{II}^{(i)}$ are the stiffness matrices arising from the finite element approximation to the elliptic problems on Ω_i with Dirichlet boundary conditions on $\partial\Omega_i$. We suppose that the interior nodes of the subdomains are numbered first in Ω_1, followed by those in $\Omega_2, \Omega_3, \cdots, \Omega_p$.

In the classical superelement technique, one eliminates the inner unknowns u_I, obtaining a reduced system $S_B u_B = g_B$, with the Schur complement $S_B = K_{BB} - K_{BI}K_{II}^{-1}K_{IB}$ and the vector $g_B = f_B - K_{BI}K_{II}^{-1}f_I$ as system matrix and right hand side, respectively. An alternative approach to solving of (5.1) is the application of the preconditioned conjugate gradient method with a symmetric and positive definite preconditioner C satisfying the following three conditions:

1) C is spectrally equivalent to K, i.e.

$$\underline{\gamma} C \leq K \leq \overline{\gamma} C \tag{5.3}$$

with positive spectral equivalence constants $\underline{\gamma}$ and $\overline{\gamma}$ such that the relative spectral condition number is as small as possible. Here the inequalities in (5.3) are to be understood in the sense of the Euclidean inner product.

2) There are fast methods for solving the preconditioning system

$$Cw = r, \tag{5.4}$$

which must be solved in every iteration step of the PCG-method.

3) The solution method for (5.4) should permit a high level of parallelism.

A good choice for such a preconditioner is the so-called DD-preconditioner

$$C = \begin{bmatrix} I_I & 0 \\ K_{BI}B_{II}^{-1} & I_B \end{bmatrix} \begin{bmatrix} \tilde{K}_{II} & 0 \\ 0 & \tilde{S}_B \end{bmatrix} \begin{bmatrix} I_I & B_{II}^{-1}K_{IB} \\ 0 & I_B \end{bmatrix}, \tag{5.5}$$

which is derived from the well-known block LDU decomposition $L_K D_K L_K^T$ of the symmetric matrix K. We have replaced the K_{II} in L_K by a not necessarily symmetric and positive definite B_{II} and replaced the K_{II} and S_B in D_K by suitable symmetric and positive definite preconditioner $\tilde{K}_{II}$ and $\tilde{S}_B$, respectively.

For a discussion of the parallel implementation of the preconditioned conjugate gradient algorithm with the DD preconditioner (5.5) for solving (5.1) which is well suited for parallel computation on Multiple Instruction Multiple data (MIMD) computers with local memory and message passing, we refer to [38].

5.3 Abstract Spectral Equivalence Bound Estimates

In this section, we will derive sharper spectral equivalence bound estimates. Let us denote the finite element space by V, then obviously $V = V_I \oplus V_B$. Here V_I is the subspace of

V which has zero nodal values on the boundaries of these subdomains $\Omega_1, \Omega_2, \cdots, \Omega_p$ and V_B is the subspace of V with $\underline{u}_I = -B_{II}^{-1} K_{IB} \underline{u}_B$, i.e. V_B is a subspace of approximately discrete harmonic finite element functions in the interior of these subdomains. We denote the energy inner-product on V by $a(\cdot, \cdot)$, where $a(\cdot, \cdot)$ is given in equation (1.4). In order to derive our bounds, we need several lemmas.

Lemma 5.3.1 *Let $u_1 \in V_I$, $u_2 \in V_B$, and $u = u_1 + u_2$. Let c_1 and c_2 be strictly positive constants. Then*

$$c_1 a(u_1, u_1) + c_2 a(u_2, u_2) \leq c a(u, u), \qquad \forall u \in V$$

if and only if

$$c \geq \frac{c_1 + c_2 + \sqrt{(c_1 + c_2)^2 - 4(1 - \eta^2) c_1 c_2}}{2(1 - \eta^2)}$$

where $\eta = cos\angle(V_I, V_B) < 1$. Conversely,

$$c_1 a(u_1, u_1) + c_2 a(u_2, u_2) \geq c a(u, u), \qquad \forall u \in V$$

if and only if

$$c \leq \frac{2 c_1 c_2}{c_1 + c_2 + \sqrt{(c_1 + c_2)^2 - 4(1 - \eta^2) c_1 c_2}}$$

Proof. Obviously we have $c \geq \max(c_1, c_2)$ from the first assumption. Simplifying the inequality, we get

$$(c - c_1) a(u_1, u_1) + 2c a(u_1, u_2) + (c - c_2) a(u_2, u_2) \geq 0$$

This inequality holds if and only if

$$(c - c_1) a(u_1, u_1) - 2c\eta \sqrt{a(u_1, u_1)} \sqrt{a(u_2, u_2)} + (c - c_2) a(u_2, u_2) \geq 0$$

Set $x = \sqrt{a(u_1, u_1)}$, $y = \sqrt{a(u_2, u_2)}$. We have

$$(c - c_1) x^2 - 2c\eta xy + (c - c_2) y^2 \geq 0 \qquad \forall x, y \geq 0$$

This condition is equivalent to

$$4c^2 \eta^2 - 4(c - c_1)(c - c_2) \leq 0$$

Finally, we get

$$c \geq \frac{c_1 + c_2 + \sqrt{(c_1 + c_2)^2 - 4(1 - \eta^2) c_1 c_2}}{2(1 - \eta^2)}.$$

Conversely, we also have $c \leq \min(c_1, c_2)$ from the second assumption. Simplifying this inequality, we get

$$(c_1 - c)a(u_1, u_1) - 2ca(u_1, u_2) + (c_2 - c)a(u_2, u_2) \geq 0.$$

This is also equivalent to

$$(c_1 - c)a(u_1, u_1) - 2c\eta\sqrt{a(u_1, u_1)}\sqrt{a(u_2, u_2)} + (c_2 - c)a(u_2, u_2) \geq 0.$$

So

$$4c^2\eta^2 - 4(c_1 - c)(c_2 - c) \leq 0.$$

We can now conclude that

$$c \leq \frac{c_1 + c_2 - \sqrt{(c_1 + c_2)^2 - 4(1 - \eta^2)c_1c_2}}{2(1 - \eta^2)} \leq \frac{2c_1c_2}{c_1 + c_2 + \sqrt{(c_1 + c_2)^2 - 4(1 - \eta^2)c_1c_2}}. \square$$

Lemma 5.3.2 *$\eta \equiv cos\angle(V_I, V_B) = \sqrt{1 - 1/\lambda}$ where $\lambda = \rho(S_B^{-1}R_B)$ is the spectral radius of $S_B^{-1}R_B$, $S_B = K_{BB} - K_{BI}K_{II}^{-1}K_{IB}$ and*

$$R_B = \begin{bmatrix} -B_{II}^{-1}K_{IB} \\ I_B \end{bmatrix}^T K \begin{bmatrix} -B_{II}^{-1}K_{IB} \\ I_B \end{bmatrix}.$$

Proof. Denote the largest singular value by $\sigma_{\max}$, the smallest eigenvalue by $\lambda_{\min}$ and the spectral radius by ρ. Take a function in V_I with a matrix representation $\begin{bmatrix} I_I \\ 0 \end{bmatrix} \underline{u}_I$, and a function in V_B with a matrix representation $\begin{bmatrix} -B_{II}^{-1}K_{IB} \\ I_B \end{bmatrix} \underline{u}_B$. Then

$$\begin{aligned} cos\angle(V_I, V_B) &= \sup_{\underline{u}_I, \underline{u}_B} \frac{\underline{u}_I^T \begin{bmatrix} I_I & 0 \end{bmatrix} K \begin{bmatrix} -B_{II}^{-1}K_{IB} \\ I_B \end{bmatrix} \underline{u}_B}{(\underline{u}_I^T K_{II} \underline{u}_I)^{1/2} (\underline{u}_B^T R_B \underline{u}_B)^{1/2}} \\ &= \sup_{\underline{v}_I, \underline{v}_B} \frac{\underline{v}_I^T K_{II}^{-1/2} \begin{bmatrix} I_I & 0 \end{bmatrix} K \begin{bmatrix} -B_{II}^{-1}K_{IB} \\ I_B \end{bmatrix} R_B^{-1/2} \underline{v}_B}{(\underline{v}_B^T \underline{v}_B)^{1/2} (\underline{v}_I^T \underline{v}_I)^{1/2}} \\ &= \sigma_{\max}(K_{II}^{-1/2} \begin{bmatrix} I_I & 0 \end{bmatrix} K \begin{bmatrix} -B_{II}^{-1}K_{IB} \\ I_B \end{bmatrix} R_B^{-1/2}) \\ &= \sqrt{\rho(R_B^{-1/2} \begin{bmatrix} -B_{II}^{-1}K_{IB} \\ I_B \end{bmatrix}^T K \begin{bmatrix} I_I \\ 0 \end{bmatrix} K_{II}^{-1} \begin{bmatrix} I_I & 0 \end{bmatrix} K \begin{bmatrix} -B_{II}^{-1}K_{IB} \\ I_B \end{bmatrix} R_B^{-1/2})} \\ &= \sqrt{\rho(R_B^{-1/2} \begin{bmatrix} -B_{II}^{-1}K_{IB} \\ I_B \end{bmatrix}^T \begin{bmatrix} K_{II} & K_{IB} \\ K_{BI} & K_{BI}K_{II}^{-1}K_{IB} \end{bmatrix} \begin{bmatrix} -B_{II}^{-1}K_{IB} \\ I_B \end{bmatrix} R_B^{-1/2})} \end{aligned}$$

$$= \sqrt{\rho(R_B^{-1/2} \begin{bmatrix} -B_{II}^{-1}K_{IB} \\ I_B \end{bmatrix}^T (K + \begin{bmatrix} 0 & 0 \\ 0 & -S_B \end{bmatrix}) \begin{bmatrix} -B_{II}^{-1}K_{IB} \\ I_B \end{bmatrix} R_B^{-1/2})}$$

$$= \sqrt{\rho(R_B^{-1/2}(R_B - S_B)R_B^{-1/2})} = \sqrt{\rho(I_B - R_B^{-1/2}S_B R_B^{-1/2})}$$

$$= \sqrt{1 - \lambda_{\min}(R_B^{-1}S_B)} = \sqrt{1 - \rho(S_B^{-1}R_B)^{-1}} = \sqrt{1 - \frac{1}{\lambda}}. \qquad \square$$

We need another variant of Lemma 5.3.2 for the main theorem of this section. Its proof is similar to that of Lemma 5.3.2.

Lemma 5.3.3 *Let $u_1 \in V_I$, $u_2 \in V_B$, and u_2 have the matrix representation*

$$\begin{bmatrix} -B_{II}^{-1}K_{IB} \\ I_B \end{bmatrix} \underline{u}_B.$$

Then

$$|a(u_1, u_2)| \leq \sqrt{1 - 1/\lambda_{\underline{u}_B}} \cdot a(u_1, u_1)^{1/2} a(u_2, u_2)^{1/2} \equiv \eta_{\underline{u}_B} \cdot a(u_1, u_1)^{1/2} a(u_2, u_2)^{1/2}$$

where

$$\lambda_{\underline{u}_B} = \frac{\underline{u}_B^T R_B \underline{u}_B}{\underline{u}_B^T S_B \underline{u}_B}$$

and R_B is the same as in Lemma 5.3.2.

For convenience, we will denote the finite element function in V, with vector representation $\underline{u}$, by $\Phi\underline{u}$. Let us assume that $\tilde{K}_{II}$ and $\tilde{S}_B$ are spectrally equivalent to K_{II} and S_B, respectively, i.e. there exist constants $\underline{\gamma}_I$, $\overline{\gamma}_I$, $\underline{\gamma}_B$ and $\overline{\gamma}_B$ such that

$$\underline{\gamma}_I K_{II} \leq \tilde{K}_{II} \leq \overline{\gamma}_I K_{II}$$

$$\underline{\gamma}_B S_B \leq \tilde{S}_B \leq \overline{\gamma}_B S_B$$

We can write $C^{-1}K$ as

$$\begin{bmatrix} I_I \\ 0 \end{bmatrix} \tilde{K}_{II}^{-1} \begin{bmatrix} I_I & 0 \end{bmatrix} K + \begin{bmatrix} -B_{II}^{-1}K_{IB} \\ I_B \end{bmatrix} \tilde{S}_B^{-1} \begin{bmatrix} -B_{II}^{-1}K_{IB} \\ I_B \end{bmatrix}^T K.$$

Let T_I be the operator from V onto V_I which corresponds to the first term of the above expression and T_B be the operator from V onto V_B corresponding to the second term. Obviously T_I and T_B are energy-symmetric. We define inner-products $b_I(\cdot,\cdot)$ and $b_B(\cdot,\cdot)$ on $V_I \times V_I$ and $V_B \times V_B$, respectively, by

$$b_I(T_I u, \phi_1) = a(u, \phi_1), \qquad \forall \phi_1 \in V_I,$$

and

$$b_B(T_B u, \phi_2) = a(u, \phi_2), \qquad \forall \phi_2 \in V_B.$$

From the definitions of T_I and T_B, we see that

$$T_I(\Phi \begin{bmatrix} \underline{u}_I \\ 0 \end{bmatrix}) = \Phi \begin{bmatrix} \tilde{K}_{II}^{-1} K_{II} \underline{u}_I \\ 0 \end{bmatrix}$$

and

$$T_B(\Phi \begin{bmatrix} -B_{II}^{-1} K_{IB} \\ I_B \end{bmatrix} \underline{u}_B) = \Phi \begin{bmatrix} -B_{II}^{-1} K_{IB} \\ I_B \end{bmatrix} \tilde{S}_B^{-1} R_B \underline{u}_B,$$

where R_B is the same as in Lemma 5.3.2. Therefore

$$\begin{aligned} b_I(\Phi \begin{bmatrix} \underline{u}_I \\ 0 \end{bmatrix}, \Phi \begin{bmatrix} \underline{v}_I \\ 0 \end{bmatrix}) &= b_I(T_I(\Phi \begin{bmatrix} K_{II}^{-1} \tilde{K}_{II} \underline{u}_I \\ 0 \end{bmatrix}), \Phi \begin{bmatrix} \underline{v}_I \\ 0 \end{bmatrix}) \\ &= a(\Phi \begin{bmatrix} K_{II}^{-1} \tilde{K}_{II} \underline{\mathbf{u}}_I \\ 0 \end{bmatrix}, \Phi \begin{bmatrix} \underline{v}_I \\ 0 \end{bmatrix}) = \underline{v}_I^T \tilde{K}_{II} \underline{u}_I^T, \end{aligned}$$

and

$$\begin{aligned} & b_B(\Phi \begin{bmatrix} -B_{II}^{-1} K_{IB} \\ I_B \end{bmatrix} \underline{u}_B, \Phi \begin{bmatrix} -B_{II}^{-1} K_{IB} \\ I_B \end{bmatrix} \underline{v}_B) \\ &= b_B(T_B \Phi \begin{bmatrix} -B_{II}^{-1} K_{IB} \\ I_B \end{bmatrix} R_B^{-1} \tilde{S}_B \underline{u}_B, \Phi \begin{bmatrix} -B_{II}^{-1} K_{IB} \\ I_B \end{bmatrix} \underline{v}_B) \\ &= a(\Phi \begin{bmatrix} -B_{II}^{-1} K_{IB} \\ I_B \end{bmatrix} R_B^{-1} \tilde{S}_B \underline{u}_B, \Phi \begin{bmatrix} -B_{II}^{-1} K_{IB} \\ I_B \end{bmatrix} \underline{v}_B) \\ &= \underline{v}_B^T R_B R_B^{-1} \tilde{S}_B \underline{u}_B = \underline{v}_B^T \tilde{S}_B \underline{u}_B. \end{aligned}$$

Using these results, we can conclude that

$$\underline{\gamma}_I a(u_1, u_1) \le b_I(u_1, u_1) \le \overline{\gamma}_I a(u_1, u_1), \qquad \forall u_1 \in V_I,$$

and

$$\frac{\underline{\gamma}_B}{\lambda} a(u_2, u_2) \le b_B(u_2, u_2) \le \frac{\overline{\gamma}_B}{\lambda_{\underline{u}_B}} a(u_2, u_2), \qquad \forall u_2 \in V_B.$$

Here λ is as in Lemma 5.3.2 and $\lambda_{\underline{u}_B}$ is given in Lemma 5.3.3. We note that it is always true that

$$1 \le \lambda_{\underline{u}_B} \le \lambda.$$

Now we can state the main result of this section.

Theorem 5.3.1

$$\kappa(C^{-1}K) \le \frac{1}{4\underline{\gamma}_I \underline{\gamma}_B} [(\lambda \overline{\gamma}_I + \overline{\gamma}_B) + \sqrt{(\lambda \overline{\gamma}_I + \overline{\gamma}_B)^2 - 4\overline{\gamma}_I \overline{\gamma}_B}] [(\lambda \underline{\gamma}_I + \underline{\gamma}_B) + \sqrt{(\lambda \underline{\gamma}_I + \underline{\gamma}_B)^2 - 4\underline{\gamma}_I \underline{\gamma}_B}]$$

where λ is as in Lemma 5.3.2. We also have

$$\kappa(C^{-1}K) \geq \frac{1}{4\overline{\gamma}_I\overline{\gamma}_B\lambda}[(\overline{\gamma}_I+\overline{\gamma}_B)\lambda+\sqrt{(\overline{\gamma}_I+\overline{\gamma}_B)^2\lambda^2-4\lambda\overline{\gamma}_I\overline{\gamma}_B}][(\lambda\underline{\gamma}_I+\underline{\gamma}_B)+\sqrt{(\lambda\underline{\gamma}_I+\underline{\gamma}_B)^2-4\underline{\gamma}_I\underline{\gamma}_B}]$$

Therefore, $\kappa(C^{-1}K)$ is uniformly bounded if and only if λ is uniformly bounded.

Proof. Our main idea of proving this theorem is to apply Lemma 2.1.2 to the spaces V_I and V_B, and the inner products $b_I(\cdot,\cdot)$ and $b_B(\cdot,\cdot)$. To estimate the minimum eigenvalue of $T_I + T_B$, we observe that

$$\begin{aligned} b_I(u_1,u_1)+b_B(u_2,u_2) &\leq \overline{\gamma}_I a(u_1,u_1)+\overline{\gamma}_B/\lambda_{\underline{u}_B}a(u_2,u_2) \\ &\leq \frac{(\overline{\gamma}_I+\overline{\gamma}_B/\lambda_{\underline{u}_B})+\sqrt{(\overline{\gamma}_I+\overline{\gamma}_B/\lambda_{\underline{u}_B})^2-4(1-\eta^2_{\underline{u}_B})\overline{\gamma}_I\overline{\gamma}_B/\lambda_{\underline{u}_B}}}{2(1-\eta^2_{\underline{u}_B})}a(u,u) \\ &= \frac{(\lambda_{\underline{u}_B}\overline{\gamma}_I+\overline{\gamma}_B)+\sqrt{(\lambda_{\underline{u}_B}\overline{\gamma}_I+\overline{\gamma}_B)^2-4\overline{\gamma}_I\overline{\gamma}_B}}{2}a(u,u) \end{aligned}$$

$\forall u \in V$ and $u = u_1 + u_2$ by Lemma 5.3.1 and Lemma 5.3.3. Here $u_1 \in V_I$, $u_2 \in V_B$ and $\eta_{\underline{u}_B}$ is the same as in Lemma 5.3.3. Therefore we have

$$\lambda_{\min}(T_I+T_B)^{-1} \leq \frac{(\lambda\overline{\gamma}_I+\overline{\gamma}_B)+\sqrt{(\lambda\overline{\gamma}_I+\overline{\gamma}_B)^2-4\overline{\gamma}_I\overline{\gamma}_B}}{2}$$

by Lemma 2.1.2.

We also have

$$\begin{aligned} b_I(u_1,u_1)+b_B(u_2,u_2) &\geq \underline{\gamma}_I a(u_1,u_1)+\underline{\gamma}_B/\lambda a(u_2,u_2) \\ &\geq ((\lambda\underline{\gamma}_I+\underline{\gamma}_B)+\sqrt{(\lambda\underline{\gamma}_I+\underline{\gamma}_B)^2-4\underline{\gamma}_B\underline{\gamma}_I})/(2\underline{\gamma}_I\underline{\gamma}_B)a(u,u) \end{aligned}$$

by Lemmas 5.3.1 and 5.3.2. We obtain

$$\lambda_{\max}(T_I+T_B) \leq \frac{(\lambda\underline{\gamma}_I+\underline{\gamma}_B)+\sqrt{(\lambda\underline{\gamma}_I+\underline{\gamma}_B)^2-4\underline{\gamma}_B\underline{\gamma}_I}}{2\underline{\gamma}_I\underline{\gamma}_B}$$

again by Lemma 2.1.2. Therefore we have the upper bound for $\kappa(T_I + T_B)$ as stated in this theorem.

Now we let $\lambda_{\min}(T_I + T_B)^{-1} = k_1$. Then

$$\underline{\gamma}_I a(u_1,u_1)+\frac{\underline{\gamma}_B}{\lambda}a(u_2,u_2) \leq b_I(u_1,u_1)+b_B(u_2,u_2) \leq k_1 a(u,u)$$

Therefore

$$k_1 \geq \frac{\underline{\gamma}_I + \underline{\gamma}_B/\lambda + \sqrt{(\underline{\gamma}_I + \underline{\gamma}_B/\lambda)^2 - 4(1-\eta^2)\underline{\gamma}_I\underline{\gamma}_B/\lambda}}{2(1-\eta^2)}$$
$$= \frac{(\lambda\underline{\gamma}_I + \underline{\gamma}_B) + \sqrt{(\lambda\underline{\gamma}_I + \underline{\gamma}_B)^2 - 4\underline{\gamma}_I\underline{\gamma}_B}}{2}.$$

Let $\lambda_{\max}(T_I + T_B) = k_2$. Then

$$\overline{\gamma}_I a(u_1, u_1) + \overline{\gamma}_B a(u_2, u_2) \geq b_I(u_1, u_1) + b_B(u_2, u_2) \geq k_2^{-1} a(u, u).$$

We obtain

$$k_2 \geq \frac{\overline{\gamma}_I + \overline{\gamma}_B + \sqrt{(\overline{\gamma}_I + \overline{\gamma}_B)^2 - 4\overline{\gamma}_I\overline{\gamma}_B}/\lambda}{2\overline{\gamma}_I\overline{\gamma}_B}.$$

The lower bound for $\kappa(T_I + T_B)$ follows easily. □

We remark that G. Hasse, U. Langer, and A. Meyer in [38] have obtained the following upper bound for $\kappa(C^{-1}K)$:

$$(\kappa(C^{-1}K) \leq \frac{\overline{\gamma}_1}{4\underline{\gamma}_1}(1 + \lambda + \sqrt{\lambda^2 + 2\lambda - 3})^2.$$

Here $\underline{\gamma}_1 = \min\{\underline{\gamma}_I, \underline{\gamma}_B\}$ and $\overline{\gamma}_1 = \max\{\overline{\gamma}_I, \overline{\gamma}_B\}$. We can see that our bound is actually sharper than their bound by a careful computation and the fact that λ is always greater than or equal to 1. We note that these authors obtained their bound completely from an algebraic point of view.

5.4 A new algorithm and some of its variants

In this section, we, for simplicity, first consider a Dirichlet problem. However, our results can be extended to general boundary conditions with slight modifications. Before describing our algorithm, we need some preparations. We first introduce a coordinate system adjacent to $\cup_{i=1}^{p}\partial\Omega_i$. We assume that these subdomains are polygonal, so that the boundaries $\partial\Omega_i$ are piecewise smooth and that at the points where there is a lack of smoothness the angles between the boundary segments belong to an interval $(\phi_0, 2\pi - \phi_0)$ for a certain $\phi_0 > 0$. Let $\Gamma = \cup_{i=1}^{p}\partial\Omega_i$ denote the union of boundaries. In a vicinity of the boundary Γ, we will define a coordinate system. Let L be the segment of Γ, which is a connected component of a smooth part of the boundary of Ω_i connecting the corners of $\partial\Omega_i$

$(1 \le i \le p)$. We assume that the curve L is parametrized as follows: $L(s) = (x(s), y(s))$ where $0 \le s \le l$ (l is the arc length of the segment L) and s defines the counter-clockwise bypassing of Ω_i.

We denote by $\overline{n}(s)$ the inner normal vector at the point $L(s)$:

$$\overline{n}(s) = (-y'(s), x'(s)), \qquad s \in (0, l).$$

Let $\overline{N}(s)$, $s = 0, l$, be the vector of unit length with the origin at the point $L(s)$, bisecting the angle between the segments of the boundary of Ω_i. Let $\phi(s)$, $s = 0, l$, denote the angle between $\overline{n}(s)$ and $\overline{N}(s)$. Note that

$$0 \le \phi(s) \le (\pi - \phi_0)/2$$

if the minimal rotation of $\overline{n}(s)$ around the point $L(s)$ to make it coincide with $\overline{N}(s)$ is performed in the counter-clockwise direction, and

$$-(\pi - \phi_0)/2 \le \phi(s) \le 0$$

otherwise.

Define

$$\phi(s) = \phi(0)(1 - s/l) + \phi(l)s/l$$

and

$$\overline{N}(s) = \overline{n}(s) \begin{bmatrix} cos\phi(s) & sin\phi(s) \\ -sin\phi(s) & cos\phi(s) \end{bmatrix}, \qquad s \in (0, l).$$

The vector $\overline{N}(s)$ is said to be the inner pseudonormal vector at the point $L(s)$ with respect to the subdomain Ω_i.

We assume that each subdomain Ω_i has a characteristic diameter on the order of H. For a given $\delta > 0$, we define in Ω_i the subregions

$$\omega(L, \delta) = \{(x, y) = L(s) + n\overline{N}(s) : \ \ s \in (0, l), \ \ n \in (0, \delta)\} \cap \Omega_i.$$

Since the boundary $\partial\Omega_i$ consists of a finite number m_i of smooth segments L_{ij} $(j = 1, \cdots, m_i)$, we can easily see that, for each $i = 1, \cdots, p$, there exists a constant $\delta_i > 0$ such that the subregion

$$\omega_i(\delta) = \cup_{j=1}^{m_i} \omega(L_{ij}, \delta), \qquad 0 < \delta \le \delta_i,$$

consists of a finite number of disjoint subregions homeomorphic to a ring. This number is equal to the connectivity number of the subdomain Ω_i.

Hereafter, we assume that $\delta > 0$ is fixed and

$$\delta \leq \delta_0 = \min_{1 \leq i \leq p} \delta_i$$

and in the vicinity $\Gamma(\delta)$, we have

$$\Gamma(\delta) = \cup_{i=1}^{p} \omega_i(\delta).$$

Introduce the coordinate system (s, n) adjacent to the boundary defined by a point on Γ and the distance along the pseudonormal. We can see that in $\Gamma(\delta)$ the transformation of (x, y) into (s, n), with δ satisfying $\delta \leq \delta_0$, determines a one-to-one correspondence and is continuous with piecewise smooth Jacobian, bounded from above by a constant $c_1(\delta)$ and bounded away from zero by a constant $c_0(\delta) > 0$.

Next we denote by $W^h(\Omega)$ the subspace of the Sobolev space $W_2^1(\Omega)$ consisting of all continuous functions linear on each triangle with characteristic diameter h of the triangulation of Ω. The scalar product and the norm in $W^h(\Omega)$ are given by the formulas

$$(u, v)_\Omega = \int_\Omega (\nabla u \cdot \nabla v + uv) dx dy,$$
$$\|u\|_\Omega = \sqrt{(u, u)_\Omega}.$$

Let

$$W^h(\Omega, \partial\Omega) = \{u \in W^h(\Omega) : u|_{\partial\Omega} = 0\}$$

and

$$W^{1/2}(\Gamma) = \{g : g = u|_{\cup_{i=1}^{p} \partial\Omega_i}, u \in W^h(\Omega, \partial\Omega\}\}.$$

To introduce a norm on $W^{1/2}(\Gamma)$, we can use either the formula

$$\|g\|_* = \inf_{u \in W^h(\Omega, \partial\Omega), u|_{\cup_{i=1}^{p} \partial\Omega_i} = g} \|u\|_\Omega$$

or the formula

$$\|g\|_{**} = \inf_{u \in W^h(\Gamma(\delta), \partial\Omega \cup \overline{\eta}), u|_{\cup_{i=1}^{p} \partial\Omega_i} = g} \|u\|_{\Gamma(\delta)},$$

where $\overline{\eta} = \partial\Gamma(\delta) \backslash \partial\Omega$. We will prove the following important lemma.

Lemma 5.4.1 *There exist positive constants c_1 and c_2, independent of h, H, and δ, such that*

$$c_1\|g\|_* \le \|g\|_{**} \le c_2\sqrt{\frac{1}{\delta}}\|g\|_*, \qquad \forall g \in W^{1/2}(\Gamma).$$

Proof. Let us take an arbitrary u $\in W^h(\Omega, \partial\Omega)$. Construct a function $v \in W^h(\Gamma(\delta), \partial\Omega \cup \overline{\eta})$ by defining its value at an arbitrary node (x, y) of the mesh domain $\Gamma(\delta)$, with the coordinate (s, n), by the formula

$$v(x, y) = (1 - \frac{n}{\delta})u(x, y).$$

Then, using the arguments in Matsokin [52], we can get the inequality

$$\|v\|_{\Gamma(\delta)} \le c\sqrt{\frac{1}{\delta}}\|u\|_\Omega,$$

where c is a constant independent of h, H, and δ.

We choose $u(x, y) \in W^h(\Omega, \partial\Omega)$ such that $\|u\|_\Omega = \|g\|_*$. Then

$$\|g\|_{**} \le \|v\|_{\Gamma(\delta)} \le c\sqrt{\frac{1}{\delta}}\|u\|_\Omega = c\sqrt{\frac{1}{\delta}}\|g\|_*.$$

The inequality, in other direction,

$$\|g\|_* \le \|g\|_{**}$$

is obvious. □

We now state a variation of above lemma, which is useful for proving the optimality of our new algorithm.

Theorem 5.4.1 *Let $g \in W^{1/2}(\Gamma)$ and define*

$$\|g\|_a^* = \inf_{u \in W^h(\Omega, \partial\Omega), u|_{\cup_{i=1}^p \partial\Omega_i} = g} \sqrt{a(u, u)},$$

$$\|g\|_a^{**} = \inf_{u \in W^h(\Gamma(\delta), \partial\Omega \cup \overline{\eta}), u|_{\cup_{i=1}^p \partial\Omega_i} = g} \sqrt{a(u, u)}.$$

Then there exist positive constants c_3 and c_4, independent of h, H, and δ, such that

$$c_3\|g\|_a^* \le \|g\|_a^{**} \le c_4 N_s^{\frac{1}{2N}} \cdot \sqrt{\frac{H}{\delta}}\|g\|_a^*.$$

Here N_s is the number of subdomains and N is the dimension of the space.

Proof. The inequality

$$\|g\|_a^* \leq \|g\|_a^{**}$$

is obvious from the definition. Let $u \in W^h(\Omega, \partial\Omega)$ such that $\sqrt{a(u,u)} = \|g\|_a^*$. Then,

$$\|g\|_a^{**} \leq \|g\|_{**} \leq c_2\sqrt{\frac{1}{\delta}}\|g\|_* \leq c_2\sqrt{\frac{1}{\delta}}\|u\|_\Omega \leq c_2 c(d)\sqrt{\frac{1}{\delta}}\sqrt{a(u,u)}$$

by using Lemma 5.4.1 and Friedrichs' inequality. $c(d)$ is a constant which only depends upon the diameter d of Ω. Therefore,

$$\|g\|_a^{**} \leq c_2 c(d)\sqrt{\frac{1}{\delta}}\|g\|_a^*.$$

We next want to remove the dependence of the constant upon d. We will use a linear transformation $y = \frac{1}{d}x$ which transform Ω to $\tilde{\Omega}$, Ω_i to $\tilde{\Omega}_i$ such that $diam(\tilde{\Omega}) = 1$ and $diam(\tilde{\Omega}_i) = H/d \equiv \tilde{H}$. This transformation also transforms $\Gamma(\delta)$ onto $\tilde{\Gamma}(\tilde{\delta})$ where $\tilde{\delta} = \delta/d$. We set $\tilde{g}(y) = g(dy)$. Then

$$\|\tilde{g}\|_a^{**} \leq c_2 c(1)\sqrt{\frac{1}{\tilde{\delta}}}\|\tilde{g}\|_a^* \equiv c\sqrt{\frac{1}{\tilde{\delta}}}\|\tilde{g}\|_a^* = c\sqrt{\frac{d}{\delta}}\|\tilde{g}\|_a^*.$$

By the transformation formula for integrals, we also have

$$\|\tilde{g}\|_a^* = d^{\frac{2-N}{2}}\|g\|_a^* \qquad \text{and} \qquad \|\tilde{g}\|_a^{**} = d^{\frac{2-N}{2}}\|g\|_a^{**}$$

where N is the dimension of the space. Therefore

$$\|g\|_a^{**} \leq c\sqrt{\frac{d}{\delta}}\|g\|_a^*.$$

From the fact that $N_s \approx (d/H)^N$, this theorem follows easily. □

Denote the vector of nodal values of g by $\underline{u}_B$. Then

$$\|g\|_a^{*2} = \underline{u}_B^T S_B \underline{u}_B$$

We also will choose B_{II}^{-1} such that

$$(\|g\|_a^{**})^2 = \underline{u}_B^T R_B \underline{u}_B$$

and set

$$A = -B_{II}^{-1} K_{IB}$$

in this case just for convenience. Therefore we can conclude by Theorem 5.4.1 that

$$S_B \leq R_B \leq c \cdot N_s^{1/N} \cdot \frac{H}{\delta} S_B,$$

where c is independent of H, h, and δ. We remark that in this case

$$R_B = \begin{bmatrix} A \\ I_B \end{bmatrix}^T K \begin{bmatrix} A \\ I_B \end{bmatrix}.$$

Finally we can estimate the condition number of $C^{-1}K$ by Theorem 5.3.1 obtaining

$$\kappa(C^{-1}K) \leq c \cdot N_s^{2/N} \cdot \frac{H^2}{\delta^2}$$

for some constant c independent of H, h, and δ.

We next discuss the implementation of our algorithm. Usually it is not easy to compute $A\underline{u}_B$ for a given vector $\underline{u}_B$ because we need to solve a Dirichlet boundary value problem to find a discrete harmonic function in the region $\Gamma(\delta)$ with given boundary values on $\cup_{i=1}^{p} \partial\Omega_i \backslash \partial\Omega$ and zero values on $\overline{\eta} \cup \partial\Omega$. The following implementation rules come from the Poisson integral formula of harmonic function on an annulus with given boundary value on the outer boundary and zero boundary value on the inner boundary. The central idea is that the values of interior points on the domain are taken as the average of values of nearby boundary points. We will use the following general and modified rules.

General rule : Let p denote an arbitrary point in $\Gamma(\delta)$ and denote by ds the distance from p to Γ. Let $q \in \Gamma$ be the point such that $d(p,q) = d(p,\Gamma)$ and L_p be the line segment on Γ with center q and length $2 \cdot ds$. To compute the nodal value at p, we first compute the average value of g on L_p and denote it by a. We then set

$$u(p) = a \cdot \frac{\delta - ds}{\delta};$$

cf. Fig. 5.1.

But this general rule cannot work well if the tangential vector on L_p changes suddenly. Let us take two points p_1 and p_2 in $\Gamma(\delta)$ which are near an interior angle bisector line of a polygonal subdomain but on opposite sides of the bisector. Then the segments L_{p_1} and L_{p_2}, over which we take the averages, are quite different even though the points p and q are near each other. Consequently, the constructed finite element function changes

abruptly near the bisector line and we cannot expect it to provide a uniformly bounded extension, cf. Fig. 5.2. So we need the following modified rule.

Modified rule: Given a point p in $\Gamma(\delta)$, draw a straight line passing through p which is parallel to Γ. Denote the intersection of the straight line and the interior angle bisector by r. Also draw two straight lines from r which are perpendicular to the two sides of $\overline{\eta}$ and denote the intersections by r_1 and r_2. Denote the quadrateral formed by r, r_1, r_2, and the corner by Q. Ignore the part of L_p which is on the boundary of the quadrateral Q and then add the other parts of Γ to L_p which are adjacent to and of the same length as the part that we ignore. Then we get $\overline{L}_p$, cf. Fig. 5.3. Take the average values of g on $\overline{L}_p$, denote it by $\overline{a}$, and obtain the nodal values of the finite element function at p

$$u(p) = \overline{a} \cdot \frac{\delta - ds}{\delta}.$$

If we follow these rules, we get another B_{II}^{-1}. Let us set

$$A_0 = -B_{II}^{-1} K_{IB} \tag{5.6}$$

in this case for convenience. We will have experimental evidence of next section that if we let

$$R_B = \begin{bmatrix} A_0 \\ I_B \end{bmatrix}^T K \begin{bmatrix} A_0 \\ I_B \end{bmatrix},$$

we also have

$$\kappa(C^{-1}K) \leq c \cdot N_s^{2/N} \cdot \frac{H^2}{\delta^2}$$

with a constant c independent of H, h, and δ. Thereafter, we will replace A by A_0 everywhere.

We remark that our arguments in this section can be applied to general mixed-type boundary condition with $\Gamma_D \neq \emptyset$ by tracing the proof of Theorem 5.4.1 and replacing Friedrichs' inequality by Poincaré-Friedrichs' inequality in this proof. Therefore the convergence rate of our algorithm will depends on the constant which appears in Lemma 1.2.3. In the case of Ω being a square and Dirichlet data assigned to at least one side of Ω, we get by Lemma 1.2.4 the same form of convergence rate bound as in the case of Dirichlet boundary condition. Next we will consider some variations of this algorithm. Let P_B be

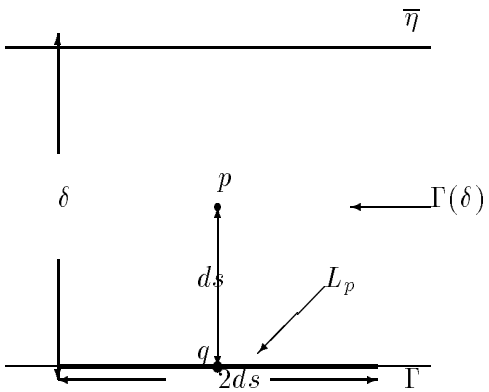

Fig. 5.1

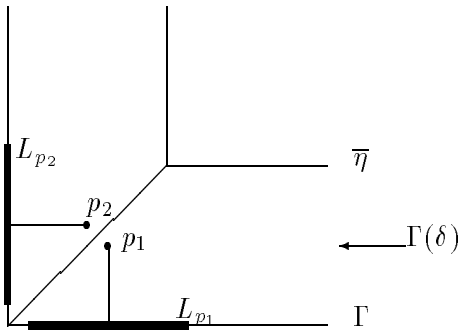

Fig. 5.2

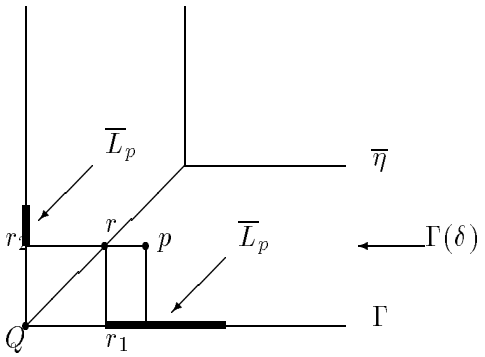

Fig. 5.3

the orthogonal projection onto the subspace V_B. P_B has the matrix representation

$$\begin{bmatrix} A_0 \\ I_B \end{bmatrix} R_B^{-1} \begin{bmatrix} A_0 \\ I_B \end{bmatrix}^T K.$$

Note that any good preconditioner for S_B is also a good preconditioner for R_B because S_B and R_B are spectrally equivalent. So we can use the preconditioners for S_B used in Smith [73], [72] to obtain $P_B u$ for a given $u \in \mathbf{V}$ and we can consider the additive Schwarz method corresponding to the energy-symmetric operator $T_I + P_B$. If we apply our previous results to $T_I + P_B$, we will find that

$$\kappa(T_I + P_B) \leq c \cdot N_s^{1/N} \cdot \frac{H}{\delta}$$

where c is a constant independent of H, h, and δ.

We can also consider a multiplicative Schwarz method corresponding to the error propagation operator $(I - T_I)(I - P_B)$ and assume that $\|T_I\|_a \leq \omega < 2$. Using the standard spectrum radius analysis for $(I - T_I)(I - P_B)$ as in Widlund [82], we will find that

$$\rho((I - T_I)(I - P_B)) \leq \sqrt{1 - \frac{(2-\omega)\lambda_{\min}(T_I + P_B)}{1 + 8\omega^2}} \leq \sqrt{1 - c(2-\omega)\frac{H}{\delta} \cdot N_s^{1/N}}.$$

We can also take B_{II}^{-1} to be a multigrid V-cycle; cf. Hackbusch [41]. Let M be the iteration matrix of one iteration for the multigrid V-cycle method with zero initial guess. M is energy-symmetric and the spectral radius of M is less than 1 uniformly with respect to mesh size. This means that

$$K_{II} M = M^T K_{II} \qquad \text{and} \qquad \|M\|_a \leq \theta < 1.$$

We can use the extension operater A_0 from the boundaries of these subdomains into the interior to provide an initial guess for this multigrid V-cycle because A_0 represents a uniformly energy bounded operator. In this case, we have

$$-B_{II}^{-1} K_{IB} = M A_0 - (I - M) K_{II}^{-1} K_{IB}. \tag{5.7}$$

To compute its transpose, we observe that

$$(-B_{II}^{-1} K_{IB})^T = A_0^T M^T - K_{BI} K_{II}^{-1} (I - M)^T$$

$$= A_0^T M^T - K_{BI}(I - M) K_{II}^{-1} = A_0^T [I - K_{II}(I - M) K_{II}^{-1}] - K_{BI}(I - M) K_{II}^{-1}$$

and note that $(I - M)K_{II}^{-1}\underline{u}_B$ is the approximation of $K_{II}^{-1}\underline{u}_B$ obtained by applying one iteration of multigrid V-cycle method with zero initial guess. We will see in next section that the preconditioner with this B_{II}^{-1} behaves better than if we used the multigrid V-cycle method with zero initial guess.

We now formulate the main algorithm of this section as follows.

Algorithm 5.4.1 *Apply DD-PPM algorithm with B_{II} given in (5.7) to linear system (5.1).*

5.5 Numerical results

In this section, we report on some numerical results to verify the theory which we have previously described. We first report these experimental data for a two-subdomain case. Let

$$\begin{aligned} \Omega_1 &\equiv (0,1) \times (0,1) \\ \Omega_2 &\equiv (0,2) \times (-1,0) \\ \gamma &\equiv (0,1) \times \{0\} \\ \Omega &\equiv \Omega_1 \cup \Omega_2 \cup \gamma. \end{aligned}$$

Consider the Dirichlet problem

$$\begin{aligned} -\Delta u &= f \qquad on \quad \Omega \\ u &= g \qquad on \quad \partial\Omega. \end{aligned}$$

Let N be an even integer, $h = 1/N$, and consider a regular square mesh of mesh size h on Ω. We divide each cell in this mesh into two triangles. Using the resulting triangulation, we discretize the problem using piecewise linear elements. For $\tilde{S}_B$ we choose $S_B^{(1)}$ where $S_B = S_B^{(1)} + S_B^{(2)}$. $\underline{u}_B^T S_B^{(i)} \underline{u}_B$ is the strain energy in Ω_i with given nodal values $\underline{u}_B$ on γ for $i = 1, 2$. By using the extension theorem given in Widlund [78], we find that S_B and $\tilde{S}_B$ are spectrally equivalent. We can apply FFT to compute $(S_B^{(1)})^{-1}\underline{u}_B$ for any given vector $\underline{u}_B$. The process is described in [6].

For $\tilde{K}_{II}$ we choose one multigrid V-cycle with zero initial guess. By the standard theory of multigrid method, K_{II} and $\tilde{K}_{II}$ are spectrally equivalent. We shall give only a

brief description of these cycles here. For the transfer of corrections from coarse to fine grids, we use piecewise bilinear interpolation. For the transfer of residuals from fine to coarse grids, we use full weighting, which is, up to a scalar factor, the Euclidean adjoint of the piecewise bilinear interpolation. For the smoothing of errors, we use red-black Gauss-Seidel iteration, with three half steps red-black-red before and after each coarse grid correction step.

In all experiments for two-subdomain case, we use the same $\tilde{S}_B$ and $\tilde{K}_{II}$, but choose different B_{II}^{-1} for comparison. We have chosen the data f and g such that the Dirichlet problem has the solution $x^2 + y^2 - x\exp(x)\cos(y)$. Because $\kappa(C^{-1}K)$ uniformly bounded, we can choose the following stopping condition

$$\|\text{residual}\|_{l^2} \leq \epsilon \|\text{approx. solution}\|_{l^2} \lambda_{\min}(C^{-1}K).$$

$\lambda_{\min}(C^{-1}K)$ is calculated by using the Lanczos method at very little extra expense. In all experiments, we use $\epsilon = 10^{-6}$; this assures that roughly six digits of the solution are correct, maybe more, regardless of the preconditioner used.

In Table 5.1, we set $B_{II}^{-1} = K_{II}^{-1}$, i.e. the exact solver is used. The subproblems can be solved by using a fast Poisson solver because of the special geometry. We include this experiment to provide a comparision between other variants with inexact solvers.

In Table 5.2, we set $B_{II}^{-1} K_{IB} = 0$, that is, we extend the functions on γ by 0 to the interior of these subdomain. In this case, we see a terrible growth of condition number as the mesh size tends to zero. It can be easily understood. If the nodal values on γ are 1 and they are extended by 0 to the interior of subdomains, we will find that the extended function has an energy $O(1/h)$.

Schwarz PCG	$1/h$				
algorithm	8	16	32	64	128
Evaluated $\kappa(C^{-1}K)$	2.412	2.435	2.434	2.424	2.408
Iteration number	8	8	8	8	8
Max. discretization error	2.242E-3	5.672E-4	1.422E-4	3.479E-5	1.609E-5

Table 5.1. *The case for exact solver* B_{II}^{-1}

Schwarz PCG algorithm	$1/h$				
	8	16	32	64	128
Evaluated $\kappa(C^{-1}K)$	20.63	70.99	262.4	1007	3945
Iteration number	15	24	41	73	249

Table 5.2. *The case for* $-B_{II}^{-1}K_{IB} = 0$

In Table 5.3, we choose $B_{II}^{-1}K_{IB}$ to be A_0 where A_0 is given as in (5.6). We see that the condition numbers are uniformly bounded if we keep δ fixed. The table also verifies that the inequality

$$\kappa(C^{-1}K) \leq c\frac{H^2}{\delta^2}$$

is right.

Schwarz PCG algorithm	$1/h$				
	8	16	32	64	128
Evaluated $\kappa(C^{-1}K)$	7.000	7.315	7.719	8.592	9.380
Iteration number	13	15	16	18	19

$\delta = 1/4$

Schwarz PCG algorithm	$1/h$				
	8	16	32	64	128
Evaluated $\kappa(C^{-1}K)$	20.63	21.13	21.52	21.79	21.97
Iteration number	15	18	19	20	23

$\delta = 1/8$

Schwarz PCG algorithm	$1/h$			
	16	32	64	128
Evaluated $\kappa(C^{-1}K)$	70.99	72.89	73.90	74.48
Iteration number	24	26	27	32

$\delta = 1/16$

Schwarz PCG algorithm	$1/h$		
	32	64	128
Evaluated $\kappa(C^{-1}K)$	262.4	269.3	272.8
Iteration number	41	42	55

$\delta = 1/32$

Table 5.3. *The case for* $-B_{II}^{-1}K_{IB} = A_0$

In Table 5.4, we set B_{II}^{-1} to be the multigrid iteration with zero initial guess. We find that the performance of this B_{II}^{-1} is almost the same as the case of exact solver when the number of V-cycles is equal to 2.

Schwarz PCG algorithm	$1/h$				
	8	16	32	64	128
Evaluated $\kappa(C^{-1}K)$	2.460	2.630	3.313	4.455	6.923
Iteration number	8	9	10	12	13

1 iteration of V-cycle

Schwarz PCG algorithm	$1/h$				
	8	16	32	64	128
Evaluated $\kappa(C^{-1}K)$	2.407	2.430	2.431	2.424	2.410
Iteration number	8	8	8	8	8

2 iteration of V-cycle

Table 5.4. *The case for $-B_{II}^{-1}K_{IB}$ to be multigrid V-cycle with zero initial guess*

In Table 5.5, we have chosen B_{II}^{-1} as a multigrid iteration using an initial guess which is obtained from the extension operator A_0. We choose δ to be $1/4$. We find that the performance of only one iteration is almost the same as for the exact solver. This initial guess is actually better than a zero initial guess for the multigrid V-cycle method.

Schwarz PCG algorithm	$1/h$				
	8	16	32	64	128
Evaluated $\kappa(C^{-1}K)$	2.432	2.447	2.450	2.468	2.475
Iteration number	8	8	8	9	9

1 iteration of V-cycle

Schwarz PCG algorithm	$1/h$				
	8	16	32	64	128
Evaluated $\kappa(C^{-1}K)$	2.437	2.442	2.433	2.415	2.405
Iteration number	8	8	8	8	8

2 iteration of V-cycle

Table 5.5. *The case for $-B_{II}^{-1}K_{IB}$ to be multigrid V-cycle with good initial guess*

In Table 5.6, we have chosen B_{II}^{-1} to be a Gauss-Seidel iteration using an initial guess which is the extended function obtained from the extension operator A_0. We choose δ

to be 1/4. We find that finite numbers of iteration cannot make the condition numbers uniformly bounded with respect to the mesh size h.

Schwarz PCG algorithm	$1/h$				
	8	16	32	64	128
Evaluated $\kappa(C^{-1}K)$	4.406	5.685	6.582	7.243	7.777
Iteration number	10	11	12	14	16

1 Gauss-Seidel iteration

Schwarz PCG algorithm	$1/h$				
	8	16	32	64	128
Evaluated $\kappa(C^{-1}K)$	3.683	5.181	6.236	6.999	7.578
Iteration number	9	10	11	13	16

2 Gauss-Seidel iteration

Table 5.6. *The case of $-B_{II}^{-1}K_{IB}$ being a Gauss-Seidel iteration with a good initial guess*

Next we report on these experimental data for a many-subdomain case. We carry out our remaining experiments on a unit square. We choose these subdomains as equal and denote this characteristic diameter by H. For $\tilde{S}_B$, we choose a preconditioner of S_B introduced in Smith [73]. This preconditioner is spectrally equivalent to S_B. The stopping condition and the ϵ are chosen as before.

In Table 5.7, we consider the case of $B_{II}^{-1} = K_{II}^{-1}$. We include this experiments also to obtain a comparison with other variants with inexact solvers.

In Table 5.8, we set $B_{II}^{-1}K_{IB} = 0$. We find that the condition number grows very much more quickly than in the case of two subdomains. Theoretically, we can see that the extended function has an energy $O((Hh)^{-1})$. This is seen in our experiments.

$1/H$	2	2	2	2	2	4	4	4	4
H/h	4	8	16	32	64	4	8	16	32
$\kappa(C^{-1}K)$	1.801	2.315	2.448	2.449	2.452	1.938	2.768	3.039	3.198
iteration no.	5	7	8	9	9	7	9	10	10

$1/H$	4	8	8	8	8	16	16	16
H/h	64	4	8	16	32	4	8	16
$\kappa(C^{-1}K)$	3.300	2.021	2.798	3.066	3.223	2.000	2.802	3.073
iteration no.	10	8	10	10	11	8	9	10

Table 5.7. *The case of many subdomains with the exact solver B_{II}^{-1}*

$1/H$	2	2	2	2	2	4	4	4
H/h	4	8	16	32	64	4	8	16
$\kappa(C^{-1}K)$	15.77	66.50	281.6	1171	4791	137.5	685.0	3062
iteration no.	9	17	29	52	94	24	45	87

$1/H$	4	8	8	8	16	16
H/h	32	4	8	16	4	8
$\kappa(C^{-1}K)$	12940	1790	9643	44090	27050	149200
iteration no.	166	59	138	282	191	447

Table 5.8. *The case of many subdomains with the inexact solver* $B_{II}^{-1}K_{IB} = 0$

In Table 5.9, we choose $B_{II}^{-1}K_{IB}$ to be A_0 where A_0 is given as in (5.6) for the case of many subdomains. We choose $\delta/H = 1/4$. It is obvious that the condition number grows quickly when we increase the number of subdomains.

$1/H$	2	2	2	2	2	4	4	4	4
H/h	4	8	16	32	64	4	8	16	32
$\kappa(C^{-1}K)$	15.77	16.29	21.17	29.38	40.92	137.5	135.2	162.1	202.7
iteration no.	9	14	18	23	28	24	28	34	43

$1/H$	4	8	8	8	8	16	16	16
H/h	64	4	8	16	32	4	8	16
$\kappa(C^{-1}K)$	248.4	1790	1744	2082	2570	27050	26290	31380
iteration no.	49	59	63	76	86	191	195	220

Table 5.9. *The case of many subdomains with* $-B_{II}^{-1}K_{IB} = A_0$ *and* $\delta/H = 1/4$

In Table 5.10, we set B_{II}^{-1} to be one multigrid V-cycle with zero initial guess. We find that the resulting algorithm is not optimal because the condition number grows if the number of subdomains increases. However, if we apply two multigrid V-cycles , then the resulting algorithm is optimal as seen from Table 5.11.

$1/H$	2	2	2	2	2	4	4	4	4
H/h	4	8	16	32	64	4	8	16	32
$\kappa(C^{-1}K)$	1.851	2.802	3.464	4.536	7.296	2.314	4.535	6.573	11.77
iteration no.	5	8	10	11	14	8	12	14	17

$1/H$	4	8	8	8	8	16	16	16
H/h	64	4	8	16	32	4	8	16
$\kappa(C^{-1}K)$	27.16	3.607	10.07	25.98	72.50	11.79	61.55	226.8
iteration no.	22	11	16	21	27	14	24	35

Table 5.10. *The case of many subdomains with $-B_{II}^{-1}K_{IB}$ being one multigrid V-cycle with a zero initial guess*

$1/H$	2	2	2	2	2	4	4	4	4
H/h	4	8	16	32	64	4	8	16	32
$\kappa(C^{-1}K)$	1.802	2.336	2.482	2.519	2.592	1.936	2.874	3.234	3.538
iteration no.	5	7	8	9	9	7	9	10	11

$1/H$	4	8	8	8	8	16	16	16
H/h	64	4	8	16	32	4	8	16
$\kappa(C^{-1}K)$	3.879	2.032	3.159	3.716	4.282	2.186	3.817	4.897
iteration no.	12	8	11	12	13	8	12	13

Table 5.11. *The case of many subdomains with $-B_{II}^{-1}K_{IB}$ being two multigrid V-cycles with a zero initial guess*

In Table 5.12, we have chosen B_{II}^{-1} as one multigrid V-cycle using an initial guess obtained from the extension operator A_0. We see that the condition number is much smaller than in Table 5.10 although the condition number still grows. In Table 5.13, we apply two multigrid V-cycles with the same initial guess. This shows that no significant difference when compared with Table 5.11.

$1/H$	2	2	2	2	2	4	4	4	4
H/h	4	8	16	32	64	4	8	16	32
$\kappa(C^{-1}K)$	1.851	2.605	2.828	2.880	2.940	2.314	3.942	4.571	4.958
iteration no.	5	7	9	10	10	8	11	12	12

$1/H$	4	8	8	8	8	16	16	16
H/h	64	4	8	16	32	4	8	16
$\kappa(C^{-1}K)$	5.250	3.607	6.216	7.574	8.398	11.79	27.34	39.01
iteration no.	13	11	14	15	16	14	20	22

Table 5.12. *The case of many subdomains with $-B_{II}^{-1}K_{IB}$ being one multigrid V-cycle with a good initial guess*

$1/H$	2	2	2	2	2	4	4	4	4
H/h	4	8	16	32	64	4	8	16	32
$\kappa(C^{-1}K)$	1.802	2.329	2.463	2.466	2.470	1.936	2.841	3.141	3.316
iteration no.	5	7	8	9	9	7	9	10	11

$1/H$	4	8	8	8	8	16	16	16
H/h	64	4	8	16	32	4	8	16
$\kappa(C^{-1}K)$	3.429	2.032	3.060	3.426	3.632	2.186	3.554	4.095
iteration no.	11	8	11	11	12	8	11	12

Table 5.13. *The case of many subdomains with $-B_{II}^{-1}K_{IB}$ being two multigrid V-cycles with a good initial guess*

Chapter 6

Iterative Substructuring Methods

6.1 Description of the Algorithm

In this chapter, we use the ideas of Chapter 5 to develop another optimal algorithm which is a variant of Smith's algorithm in [73]. Let us first describe Smith's algorithm in two dimension. For simplicity, we assume that the elliptic problem (1.1)-(1.3) has a homogeneous Dirichlet boundary condition.

We consider the Schur complement problem,

$$S_B u_B = g_B, \tag{6.1}$$

obtained by eliminating the interior variables as described in Section 1.4. Smith's algorithm is an additive Schwarz scheme on the interfaces between the substructures.

The variational problem is to find $u \in \tilde{V}^h$ such that

$$a(u, v) = f(v), \qquad \forall v \in \tilde{V}^h,$$

where $\tilde{V}^h$ is the subspace of V^h of all continuous functions linear on each triangle with characteristic diameter h of the triangulation of Ω, which are discrete harmonic in the interiors of the subdomains.

Let the union of interfaces of the subdomains be Γ. Smith worked with subspaces of $\tilde{V}^h$ associated with overlapping subsets of Γ and a global coarse space. We can regard the boundaries of the substructures as consisting of two parts: the substructure vertices and the edges between the substructure vertices. The subset Γ^{E_i} are formed by individual edges of the substructures. Smith chose Γ^{V_j} to be regions consisting of a vertex and an

overlap of order H onto all of the adjacent edges. We can easily constrain the overlap so that no portion of Γ is covered more than four times.

The subspaces of $\tilde{V}^h$ are given by

$$V^H,$$

$$\tilde{V}^h_{E_i} = \{\phi \in \tilde{V}^h : \mathrm{supp}(\phi|_\Gamma) \subseteq \Gamma^{E_i}\},$$

and

$$\tilde{V}^h_{V_j} = \{\phi \in \tilde{V}^h : \mathrm{supp}(\phi|_\Gamma) \subseteq \Gamma^{V_j}\}.$$

It is easy to see that each of these spaces is a subspaces of $\tilde{V}^h$ and that

$$\tilde{V}^h = (\sum_i \tilde{V}^h_{E_i}) + (\sum_j \tilde{V}^h_{V_k}).$$

The projection onto each edge subspace is given by

$$P_{E_i} = R^T_{E_i} S^{-1}_{E_i} R_{E_i} S_B,$$

where R_{E_i} is the restriction operator, which returns only those unknowns which are associated with Γ^{E_i}, and S_{E_i} is the principal minor of the Schur complement S_B which is associated with that same set of unknowns. The projections onto vertex subspaces are formed in the same manner.

For the coarse space V^H, the projection operator is of the form

$$P_H = R^T_H K^{-1}_H R_H S_B.$$

The operator R^T_H represents linear interpolation from V^H to $\tilde{V}^h$ and K_H is the stiffness matrix for the original problem treating the substructures as elements.

We now describe the Smith's algorithm. We note that many opportunities exist for parallelism between and within each step.

Algorithm 6.1.1 (Smith's algorithm) *Apply the preconditioned conjugate gradient method to the linear system (6.1) using the preconditioner $\tilde{S}_B$ with*

$$\tilde{S}^{-1}_B = R^T_H K^{-1}_H R_H + \sum_i R^T_{E_i} S^{-1}_{E_i} R_{E_i} + \sum_j R^T_{V_j} S^{-1}_{V_j} R_{V_j}.$$

Smith's algorithm is optimal which can be seen in the following theorem.

Theorem 6.1.1 *The condition number of the preconditioned system for Smith's method is bounded independently of the size of the subdomains H and the size of the elements h, i.e.*

$$\kappa(\tilde{S}_B^{-1} S_B) \leq C.$$

We will derive a variation of the above algorithm. We first observe that S_{E_i} is the Schur complement of a domain which has only two subdomains with homogeneous Dirichlet boundary condition. According to Section 5.5, we can find another operator $S_{E_i}^{(A_0)}$ which is spectrally equivalent to S_{E_i} such that

$$S_{E_i} \leq S_{E_i}^{(A_0)} \leq c\frac{H}{\delta} S_{E_i} \tag{6.2}$$

where c is a constant independent of H, h, and δ and the superscript A_0 indicates that the construction of $S_{E_i}^{(A_0)}$ is closely related to the operator A_0 given in equation (5.6). Similarly, S_{V_j} is a principal minor of the Schur complement of a domain which by shape regularity is given by a finite number of contributions from neighbouring subdomains. Therefore, we can also find another operator $S_{V_j}^{(A_0)}$ which is spectrally equivalent to S_{V_j} such that

$$S_{V_j} \leq S_{V_j}^{(A_0)} \leq c\frac{H}{\delta} S_{V_j}. \tag{6.3}$$

If we replace S_{E_i} and S_{V_j} by $S_{E_i}^{(A_0)}$ and $S_{V_j}^{(A_0)}$ in Smith's algorithm, we will obtain an alternative preconditioner $\tilde{R}_B$ for S_B where

$$(\tilde{R}_B)^{-1} = R_H^T K_H^{-1} R_H + \sum_i R_{E_i}^T (S_{E_i}^{(A_0)})^{-1} R_{E_i} + \sum_j R_{V_j}^T (S_{V_j}^{(A_0)})^{-1} R_{V_j}.$$

We see that

$$\tilde{S}_B \leq \tilde{R}_B \leq c\frac{H}{\delta} \tilde{S}_B.$$

Therefore,

$$\kappa((\tilde{R}_B)^{-1} S_B) \leq C\frac{H}{\delta}$$

where C is a constant independent of H, h, and δ. This follows from the fact that Smith's algorithm is optimal. We conclude that we have obtained optimal algorithm. Let us summarize the main results of this section as follows.

Algorithm 6.1.2 (Modified algorithm) *Apply the preconditioned conjugate gradient method to the linear system (6.1), using the preconditioner $\tilde{R}_B$, with*

$$\tilde{R}_B^{-1} = R_H^T K_H^{-1} R_H + \sum_i R_{E_i}^T (S_{E_i}^{(A_0)})^{-1} R_{E_i} + \sum_j R_{V_j}^T (S_{V_j}^{(A_0)})^{-1} R_{V_j}.$$

Theorem 6.1.2 *The condition number of the preconditioned system for modified algorithm is bounded by i.e.*

$$\kappa(\tilde{R}_B^{-1} S_B) \leq C \frac{H}{\delta}$$

where C is a constant independent of H, h, and δ.

If we use the new algorithm rather than Smith's algorithm, we get two benefits. First, it is easier to form $S_{E_i}^{(A_0)}$ and $S_{V_j}^{(A_0)}$ than S_{E_i} and S_{V_j} because we need not solve the problem in the interiors of the subdomains exactly. Second, $S_{E_i}^{(A_0)}$ and $S_{V_j}^{(A_0)}$ are sparse matrices and $S_{E_i}^{(A_0)}$ are band matrices with band width $O(\delta/h)$. It is then easier to factor $S_{E_i}^{(A_0)}$ and $S_{V_j}^{(A_0)}$ than S_{E_i} and S_{V_j}. However, the condition number of $(\tilde{R}_B)^{-1} S_B$ is bigger than that of $\tilde{S}_B^{-1} S_B$.

For completeness, we also consider the case of vertex spaces with small overlap δ_0. This work has been inspired by several series of numerical experiments that indicate that the rate of convergence is quite satisfactory even for a small overlap and that the running time of the programs is often the smallest when the overlap is at a minimum. The number of conjugate gradient iterations is typically higher in such a case but this can be compensated for by the fact that the local problems are smaller and therefore cheaper to solve; cf. Skogen [70], Bjørstad and Skogen [5], and Cai, Gropp and Keyes [22]. We now state a theorem about Smith's method for the case of small overlap. A proof can be found in Dryja and Widlund [34].

Theorem 6.1.3 *The condition number of the preconditioned system for Smith's method with small overlap size δ_0 satisfies*

$$\kappa(\tilde{S}_B^{-1} S_B) \leq C(1 + \log(H/\delta_0))^2.$$

The constant is independent of H, h and δ_0.

If we consider our modified algorithm for the case of small overlap, we can obtain the following theorem by using equations (6.2) and (6.3)

Theorem 6.1.4 *The condition number of the preconditioned system for the modified method with small overlap size δ_0 satisfies*

$$\kappa(\tilde{R}_B^{-1} S_B) \leq C \cdot H/\delta \cdot (1 + \log(H/\delta_0))^2.$$

The constant is independent of H, h, δ and δ_0.

We will see from the experiments of next section that if we choose $H/\delta = 1/4$, then there is no significant difference between the condition numbers of the two methods for the case of small overlap.

6.2 Numerical Results

In this section, we report on some numerical experiments which compare Smith's algorithm and our modified algorithm. We have carried out these experiments on the unit square. We choose these subdomains to be equal and denote the characteristic diameter of these subdomains by H. The stopping condition and the ϵ are chosen as in Chapter 5.

In Tables 6.1, 6.2, 6.3 and 6.4, we choose the overlap of an edge space and an vertex space to be $H/4$. From Tables 6.2, 6.3 and 6.4, we see that our modified algorithm is optimal as long as the value of H/δ is fixed. Furthermore, the condition number bound, provided by Theorem 6.1.2, is sharp.

$1/H$	2	2	2	2	2	4	4	4	4
H/h	4	8	16	32	64	4	8	16	32
$\kappa(\tilde{S}_B^{-1} S_B)$	1.801	2.360	2.293	2.262	2.243	1.938	2.661	2.771	2.860
iteration no.	4	6	7	8	8	8	10	10	10

$1/H$	4	8	8	8	8	16	16	16
H/h	64	4	8	16	32	4	8	16
$\kappa(\tilde{S}_B^{-1} S_B)$	2.927	2.025	2.593	2.741	2.835	2.018	2.568	2.710
iteration no.	10	8	9	10	10	8	9	9

Table 6.1. *Smith's algorithm*

$1/H$	2	2	2	2	2	4	4	4	4
H/h	4	8	16	32	64	4	8	16	32
$\kappa(\tilde{S}_B^{-1} S_B)$	2.178	2.741	2.737	2.842	3.250	2.305	3.374	3.524	3.514
iteration no.	6	8	9	10	11	9	11	12	12

$1/H$	4	8	8	8	8	16	16	16
H/h	64	4	8	16	32	4	8	16
$\kappa(\tilde{S}_B^{-1}S_B)$	4.072	2.375	3.488	3.636	3.589	2.406	3.487	3.646
iteration no.	13	9	12	12	12	9	11	12

Table 6.2. *The modified algorithm with* $\delta/H = 1/4$

$1/H$	2	2	2	2	4	4	4
H/h	8	16	32	64	8	16	32
$\kappa(\tilde{S}_B^{-1}S_B)$	4.601	4.824	4.941	4.581	5.724	6.173	6.178
iteration no.	10	12	13	13	15	16	16

$1/H$	4	8	8	8	16	16
H/h	64	8	16	32	8	16
$\kappa(\tilde{S}_B^{-1}S_B)$	5.953	5.780	6.221	6.226	5.820	6.257
iteration no.	16	15	16	17	15	16

Table 6.3. *The modified algorithm with* $\delta/H = 1/8$

$1/H$	2	2	2	4	4	4	8	8	16
H/h	16	32	64	16	32	64	16	32	16
$\kappa(\tilde{S}_B^{-1}S_B)$	9.400	9.695	9.472	10.68	11.64	11.68	10.72	11.67	10.79
iteration no.	16	17	17	22	23	23	22	23	21

Table 6.4. *The modified algorithm with* $\delta/H = 1/16$

In Tables 6.5 and 6.6, we consider the case of small overlap. We take the small overlap δ_0 to be just one mesh size h. In Table 6.6, the value of H/δ is chosen to be $1/4$. From these tables, we conclude that there is little difference in the condition numbers of these two methods. However, our modified algorithm is cheaper to achieve for each iteration step. Therefore we can say that our modified algorithm is better to implement than Smith's algorithm in this case. These tables also show that the bounds of Theorems 6.1.3 and 6.1.4 are sharp.

$1/H$	2	2	2	2	2	4	4	4	4
H/h	4	8	16	32	64	4	8	16	32
$\kappa(\tilde{S}_B^{-1}S_B)$	1.801	2.368	3.068	3.884	4.812	1.938	2.688	3.627	4.716
iteration no.	4	5	6	6	7	8	9	11	12

$1/H$	4	8	8	8	8	16	16	16
H/h	64	4	8	16	32	4	8	16
$\kappa(\tilde{S}_B^{-1}S_B)$	5.962	2.025	2.813	3.784	4.913	2.018	2.798	3.807
iteration no.	13	8	10	11	12	8	9	11

Table 6.5. *Smith's algorithm for the case of small overlap*

$1/H$	2	2	2	2	2	4	4	4	4
H/h	4	8	16	32	64	4	8	16	32
$\kappa(\tilde{S}_B^{-1}S_B)$	2.178	2.604	3.442	4.430	5.593	2.305	2.562	3.473	4.517
iteration no.	6	7	8	8	9	9	9	10	12

$1/H$	4	8	8	8	8	16	16	16
H/h	64	4	8	16	32	4	8	16
$\kappa(\tilde{S}_B^{-1}S_B)$	5.799	2.375	2.634	3.763	4.930	2.406	2.676	3.823
iteration no.	13	9	10	11	13	9	10	11

Table 6.6. *The modified algorithm with $H/\delta = 1/4$ for the case of small overlap*

Bibliography

[1] R. A. Adams, *Sobolev Spaces*, Academic Press, New York, 1975.

[2] P. E. Bjørstad, *Multiplicative and Additive Schwarz Methods: Convergence in the 2 domain case*, in Domain Decomposition Methods, T. Chan, R. Glowinski, J. Périaux, and O. Widlund, eds., Philadelphia, PA, 1989, SIAM.

[3] P. E. Bjørstad and J. Mandel, *On the spectra of sums of orthogonal projections with applications to parallel computing*, BIT, 31 (1991), pp. 76–88.

[4] P. E. Bjørstad, R. Moe, and M. Skogen, *Parallel domain decomposition and iterative refinement algorithms*, in Parallel Algorithms for PDEs, Proceedings of the 6th GAMM-Seminar held in Kiel, Germany, January 19–21, 1990, W. Hackbusch, ed., Braunschweig, Wiesbaden, 1990, Vieweg-Verlag.

[5] P. E. Bjørstad and M. Skogen, *Domain decomposition algorithms of Schwarz type, designed for massively parallel computers*, in Fifth International Symposium on Domain Decomposition Methods for Partial Differential Equations, T. F. Chan, D. E. Keyes, G. A. Meurant, J. S. Scroggs, and R. G. Voigt, eds., Philadelphia, PA, 1992, SIAM.

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

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

[8] F. Bornemann and H. Yserentant, *A basic norm equivalence for the theory of multilevel methods*, Numerische Mathematik, 64 (1993), pp. 455–476.

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

[10] J. H. Bramble and J. E. Pasciak, *New estimates for multilevel algorithms including the V-cycle*, tech. report, Cornell University, New York, 1991. PREPRINT.

[11] J. H. Bramble, J. E. Pasciak, and A. H. Schatz, *The construction of preconditioners for elliptic problems by substructuring, I*, Math. Comp., 47 (1986), pp. 103–134.

[12] ——, *The construction of preconditioners for elliptic problems by substructuring, II*, Math. Comp., 49 (1987), pp. 1–16.

[13] ——, *The construction of preconditioners for elliptic problems by substructuring, III*, Math. Comp., 51 (1988), pp. 415–430.

[14] ——, *The construction of preconditioners for elliptic problems by substructuring, IV*, Math. Comp., 53 (1989), pp. 1–24.

[15] J. H. Bramble, J. E. Pasciak, J. Wang, and J. Xu, *Convergence estimates for multigrid algorithms without regularity assumptions*, Math. Comp., 57 (1991), pp. 23–45.

[16] ——, *Convergence estimates for product iterative methods with applications to domain decomposition*, Math. Comp., 57 (1991), pp. 1–21.

[17] J. H. Bramble, J. E. Pasciak, and J. Xu, *Parallel multilevel preconditioners*, Math. Comp., 55 (1990), pp. 1–22.

[18] J. H. Bramble and J. Xu, *Some estimates for a weighted L^2 projection*, Math. Comp., 56 (1991), pp. 463–476.

[19] P. L. Butzer and H. Berens, *Semi-Groups of Operators and Approximation*, Springer, Berlin/Heidelberg/New York/Tokyo, 1967.

[20] 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, PA, 1990.

[21] ——, *An optimal two-level overlapping domain decomposition method for elliptic problems in two and three dimensions*, SIAM J. Sci. Comp., 14 (1993), pp. 239–247.

[22] X.-C. Cai, W. D. Gropp, and D. E. Keyes, *Convergence rate estimate for a domain decomposition method*, tech. report, Yale University, 1990. Submitted to Numer. Math.

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

[24] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, 1978.

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

[26] ——, *A method of domain decomposition for 3-D finite element problems*, in First International Symposium on Domain Decomposition Methods for Partial Differential Equations, R. Glowinski, G. H. Golub, G. A. Meurant, and J. Périaux, eds., Philadelphia, PA, 1988, SIAM.

[27] M. Dryja, W. Proskurowski, and O. Widlund, *A method of domain decomposition with crosspoints for elliptic finite element problems*, in Optimal Algorithms, B. Sendov, ed., Sofia, Bulgaria, 1986, Bulgarian Academy of Sciences, pp. 97–111.

[28] M. Dryja, B. F. Smith, and O. B. Widlund, *Schwarz analysis of iterative substructuring algorithms for elliptic problems in three dimensions*, Tech. Report 638, Department of Computer Science, Courant Institute, 1993.

[29] M. Dryja and O. B. Widlund, *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 Institute, 1987.

[30] ——, *On the optimality of an additive iterative refinement method*, in Proceedings of the Fourth Copper Mountain Conference on Multigrid Methods, Philadelphia, PA, 1989, SIAM, pp. 161–170.

[31] ——, *Some domain decomposition algorithms for elliptic problems*, in Iterative Methods for Large Linear Systems, L. Hayes and D. Kincaid, eds., San Diego, California, 1989, Academic Press, pp. 273–291. Proceeding of the Conference on Iterative Methods for Large Linear Systems held in Austin, Texas, October 19 - 21, 1988, to celebrate the sixty-fifth birthday of David M. Young, Jr.

[32] ——, *Towards a unified theory of domain decomposition algorithms for elliptic problems*, in Third International Symposium on Domain Decomposition Methods for Partial Differential Equations, held in Houston, Texas, March 20-22, 1989, T. Chan, R. Glowinski, J. Périaux, and O. Widlund, eds., SIAM, Philadelphia, PA, 1990.

[33] ——, *Multilevel additive methods for elliptic finite element problems*, in Parallel Algorithms for Partial Differential Equations, Proceedings of the Sixth GAMM-Seminar, Kiel, January 19–21, 1990, W. Hackbusch, ed., Braunschweig, Germany, 1991, Vieweg & Son.

[34] ——, *Domain decomposition algorithms with small overlap*, Tech. Report 606, Department of Computer Science, Courant Institute, May 1992. To appear in SIAM J. Sci. Stat. Comput.

[35] N. Dunford and J. T. Schwartz, *Linear operators*, Interscience Publishers, New York/London/Sydney, 1967.

[36] R. E. Ewing, R. D. Lazarov, and P. S. Vassilevski, *Local refinement techniques for elliptic problems on cell-centered grids*, tech. report, University of Wyoming, 1988.

[37] U. L. G. Hasse and A. Meyer, *The approximate Dirichlet domain decomposition method. Part ii: Applications to 2nd-order elliptic B.V.P.s*, Computing, 47 (1991), pp. 153–167.

[38] ——, *Domain decomposition preconditioners with inexact subdomain solvers*, tech. report, Technical University of Chemnitz, Germany, 1991.

[39] G. H. Golub and C. F. V. Loan, *Matrix Computations*, Johns Hopkins Univ. Press, 1989. Second Edition.

[40] P. Grisvard, *Elliptic problems in nonsmooth domains*, Pitman Publishing, Boston, 1985.

[41] W. Hackbusch, *Multigrid methods and applications*, Springer, Berlin, Heidelberg, New York, Tokyo, 1985.

[42] L. Hart and S. McCormick, *Asynchronous multilevel adaptive methods for solving partial differential equations on multiprocessors: Computational analysis*, tech. report, Comp. Math. Group, Univ. of Colorado at Denver, 1987. Submitted to Parallel Computing.

[43] C. Lanczos, *An iteration method for the solution of the eigenvalue problem of linear differential and integral operators*, J. Res. Nat. Bur. Standards, 45 (1950), pp. 255–282.

[44] P. Lax and A. Milgram, *Contributions to the theory of partial differential equations*, in Annals of Math. Studies 33, New Jersey, 1954, Princeton University Press, pp. 167–190.

[45] J.-L. Lions and E. Magenes, *Nonhomogeneous Boundary Value Problems and Applications*, vol. I, Springer, New York, Heidelberg, Berlin, 1972.

[46] P. L. Lions, *On the Schwarz alternating method. I.*, in First International Symposium on Domain Decomposition Methods for Partial Differential Equations, R. Glowinski, G. H. Golub, G. A. Meurant, and J. Périaux, eds., Philadelphia, PA, 1988, SIAM.

[47] D. G. Luenberger, *Introduction to Linear and Nonlinear Programming*, Addison-Wesley, New York, 1973.

[48] J. Mandel and S. McCormick, *Iterative solution of elliptic equations with refinement: The model multi-level case*, in Domain Decomposition Methods, T. Chan, R. Glowinski, J. Périaux, and O. Widlund, eds., Philadelphia, PA, 1989, SIAM.

[49] ——, *Iterative solution of elliptic equations with refinement: The two-level case*, in Domain Decomposition Methods, T. Chan, R. Glowinski, J. Périaux, and O. Widlund, eds., Philadelphia, PA, 1989, SIAM.

[50] T. P. Mathew, *Schwarz alternating and iterative refinement methods for mixed formulations of elliptic problems, part I: Algorithms and Numerical results*, Numerische Mathematik, (1993). To appear.

[51] ——, *Schwarz alternating and iterative refinement methods for mixed formulations of elliptic problems, part II: Theory*, Numerische Mathematik, (1993). To appear.

[52] A. Matsokin, *Norm preserving prolongation of mesh functions*, Sov. J. Numer. Anal. Math. Modelling, 3 (1988), pp. 137–149.

[53] A. M. Matsokin and S. V. Nepomnyaschikh, *A Schwarz alternating method in a subspace*, Soviet Mathematics, 29(10) (1985), pp. 78–84.

[54] S. McCormick, *Fast adaptive composite grid (FAC) methods*, in Defect Correction Methods: Theory and Applications, K. Böhmer and H. J. Stetter, eds., Computing Supplementum 5, Springer-Verlag, Wien, 1984, pp. 115–121.

[55] S. McCormick and J. Thomas, *The fast adaptive composite grid (FAC) method for elliptic equations*, Math. Comp., 46 (1986), pp. 439–456.

[56] S. F. McCormick, *Multigrid Methods*, SIAM, Philadelphia, PA, 1987.

[57] J. A. Meijerink and H. Van der vorst, *An iterative solution method for linear equation systems of which the coefficient matrix is a symmetric M-matrix*, Math. Comp., 31 (1977), pp. 148–162.

[58] A. Meyer, *A parallel preconditioned conjugate gradient method using domain decomposiyion and inexact solvers*, Computing, 45 (1990), pp. 217–234.

[59] R. Moe, *Iterative Local Uniform Mesh Refinement Methods and Parallel Processing*, PhD thesis, University of Bergen, Norway, June 1992. Department of Informatics.

[60] S. V. Nepomnyaschikh, *Domain Decomposition and Schwarz Methods in a Subspace for the Approximate Solution of Elliptic Boundary Value Problems*, PhD thesis, Computing Center of the Siberian Branch of the USSR Academy of Sciences, Novosibirsk, USSR, 1986.

[61] J. Nečas, *Les méthodes directes en théorie des équations elliptiques*, Academia, Prague, 1967.

[62] J. T. Oden and J. N. Reddy, *An Introduction to the Mathematical Theory of Finite Elements*, John Wiley and Sons, New York, 1982.

[63] P. Oswald, *On function spaces related to finite element approximation theory*, Z. Anal. Anwendungen, (1990), pp. 43–64.

[64] B. N. Parlett, *The symmetric eigenvalue problem*, Prentice-Hall, 1980.

[65] L. F. Pavarino, *An additive Schwarz method for the p-version finite element method*, Tech. Report 580, Courant Institute of Mathematical Sciences, Department of Computer Science, September 1991. A revised version has been accepted for Numer. Math.

[66] J. S. Przemieniecki, *Matrix structural analysis of substructures.*, Am. Inst. Aero. Astro. J., 1 (1963), pp. 138–147.

[67] Y. Saad and M. H. Schultz, *GMRES: A generalized minimum residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comp., 7 (1986), pp. 856–869.

[68] H. A. Schwarz, *Gesammelte Mathematische Abhandlungen*, vol. 2, Springer, Berlin, 1890, pp. 133–143. First published in Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich, volume 15, 1870, pp. 272–286.

[69] R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493.

[70] M. D. Skogen, *Schwarz Methods and Parallelism*, PhD thesis, Department of Informatics, University of Bergen, Norway, February 1992.

[71] B. F. Smith, *Domain Decomposition Algorithms for the Partial Differential Equations of Linear Elasticity*, PhD thesis, Courant Institute of Mathematical Sciences, September 1990. Tech. Rep. 517, Department of Computer Science, Courant Institute.

[72] ——, *A domain decomposition algorithm for elliptic problems in three dimensions*, Numer. Math., 60 (1991), pp. 219–234.

[73] ——, *An optimal domain decomposition preconditioner for the finite element solution of linear elasticity problems*, SIAM J. Sci. Stat. Comput., 13 (1992), pp. 364–378.

[74] S. L. Sobolev, *L'Algorithme de Schwarz dans la Théorie de l'Elasticité*, Comptes Rendus (Doklady) de l'Académie des Sciences de l'URSS, IV (1936), pp. 243–246.

[75] E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton University Press, New Jersey, 1970.

[76] G. Strang, *Approximation in the finite element method*, Numer. Math., 19 (1972), pp. 81–98.

[77] G. Strang and G. J. Fix, *An Analysis of the Finite Element Method*, Prentice-Hall, Englewood Cliffs, N.J., 1973.

[78] O. B. Widlund, *An extension theorem for finite element spaces with three applications*, in Numerical Techniques in Continuum Mechanics, W. Hackbusch and K. Witsch, eds., Braunschweig/Wiesbaden, 1987, Notes on Numerical Fluid Mechanics, v. 16, Friedr. Vieweg und Sohn, pp. 110–122. Proceedings of the Second GAMM-Seminar, Kiel, January, 1986.

[79] ——, *Iterative substructuring methods: Algorithms and theory for elliptic problems in the plane*, in First International Symposium on Domain Decomposition Methods for Partial Differential Equations, R. Glowinski, G. H. Golub, G. A. Meurant, and J. Périaux, eds., Philadelphia, PA, 1988, SIAM.

[80] ——, *Iterative substructuring methods: The general elliptic case*, in Computational Processes and Systems, 6, Moscow, 1988, Nauka. Proceedings of Modern Problems in

Numerical Analysis, a conference held in Moscow, USSR, September 1986. In Russian, also available from the author, in English.

[81] ——, *Optimal iterative refinement methods*, in Domain Decomposition Methods, T. Chan, R. Glowinski, J. Périaux, and O. Widlund, eds., Philadelphia, PA, 1989, SIAM.

[82] ——, *Some Schwarz methods for symmetric and nonsymmetric elliptic problems*, in Fifth International Symposium on Domain Decomposition Methods for Partial Differential Equations, T. F. Chan, D. E. Keyes, G. A. Meurant, J. S. Scroggs, and R. G. Voigt, eds., Philadelphia, PA, 1992, SIAM.

[83] J. Xu, *Iterative methods by space decomposition and subspace correction*, SIAM Review, 34 (1992), pp. 581–613.

[84] H. Yserentant, *On the multi-level splitting of finite element spaces*, Numer. Math., 49 (1986), pp. 379–412.

[85] X. Zhang, *Domain decomposition algorithms for the biharmonic Dirichlet problem*, in Fifth International Symposium on Domain Decomposition Methods for Partial Differential Equations, T. F. Chan, D. E. Keyes, G. A. Meurant, J. S. Scroggs, and R. G. Voigt, eds., Philadelphia, PA, 1992, SIAM.

[86] ——, *Multilevel Schwarz methods*, Numerische Mathematik, 63 (1992), pp. 521–539.