Some Domain Decomposition and Multigrid Preconditioners for Hybrid Mixed Finite Elements

Lawrence C. Cowsar

CRPC-TR94465-S
April, 1994

Center for Research on Parallel Computation
Rice University
P.O. Box 1892
Houston, TX 77251-1892

RICE UNIVERSITY

Some Domain Decomposition and Multigrid Preconditioners for Hybrid Mixed Finite Elements

by

Lawrence Charles Cowsar

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE:

Mary Fanett Wheeler, Chairman
Professor of Computational and Applied Mathematics

Todd Arbogast
Assistant Professor of Computational and Applied Mathematics

William Symes
Professor of Computational and Applied Mathematics

Bala Ramaswamy
Assistant Professor of Mechanical Engineering and Materials Science

Houston, Texas
April, 1994

Abstract

Some Domain Decomposition and Multigrid Preconditioners for Hybrid Mixed Finite Elements

by

Lawrence Charles Cowsar

Discretizations of self-adjoint, linear, second-order, uniformly elliptic partial differential equations by hybrid mixed finite elements lead to large, ill-conditioned saddle-point problems. By eliminating the flux variable, a reduced problem is formed that is symmetric and positive definite but still large and ill-conditioned. Several domain decomposition and multigrid preconditioners are applied to the reduced problem, and bounds on their asymptotic rates of convergence are derived.

Two Schwarz domain decomposition methods are shown to converge at least as fast asymptotically as the same methods applied to conforming linear finite element discretizations. In particular, for both the standard additive overlapping Schwarz method of Dryja and Widlund and one of the interfacial Schwarz methods of Smith, it is proven that the rates of convergence of the methods are uniformly bounded with respect to the mesh size in both two and three dimensions under standard assumptions.

Several multigrid preconditioners are constructed for the reduced problem including a generalization of a method due to Bramble, Pasciak and Xu and an adaptation of methods of Wohlmuth and Hoppe. A common feature of these multigrid methods is the use of conforming finite element spaces on the coarser grids. Uniform convergence rates are proven for most of the methods and numerical results that verify the bounds are reported.

A mixed finite element discretization of a simplified model of sediment transport in a two dimensional periodic channel is also described. The results of two simulations that employ one of the multigrid preconditioners are reported.

Acknowledgments

Dedicated to the memory of Dr. Alan Weiser

I would like to thank Roland Glowinksi, Jan Mandel and the members of my committee for their guidance in the preparation of the material presented in this thesis. I am especially grateful for the time and energy my advisor Mary Wheeler has devoted to me as both a mentor and a friend.

I am indebted to all of the faculty and students of Rice University for making both my undergraduate and graduate days the most engaging, exciting and enjoyable years of my life. A few people that deserve special mention in this regard are Pat Shopbell, Matt Biggio, Scott Davidson, Tony Kearsley, Clint Dawson, Todd Arbogast, Cecil and Beth.

Most of all, I wish to thank my parents, Don and Carol, for their constant support, unconditional love and unique appreciation of my rainbows. A note of heartfelt gratitude also goes to my brothers, Paul and Chet, and my Aunt Margaret and Uncle Carter for their encouragement.

Finally, I want to thank Victoria Sanchez for her understanding, her patience and, most of all, her love.

Contents

Illustrations

Tables

Chapter 1

Mixed and Hybrid Mixed Finite Elements

1.1 Introduction

We consider the asymptotic convergence properties of several iterative methods for the solution of the hybrid mixed finite element discretization of the following elliptic problem for p on the connected polygonal domain $\Omega \subset \mathbb{R}^n$, $n = 2, 3$, with boundary $\partial\Omega$:

$$-\nabla \cdot A \nabla p = f \quad \text{in } \Omega, \tag{1.1}$$

$$p = 0 \quad \text{on } \partial\Omega. \tag{1.2}$$

The tensor A is assumed to be uniformly positive definite, bounded and symmetric, and $f \in L^2(\Omega)$. The choice of homogeneous Dirichlet boundary conditions and a connected polygonal domain is merely for convenience. The extensions to other boundary conditions, domains with multiple components and curved boundaries are straightforward.

Though not initially viewed as such, mixed finite element methods have in fact been used since the 1950's in their earliest incarnation as cell-centered finite differences as pointed out by Russell and Wheeler [72]. Because mixed discretizations possess inherent mass conservation properties and yield high quality approximations to both the scalar variable p and its flux $-A\nabla p$, higher order mixed finite element methods, as well as cell-centered finite differences, continue to be used to discretize elliptic problems such as (1.1)–(1.2) arising in industry. One such example is the work of Durlofsky and Chien [36] in which mixed finite element methods are applied to a problem arising in reservoir simulation.

Like standard conforming Galerkin discretizations, the use of a mixed finite element method leads to a large ill-conditioned linear system. For most problems of interest, direct methods, including sparse direct methods, are not applicable because of the size of the problem. In such cases, preconditioned iterative methods are needed.

Unlike the standard conforming Galerkin discretization in which the linear problem is symmetric and positive definite, the linear problems arising from the mixed and hybrid mixed discretizations are saddle point problems which are symmetric but indefinite. In general, iterative methods seem to work best for symmetric positive definite systems; therefore, in Section 1.3 the saddle point problem is reduced to one that is symmetric and positive definite defined solely in terms of the approximations to the scalar variable p.

The construction and analysis of effective preconditioners for this positive definite problem is the focus of this thesis. We construct preconditioners suitable for use with the conjugate gradient method [46, 70, 43]. Recall that the error (measured in the energy norm) after m iterations of the conjugate gradient method is bounded by $2((\sqrt{\kappa}-1)/(\sqrt{\kappa}+1))^m$ times the initial error where κ is the condition number of the preconditioned operator (see, e.g., [45]). We derive upper bounds on the condition number of the preconditioned operator, and hence, bounds on the rate of convergence of the conjugate gradient method. In particular, we construct domain decomposition and multigrid preconditioners with rates of convergence that do not deteriorate or deteriorate very slowly under refinement of the mesh on which the mixed finite element method is defined.

The key tool in the analysis of the preconditioners is an isomorphism between the hybrid mixed finite element space and a space of functions that are continuous and piecewise linear. The isomorphism between the two function spaces was first used by the author in the analysis of a substructuring domain decomposition method in joint work with Mandel and Wheeler [25] and is recalled with proof in Section 1.4.

In Chapter 2, the isomorphism is used as a theoretical tool to analyze the asymptotic rate of convergence of the application to hybrid mixed finite elements of two Schwarz domain decomposition methods: the overlapping additive method due to Dryja and Widlund [33] and Nepomnyaschikh [61] and a substructuring Schwarz method due to Smith [78]. Using the isomorphism, the analysis of these Schwarz methods follows with only modest modification from the existing theory for the methods applied to conforming piecewise linear discretizations.

We use the isomorphism and a fictitious domain lemma of Nepomnyaschikh [62] in Chapter 3 to show how existing implementations of preconditioners for conforming piecewise linear elements can be used to precondition hybrid mixed finite element discretizations with the same asymptotic effectiveness. A multigrid method based on this idea is implemented and compared with other standard algebraic preconditioners and a second multigrid method based on an extension of an algorithm proposed by Bramble, Pasciak and Xu in [12]. To avoid possible confusion, we note that the second multigrid algorithm is not the popular BPX multilevel preconditioner of [13] defined by the same authors. Both multigrid methods use conforming piecewise bilinear functions for the coarser spaces and are implemented using Dendy's Black Box Multigrid code [28, 29].

While the focus of this thesis is the construction of preconditioners for the hybrid mixed finite element discretizations, the techniques used in Chapters 2 and 3 can be used in a straightforward manner to construct and analyze preconditioners for several other types of discretizations of second order elliptic problems. The extended range of applicability is discussed in more detail in Chapter 4.

In Chapter 5, a highly simplified model for the evolution of erodible beds of sediments in channels under the influence of bed-load transport is presented. A numerical method is proposed, and the multigrid method used in Chapter 3 is applied to the resulting linear systems. The results of two simulations are reported.

1.2 Hybrid Mixed Finite Element Discretizations

In order to set some notation, we recall from [25] the formulation of the mixed and hybrid mixed finite element method. Readers who are familiar with hybrid mixed finite elements may simply wish to skim this section to set some notation. We refer the readers who are unfamiliar with mixed methods and their hybrid formulation to the more complete expositions by Roberts and Thomas [71] and Brezzi and Fortin [19] for more detail.

Let dx denote the standard Lebesgue n-dimensional measure and ds the $(n-1)$-dimensional surface measure. For a bounded open set $\Omega \subseteq \mathbb{R}^n$, let $|\Omega|$ denote the measure of the set, $\overline{\Omega}$ its closure and ν_Ω its outward directed normal. Let $L^2(\Omega)$, $(L^2(\Omega))^n$, $L^2(\partial\Omega)$, $H^s(\Omega)$, $(H^s(\Omega))^n$, $H^s(\partial\Omega)$ denote the standard Sobolev spaces of real-valued functions defined on Ω and $\partial\Omega$ (see, e.g., [1, 50]). We denote the natural semi-norms on $H^s(\Omega)$ and $H^s(\partial\Omega)$ by $|\cdot|_{s,\Omega}$ and $|\cdot|_{s,\partial\Omega}$, respectively. Let $\mathbf{H}(\Omega;\mathrm{div})$ denote the subspace of functions in $(L^2(\Omega))^n$ with divergences in $L^2(\Omega)$.

Let $\mathcal{T}_h$ be a quasi-regular triangulation of the polygonal domain Ω with characteristic mesh size h. The elements of $\mathcal{T}_h$ are not limited to triangles (tetrahedra in 3-D), but can, more generally, include rectangles, parallelograms, and rectangular solids. Denote by $\partial\mathcal{T}_h$ the set of edges if $n = 2$ or faces if $n = 3$ of $\mathcal{T}_h$.

The mixed and hybrid mixed finite elemenet spaces admit a standard element-wise construction based on spaces defined on a reference element. Let $\tilde{\tau}$ be a fixed reference element, and let

$$W_h(\tilde{\tau}) \times \mathbf{V}_h(\tilde{\tau}) \subset L^2(\tilde{\tau}) \times \mathbf{H}(\tilde{\tau}; \mathrm{div})$$

be one of the mixed finite element spaces of fixed degree from one of the following families of elements: the RTN spaces [69, 60], the BDM spaces [18, 16], and the BDFM spaces [17]. Without loss of generality, we assume that the reference element is chosen to be a regular polygon or polyhedron. By the assumed regularity of the mesh (see, e.g., [22]), there exists a family of bijective affine maps from the reference element to the elements in the triangulation $\mathcal{T}_h$ of the form

$$\{F_\tau \equiv b_\tau + B_\tau \tilde{x} : \tilde{\tau} \rightarrow \tau, \tau \in \mathcal{T}_h\}$$

satisfying

$$|\mathrm{det} B_\tau| = |\tau|/|\tilde{\tau}|, \quad \|B_\tau\| \leq C|\tau|^{1/n}, \quad \|B_\tau^{-1}\| \leq C|\tau|^{-1/n} \quad \tau \in \mathcal{T}_h. \tag{1.3}$$

Here, and throughout this paper, C will denote a generic positive constant, not necessarily the same from line to line, but always independent of the mesh parameter h.

For each element, let

$$W_h(\tau) = \{p \mid p = \tilde{p} \circ F_\tau^{-1}, \tilde{p} \in W_h(\tilde{\tau})\},$$

$$\mathbf{V}_h(\tau) = \{\mathbf{v} \mid \mathbf{v} = |\mathrm{det} B_\tau|^{-1} B_\tau \tilde{\mathbf{v}} \circ F_\tau^{-1}, \tilde{\mathbf{v}} \in \mathbf{V}_h(\tilde{\tau})\};$$

and for $\omega \subset \Omega$, the union of elements of $\mathcal{T}_h$, let

$$W_h(\omega) = \{p \in L^2(\Omega) \mid \mathrm{support}(p) \subset \overline{\omega} \text{ and } p_{|\tau} \in W_h(\tau)\ \forall \tau \in \mathcal{T}_h\},$$

$$\widetilde{\mathbf{V}}_h(\omega) = \{\mathbf{v} \in (L^2(\Omega))^n \mid \mathrm{support}(\mathbf{v}) \subset \overline{\omega} \text{ and } \mathbf{v}_{|\tau} \in \mathbf{V}_h(\tau)\ \forall \tau \in \mathcal{T}_h\},$$

$$\Lambda_h(\omega) = \left\{ \mu \in L^2 \left(\bigcup_{e \in \partial \mathcal{T}_h,\ e \subset \overline{\omega}} e \right) \mid \mu_{|e} \in (\widetilde{\mathbf{V}}_h(\omega) \cdot \nu)_{|e}\ \forall e \in \partial \mathcal{T}_h \right\}.$$

Let $\Lambda_h^0(\Omega)$ denote the set of functions in $\Lambda_h(\Omega)$ that vanish on the boundary of Ω, and let $\mathbf{V}_h(\Omega)$ denote the subset of $\widetilde{\mathbf{V}}_h(\Omega)$ composed of functions such that the normal component is continuous across the boundaries of elements. The continuity of the normal component and the fact that $\mathbf{V}_h(\tau) \subset \mathbf{H}(\tau; \mathrm{div})$ insure that $\mathbf{V}_h(\Omega) \subset \mathbf{H}(\Omega; \mathrm{div})$. Note also that if $\mathbf{v} \in \widetilde{\mathbf{V}}_h(\Omega)$, then $\mathbf{v} \in \mathbf{V}_h(\Omega)$ if, and only if,

$$\sum_{\tau \in \mathcal{T}_h} \int_{\partial \tau} \mu \mathbf{v} \cdot \nu_\tau \, ds = 0 \quad \forall \mu \in \Lambda_h^0(\Omega). \tag{1.4}$$

The analysis presented in subsequent sections will be applicable to all the elements in the RTN, BDM and BDFM families of mixed finite elements. We will, however, use the RT elements of lowest order defined on rectangles as an example to make concrete some of our constructions. Recall that the lowest order RT space consists of

$$W_h(\tilde{\tau}) = \mathrm{span}\,\{1\},$$

$$\mathbf{V}_h(\tilde{\tau}) = \mathrm{span}\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} x \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ y \end{pmatrix} \right\}.$$

For this space, $\Lambda_h(\Omega)$ is the space a functions that are piecewise constant on each edge in $\partial \mathcal{T}_h$.

A weak form conducive to approximation by mixed finite elements is arrived at by first rewriting (1.1) as the first order system

$$A^{-1}\mathbf{u} + \nabla p = 0 \quad \text{in } \Omega, \tag{1.5}$$

$$\nabla \cdot \mathbf{u} = f \quad \text{in } \Omega. \tag{1.6}$$

After multiplying by appropriate test functions, integrating the second term in (1.5) by parts and using the boundary condition (1.2), we arrive at the problem of finding

$$\{\mathbf{u}, p\} \in \mathbf{H}(\Omega; \mathrm{div}) \times L^2(\Omega)$$

such that

$$\int_\Omega A^{-1}\mathbf{u} \cdot \mathbf{v} \, dx - \int_\Omega p \nabla \cdot \mathbf{v} \, dx = 0 \quad \forall \mathbf{v} \in \mathbf{H}(\Omega; \mathrm{div}), \tag{1.7}$$

$$\int_\Omega q \nabla \cdot \mathbf{u} \, dx = \int_\Omega q f \, dx \quad \forall q \in L^2(\Omega). \tag{1.8}$$

We now introduce two equivalent finite dimensional approximations of (1.7)–(1.8). By the *mixed finite element approximation* to (1.7)–(1.8), we mean the pair

$$\{\mathbf{u}_h, p_h\} \in \mathbf{V}_h(\Omega) \times W_h(\Omega)$$

satisfying

$$\int_\Omega A^{-1}\mathbf{u}_h \cdot \mathbf{v}\,dx - \int_\Omega p_h \nabla\cdot\mathbf{v}\,dx = 0 \quad \forall \mathbf{v} \in \mathbf{V}_h(\Omega), \tag{1.9}$$

$$\int_\Omega q\nabla\cdot\mathbf{u}_h\,dx = \int_\Omega qf\,dx \quad \forall q \in W_h(\Omega). \tag{1.10}$$

By the *hybrid mixed finite element approximation* to (1.1)–(1.2), we mean the triple

$$\{\mathbf{u}_h, p_h, \lambda_h\} \in \widetilde{\mathbf{V}}_h(\Omega) \times W_h(\Omega) \times \Lambda_h^0(\Omega)$$

satisfying

$$\sum_{\tau\in\mathcal{T}_h} \left(\int_\tau A^{-1}\mathbf{u}_h \cdot \mathbf{v}\,dx - \int_\tau p_h\nabla\cdot\mathbf{v}\,dx + \int_{\partial\tau} \lambda_h \mathbf{v}\cdot\nu_\tau\,ds\right) = 0 \quad \forall \mathbf{v} \in \widetilde{\mathbf{V}}_h(\Omega), \tag{1.11}$$

$$-\sum_{\tau\in\mathcal{T}_h} \int_\tau q\nabla\cdot\mathbf{u}_h\,dx = -\int_\Omega qf\,dx \quad \forall q \in W_h(\Omega), \tag{1.12}$$

$$\sum_{\tau\in\mathcal{T}_h} \int_{\partial\tau} \mu\mathbf{u}_h\cdot\nu_\tau\,ds = 0 \quad \forall \mu \in \Lambda_h^0(\Omega). \tag{1.13}$$

For issues related to the well-posedness and convergence properties of these approximations, we refer the reader to the expositions in [30, 71, 19] and [69, 60, 18, 16, 17] in which the families of elements were first defined.

In the hybrid mixed finite element method, we have relaxed the continuity condition on the space of fluxes, replacing $\mathbf{V}_h(\Omega)$ with $\widetilde{\mathbf{V}}_h(\Omega)$, and imposed it variationally through (1.13). The equivalence of (1.11)–(1.13) with (1.9)–(1.10) is a simple consequence of (1.4). The relaxation of the continuity requirement was proposed by Fraeijs de Veubeke [39] and analyzed carefully in the context of mixed methods by Arnold and Brezzi [4].

Note that the hybrid formulation (1.11)–(1.13) is just the first order necessary conditions from the constrained energy form

$$\min_{\substack{\mathbf{v}\in\widetilde{\mathbf{V}}_h(\Omega),\\ \text{s.t. } \mathbf{v}\in\mathbf{H}(\Omega;\mathrm{div}),\\ \text{and } \nabla\cdot\mathbf{v} = P_{W_h(\Omega)}f}} \int_\Omega A^{-1}\mathbf{v}\cdot\mathbf{v}\,dx,$$

where $P_{W_h(\Omega)}$ is $L^2(\Omega)$-projection onto $W_h(\Omega)$. Adopting some terminology from optimization, we refer to the functions in $\widetilde{\mathbf{V}}_h(\Omega)$ as the *primal variables* and functions in $W_h(\Omega) \times \Lambda_h(\Omega)$ as the *dual variables.* We see that p_h is nothing more than the Lagrange multiplier associated with the constraint $\nabla\cdot\mathbf{u}_h = P_{W_h(\Omega)}f$, and λ_h is the multiplier associated with the constraint that $\mathbf{v} \in \mathbf{H}(\Omega; \mathrm{div})$, i.e. the continuity of fluxes across edges express in (1.4). Equation (1.11) is also a weak form of the constitutive relationship (1.5) which, after multiplying by a test function and integrating the second term by parts, yields

$$\int_\tau A^{-1}\mathbf{u} \cdot \mathbf{v}\, dx - \int_\tau p\nabla\cdot\mathbf{v}\, dx + \int_{\partial\tau} p\mathbf{v}\cdot\nu_\tau\, ds = 0. \tag{1.14}$$

Comparing with (1.11), we see that λ_h is naturally interpreted as an approximation to the trace of p on the boundaries of the elements.

The RTN, BDM and BDFM mixed finite element families share several common properties that we will use in our analysis. Primarily, we use the fact that the spaces $W_h(\Omega)$ and $\Lambda_h(\Omega)$ admit natural nodal bases defined by degrees of freedom in the interior of elements for $W_h(\Omega)$ and on the boundary of elements for $\Lambda_h(\Omega)$. For example, one can take the nodes to be the center of the element for $W_h(\Omega)$ and the center of each edge for $\Lambda_h(\Omega)$ for the lowest order RT space defined above. Other examples can be found in [44]. Because of (1.14), the values attained at these nodes have the natural interpretation as values of the scalar variable at the nodes, in the interior as well as on the edges (faces in 3D) of the elements. We will not differentiate between functions in $W_h(\Omega) \times \Lambda_h(\Omega)$ and the values they attain at the nodal degrees of freedom.

Additionally, the following two properties satisfied by the RTN, BDM and BDFM families are used to prove Theorem 1.1 in Section 1.3:

- the approximating spaces for the scalar variable and its flux are related by the inclusion

$$\mathrm{div}(\mathbf{V}_h(\tau)) \subseteq W_h(\tau),$$

  and

- there exists a projection

$$\Pi_h : \mathbf{H}(\tau; \mathrm{div}) \cap \{\mathbf{v}\cdot\nu_\tau \in L^2(\partial\tau),\ \tau \in \mathcal{T}_h\} \to \mathbf{V}_h(\tau)$$

  that satisfies, among other properties, that for every $\tau \in \mathcal{T}_h$ and edge (face in 3-D) e_i of the boundary of τ that

$$\int_{e_i} (\Pi_h\mathbf{u} - \mathbf{u})\cdot\nu_\tau\mathbf{v}\cdot\nu_\tau\, ds = 0 \quad \forall\mathbf{v} \in \mathbf{V}_h(\tau), \tag{1.15}$$

$$\int_\tau \nabla\cdot(\Pi_h\mathbf{u} - \mathbf{u})q\, dx = 0 \quad \forall q \in W_h(\tau). \tag{1.16}$$

1.3 The Dual Variable Problem

Henceforth, we are concerned only with the solution of the finite dimensional problem (1.11)–(1.13) and drop the subscript h from $\mathbf{u}_h$, p_h and λ_h.

The hybrid mixed finite element problem (1.11)–(1.13) is symmetric, but indefinite. In applying iterative methods, it is often advantageous to reduce to a symmetric positive definite form if one can accomplish that at modest cost. To that end, we eliminate the flux variable $\mathbf{u}$ in (1.11)–(1.13) by introducing a discretization of the flux operator $A\nabla$ denoted

$$\nabla_h^A : W_h(\Omega) \times \Lambda_h(\Omega) \to \widetilde{\mathbf{V}}_h(\Omega),$$

and defined by

$$\sum_{\tau\in\mathcal{T}_h} \int_\tau A^{-1}\nabla_h^A[q,\mu]\cdot\mathbf{v}\,dx = \sum_{\tau\in\mathcal{T}_h}\Big(-\int_\tau q\nabla\cdot\mathbf{v}\,dx + \int_{\partial\tau}\mu\mathbf{v}\cdot\nu_\tau\,ds\Big) \quad \forall\mathbf{v}\in\widetilde{\mathbf{V}}_h(\Omega). \tag{1.17}$$

Since $\widetilde{\mathbf{V}}_h(\Omega)$ is the direct sum of spaces defined on each element, we note that (1.17) holds element by element. Hence, $\nabla_h^A[q,\mu]$ is defined element-wise in terms of the values of q and μ restricted to $\overline{\tau}$. By restriction we consider ∇_h^A as a map from $W_h(\omega)\times\Lambda_h(\omega)$ into $\widetilde{\mathbf{V}}_h(\omega)$ for any set $\omega\subset\Omega$ which is the union of elements of $\mathcal{T}_h$.

Define bilinear forms

$$d_\omega : [W_h(\omega)\times\Lambda_h(\omega)]\times[W_h(\omega)\times\Lambda_h(\omega)] \to \mathbb{R}$$

by

$$d_\omega([p,\lambda],[q,\mu]) \equiv \sum_{\substack{\tau\in\mathcal{T}_h,\\ \tau\subset\omega}} \int_\tau A^{-1}\nabla_h^A[p,\lambda]\cdot\nabla_h^A[q,\mu]\,dx. \tag{1.18}$$

By the *dual variable problem*, we mean finding the pair

$$[p,\lambda]\in W_h(\Omega)\times\Lambda_h^0(\Omega)$$

satisfying

$$d_\Omega([p,\lambda],[q,\mu]) = \int_\Omega fq\,dx \quad \forall[q,\mu]\in W_h(\Omega)\times\Lambda_h^0(\Omega). \tag{1.19}$$

The dual variable problem is nothing more than the variational equivalent of forming the Schur complement of (1.11)–(1.13) with respect to the dual variables p and λ. Hence, solving the dual variable problem is equivalent to solving (1.11)–(1.13), a fact that is demonstrated in [4], and the flux can be recovered as $\mathbf{u} = -\nabla_h^A[p,\lambda]$.

Moreover, the bilinear form $d_\Omega(\cdot,\cdot)$ is obviously symmetric and its positive definiteness on $W_h(\Omega) \times \Lambda_h^0(\Omega)$ is a simple corollary of Theorem 1.1 below.

The evaluation of the bilinear form $d_\omega([p,\lambda],[q,\mu])$ might appear prohibitively expensive, requiring the calculation of both $\nabla_h^A[p,\lambda]$ and $\nabla_h^A[q,\mu]$; however, by using (1.18), we see that the bilinear form can be computed as

$$d_\omega([p,\lambda],[q,\mu]) = \sum_{\substack{\tau \in \mathcal{T}_h, \\ \tau \subset \omega}} \left(-\int_\tau q \nabla\cdot(\nabla_h^A[p,\lambda])\, dx + \int_{\partial\tau} \mu(\nabla_h^A[p,\lambda])\cdot\nu_\tau\, ds \right).$$

Moreover, since there are no continuity requirements across elements for functions in $\widetilde{\mathbf{V}}_h(\Omega)$, $\nabla_h^A[p,\lambda]$ can be constructed element by element.

Recall that two quadratic forms, $\mathcal{Q}_1$ and $\mathcal{Q}_2$, with domain $\mathcal{D}$ are said to be *equivalent* if there exist constants $c_1, c_2 > 0$ such that

$$c_1\mathcal{Q}_1(\phi,\phi) \le \mathcal{Q}_2(\phi,\phi) \le c_2\mathcal{Q}_1(\phi,\phi) \quad \forall \phi \in \mathcal{D}.$$

We will denote this equivalence by $\mathcal{Q}_1 \simeq \mathcal{Q}_2$. The constants suppressed by this notation will always be independent of the mesh parameter h, but can, unless explicitly stated to the contrary, depend on the minimum and maximum eigenvalues of the coefficient A in (1.1), the choice of family and degree of the mixed finite elements, and the regularity of the triangulation $\mathcal{T}_h$.

The following theorem relates the quadratic form induced by $d_\omega(\cdot,\cdot)$ to an equivalent quadratic form in terms of the nodal degrees of freedom. In stating the equivalence more sharply, we decompose A as

$$A(x) = \alpha(x)\widehat{A}(x)$$

where α is an arbitrary piecewise constant function on each element of the triangulation. By the uniform positivity of A, there exist positive constants C_1 and C_2 such that

$$C_1\xi^t\xi \le \xi^t\widehat{A}(x)\xi \le C_2\xi^t\xi \quad \forall x \in \Omega,\ \xi \in \mathbb{R}^n. \tag{1.20}$$

An appropriate choice for α is the average eigenvalue of A over each cell; that is,

$$\alpha_{|\tau} = \frac{1}{n|\tau|}\int_\tau \operatorname{trace}(A(x))\, dx.$$

Theorem 1.1 *(Theorem 4.1 of [25])* Let $\widehat{p} = [p,\lambda] \in W_h(\omega) \times \Lambda_h(\omega)$ for $\omega \subseteq \Omega$ composed of elements of the triangulation $\mathcal{T}_h$. Then

$$d_\omega(\widehat{p},\widehat{p}) \simeq \sum_{\tau\in\mathcal{T}_h,\, \tau\subset\omega} \alpha_{|\tau}|\tau|^{1-2/n} \sum_{\substack{\textbf{nodes}: \\ n_i, n_j \in \overline{\tau}}} (\widehat{p}(n_i) - \widehat{p}(n_j))^2 \tag{1.21}$$

with constants independent of h, α, A and $|\omega|$, but depending on the eigenvalues of $\widehat{A}$, the choice of the mixed finite element space, and the regularity of the triangulation $\mathcal{T}_h$.

The proof of Theorem 1.1 is a direct consequence of the following two lemmas which we recall along with their proofs from [25] for completeness.

Lemma 1.1 *(Lemma 4.2 of [25])* The kernel of $d_\tau(\cdot,\cdot)$ consists of the constant functions on $\overline{\tau}$; that is,

$$d_\tau([p,\lambda],[q,\mu]) = 0 \quad \forall [q,\mu] \in W_h(\tau) \times \Lambda_h(\tau), \tag{1.22}$$

if, and only if, $[p,\lambda]$ has the same value on all nodes of $\overline{\tau}$.

Proof We first check that if $[p,\lambda] = [K,K]$ for some constant K on $\overline{\tau}$, then it is in the kernel of $d_\tau(\cdot,\cdot)$. Letting $\mathbf{v} = \nabla_h^A[q,\mu] \in \mathbf{V}_h(\tau)$, we see that

$$\begin{aligned} d_\tau([p,\lambda],[q,\mu]) &= \int_\tau A^{-1}\nabla_h^A[p,\lambda]\cdot \mathbf{v}\,dx \\ &= -\int_\tau p\nabla\cdot\mathbf{v}\,dx + \int_{\partial\tau} \lambda \mathbf{v}\cdot\nu_\tau \\ &= K\left[-\int_\tau \nabla\cdot\mathbf{v}\,dx + \int_{\partial\tau}\mathbf{v}\cdot\nu\,ds\right], \end{aligned} \tag{1.23}$$

which is zero by the Divergence Theorem.

To prove the converse, let $[\tilde{p},\tilde{\lambda}] \in W_h(\tau)\times\Lambda_h(\tau)$ be such that

$$d_\tau([\tilde{p},\tilde{\lambda}],[q,\mu]) = 0 \quad \forall [q,\mu] \in W_h(\tau)\times\Lambda_h(\tau). \tag{1.24}$$

It is enough to show that $[\tilde{p},\tilde{\lambda}]$ is zero if it is orthogonal to constants, i.e.

$$\int_\tau K\tilde{p}\,dx + \int_{\partial\tau} K\tilde{\lambda}\,ds = 0 \quad \forall K \in \mathbb{R}. \tag{1.25}$$

By (1.24),

$$\int_\tau A^{-1}\nabla_h^A[\tilde{p},\tilde{\lambda}]\cdot\nabla_h^A[\tilde{p},\tilde{\lambda}]\,dx = 0,$$

and hence,

$$\nabla_h^A[\tilde{p},\tilde{\lambda}] = 0. \tag{1.26}$$

Since (1.25) holds, there exists a solution ϕ (unique up to a constant) to the Neumann problem

$$-\Delta\phi = \tilde{p} \quad \text{in } \tau, \tag{1.27}$$

$$\nabla\phi\cdot\nu = \tilde{\lambda} \quad \text{on } \partial\tau. \tag{1.28}$$

Setting $\tilde{\mathbf{v}} = \Pi_h \nabla\phi$ and using properties (1.15) and (1.16), we see that $\tilde{\mathbf{v}} \in \mathbf{V}_h(\tau)$ with $\nabla\cdot\tilde{\mathbf{v}} = -\tilde{p}$ and $\tilde{\mathbf{v}}\cdot\nu = \tilde{\lambda}$. Using $\tilde{\mathbf{v}}$ in (1.17) with (1.26), we have

$$0 = -\int_\tau \tilde{p}\nabla\cdot\tilde{\mathbf{v}}\,dx + \int_{\partial\tau} \tilde{\lambda}\tilde{\mathbf{v}}\cdot\nu_\tau\,ds = \int_\tau \tilde{p}^2\,dx + \int_{\partial\tau} \tilde{\lambda}^2\,ds; \tag{1.29}$$

therefore, $[\tilde{p}, \tilde{\lambda}]$ is zero. □

The following lemma was suggested by Joseph Pasciak [67].

Lemma 1.2 *(Lemma 4.3 of [25])* For any $\tau \in \mathcal{T}_h$ and for all $\hat{p} = [p, \lambda] \in W_h(\tau) \times \Lambda_h(\tau)$

$$d_\tau(\hat{p}, \hat{p}) \simeq \alpha_{|\tau}|\tau|^{1-2/n} \sum_{\substack{\text{nodes}: \\ n_i, n_j \in \tau}} (\hat{p}(n_i) - \hat{p}(n_j))^2, \tag{1.30}$$

with constants independent of h, α, A and the particular choice of τ, but depending on the eigenvalues of $\hat{A}$, the choice the mixed finite element space, and the regularity of $\mathcal{T}_h$.

Proof Let $\tilde{\tau}$ denote the reference element, $W_h(\tilde{\tau})$ and $\mathbf{V}_h(\tilde{\tau})$ the reference spaces, and $F_\tau(\tilde{x}) = b_\tau + B_\tau\tilde{x}$ the affine mapping of $\tilde{\tau}$ onto τ introduced in Section 1.2. For $[\tilde{q}, \tilde{\mu}] \in W_h(\tilde{\tau}) \times \Lambda_h(\tilde{\tau})$, define $\widetilde{\nabla}_h[\tilde{q}, \tilde{\mu}] \in \mathbf{V}_h(\tilde{\tau})$ by

$$\int_{\tilde{\tau}} \widetilde{\nabla}_h[\tilde{q}, \tilde{\mu}] \cdot \tilde{\mathbf{v}}\,dx = -\int_{\tilde{\tau}} \tilde{q}\nabla\cdot\tilde{\mathbf{v}}\,dx + \int_{\partial\tilde{\tau}} \tilde{\mu}\tilde{\mathbf{v}}\cdot\nu_{\tilde{\tau}}\,ds \quad \forall\tilde{\mathbf{v}} \in \mathbf{V}_h(\tilde{\tau}). \tag{1.31}$$

Using Lemma 1.1, one has

$$\int_{\tilde{\tau}} \widetilde{\nabla}_h[\tilde{q}, \tilde{\mu}] \cdot \widetilde{\nabla}_h[\tilde{q}, \tilde{\mu}]\,dx \simeq \sum_{\substack{\text{nodes}: \\ n_i, n_j \in \tilde{\tau}}} ([\tilde{q}, \tilde{\mu}](n_i) - [\tilde{q}, \tilde{\mu}](n_j))^2 \tag{1.32}$$

since both quadratic forms induce norms on $W_h(\tilde{\tau}) \times \Lambda_h(\tilde{\tau})$ modulo constant functions with equivalence constants depending only on the reference element and choice of mixed finite element space.

Letting $J_\tau = |\det B_\tau|$, define the following functions on the reference element

$$\tilde{p} = p \circ F_\tau, \qquad \tilde{\lambda} = \lambda \circ F_\tau, \quad \tilde{\mathbf{u}} = (J_\tau B_\tau^{-1}\nabla_h^A[p, \lambda]) \circ F_\tau, \quad \tilde{A} = A \circ F_\tau.$$

Under this change of variables for the mixed finite element spaces (see, [82, 19]), we have $\tilde{\mathbf{u}} \in \mathbf{V}_h(\tilde{\tau})$ satisfying (1.17) in the form

$$\int_{\tilde{\tau}} \frac{1}{J_\tau} B_\tau^t \tilde{A}^{-1} B_\tau \tilde{\mathbf{u}} \cdot \tilde{\mathbf{v}}\,dx = -\int_{\tilde{\tau}} \tilde{p}\nabla\cdot\tilde{\mathbf{v}}\,dx + \int_{\partial\tilde{\tau}} \tilde{\lambda}\tilde{\mathbf{v}}\cdot\nu_{\tilde{\tau}}\,ds \quad \forall\tilde{\mathbf{v}} \in \mathbf{V}_h(\tilde{\tau}), \tag{1.33}$$

and

$$d_\tau(\widehat{p},\widehat{p}) = \int_\tau A^{-1}\nabla_h^A[p,\lambda]\cdot\nabla_h^A[p,\lambda]\,dx = \int_{\tilde{\tau}} \frac{1}{J_\tau} B_\tau^t \tilde{A}^{-1} B_\tau \tilde{\mathbf{u}}\cdot\tilde{\mathbf{u}}\,dx. \tag{1.34}$$

Comparing (1.31) and (1.33), we see that

$$\int_{\tilde{\tau}} \frac{1}{J_\tau} B_\tau^t \tilde{A}^{-1} B_\tau \tilde{\mathbf{u}}\cdot\tilde{\mathbf{v}}\,dx = \int_{\tilde{\tau}} \widetilde{\nabla}_h[\tilde{p},\tilde{\lambda}]\cdot\tilde{\mathbf{v}}\,dx \quad \forall\tilde{\mathbf{v}}\in\mathbf{V}_h(\tilde{\tau}).$$

Using Cauchy-Schwarz with $\mathbf{v}=\tilde{\mathbf{u}}$ and $\mathbf{v}=\widetilde{\nabla}_h[\tilde{p},\tilde{\lambda}]$, we deduce that

$$\int_{\tilde{\tau}} \frac{1}{J_\tau} B_\tau^t \tilde{A}^{-1} B_\tau \tilde{\mathbf{u}}\cdot\tilde{\mathbf{u}}\,dx \le C_2\alpha_{|\tau} J_\tau \|B_\tau^{-1}\|^2 \int_{\tilde{\tau}} \widetilde{\nabla}_h[\tilde{p},\tilde{\lambda}]\cdot\widetilde{\nabla}_h[\tilde{p},\tilde{\lambda}]\,dx, \tag{1.35}$$

$$\int_{\tilde{\tau}} \frac{1}{J_\tau} B_\tau^t \tilde{A}^{-1} B_\tau \tilde{\mathbf{u}}\cdot\tilde{\mathbf{u}}\,dx \ge C_1\alpha_{|\tau} J_\tau \|B_\tau\|^{-2} \int_{\tilde{\tau}} \widetilde{\nabla}_h[\tilde{p},\tilde{\lambda}]\cdot\widetilde{\nabla}_h[\tilde{p},\tilde{\lambda}]\,dx, \tag{1.36}$$

where C_1 and C_2 are the positive constants in (1.20).

The proof of the lemma is completed by using (1.32), (1.34), (1.35), (1.36), and (1.3). □

1.4 An Isomorphism with a Conforming Space

In this section, we recall the construction of an isomorphism between $W_h(\omega)\times\Lambda_h(\omega)$ and a conforming space of piecewise linear functions. This isomorphism was first used in joint work with Mandel and Wheeler [25] to analyze the application of Mandel's Balancing Domain Decomposition method to hybrid mixed finite element discretizations. This type of isomorphism was also used by the author in [23] to analyze three domain decomposition methods for nonconforming elements of Lagrange type. A similar isomorphism was used independently by Sarkis in [76] in the analysis of Schwarz methods for piecewise linear nonconforming finite elements.

The isomorphism is constructed by first building a refined triangulation $\widehat{\mathcal{T}}_{\tilde{\tau}}$ of the reference element $\tilde{\tau}$. The vertices used in the refinement are the vertices of $\tilde{\tau}$ and the nodal points in $\tilde{\tau}$ pertaining to the degrees of freedom of $W_h(\tilde{\tau})\times\Lambda_h(\tilde{\tau})$. Several examples are given in Figures 1.1 and 1.2. A refinement using only these vertices is always possible. For example in two dimensions, the refined triangulation can be constructed by first partitioning $\tilde{\tau}$ into triangles by adding edges connecting some of the vertices of $\tilde{\tau}$. If the reference element is a triangle, no additional edges are added in this step. The nodal points can then be added one at a time by adding edges connecting the nodal point to the vertices of all triangles that contain the nodal point in their interior or on their boundary in the current subtriangulation. Triangulations of three-dimensional elements can be constructed in an analogous manner.

The mappings F_τ from the reference element to the elements of $\mathcal{T}_h$ induce a refined triangulation $\hat{\mathcal{T}}$ of $\mathcal{T}_h$ by

$$\hat{\mathcal{T}} = \bigcup_{\tau \in \mathcal{T}_h} F_\tau(\hat{\mathcal{T}}_{\hat{\tau}}).$$

Note that the regularity of the refined mesh $\hat{\mathcal{T}}$ is a function only of the regularity of the original mesh $\mathcal{T}_h$ and the regularity of the refinement $\hat{\mathcal{T}}_{\hat{\tau}}$ of the reference element.

A vertex of $\hat{\mathcal{T}}$ will be called *primary* if it was a nodal point corresponding to a degree of freedom of $W_h(\Omega) \times \Lambda_h(\Omega)$; otherwise, we call the vertex *secondary*. We say that two vertices of the triangulation $\hat{\mathcal{T}}$ are *adjacent* if there exists an edge of $\hat{\mathcal{T}}$ connecting the vertices.

Let $U_h(\Omega)$ denote the space of functions that are continuous, piecewise linear with respect to the refined triangulation $\hat{\mathcal{T}}$, and vanish on $\partial\Omega$. For $\omega \subset \Omega$, a union of elements in $\mathcal{T}_h$, let $U_h(\omega) \subset U_h(\Omega)$ denote those functions that vanish outside of ω. Since the functions in $U_h(\omega)$ are naturally parameterized by the values they attain at the vertices in ω, we define a pseudo-interpolation mapping $\mathcal{I}^\omega$ into $U_h(\omega)$ for any

Figure 1.1 Examples of subtriangulations of two commonly used elements

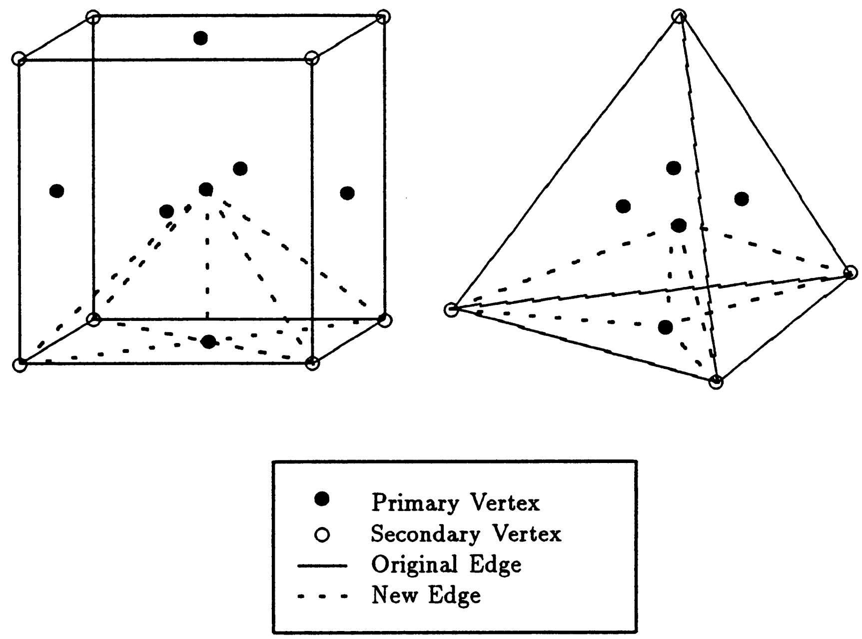

Figure 1.2 Partial subtriangulation of the lowest order Raviart-Thomas-Nedelec elements

function ϕ defined at the primary vertices contained in ω by

$$\mathcal{I}^\omega\phi(x) = \begin{cases} 0, \text{ if } x \in \partial\Omega; \\ \phi(x), \text{ if } x \text{ is a primary vertex;} \\ \text{The average of all adjacent primary vertices, if } x \text{ is a} \\ \quad \text{secondary vertex in the interior of } \omega; \\ \text{The average of all adjacent primary vertices on } \partial\omega, \text{ if} \\ \quad x \text{ is a secondary vertex on } \partial\omega; \\ \text{The continuous piecewise linear interpolant of the} \\ \quad \text{above vertex values, if } x \text{ is not a vertex of } \hat{\mathcal{T}}. \end{cases} \tag{1.37}$$

Since $\mathcal{I}^\omega$ is defined for any function defined at primary vertices, by an abuse of notation, we can understand $\mathcal{I}^\omega$ as a linear map from $W_h(\omega) \times \Lambda_h(\omega)$ into $U_h(\omega)$ and a map from $U_h(\omega)$ into $U_h(\omega)$. Let $\widetilde{U}_h(\omega) \subset U_h(\omega)$ denote the range of $\mathcal{I}^\omega$; that is,

$$\widetilde{U}_h(\omega) = \{\psi \in U_h(\omega) \mid \psi = \mathcal{I}^\omega \phi, \phi \in U_h(\omega)\}.$$

Let

$$a_\omega(\phi, \psi) = \int_\omega A\nabla\phi \cdot \nabla\psi \, dx.$$

Since $\mathcal{I}^\Omega$ is clearly a bijective map between $W_h(\Omega) \times \Lambda_h^0(\Omega)$ and $\widetilde{U}_h(\Omega)$ by construction, the following theorem with $\omega = \Omega$ proves that $\mathcal{I}^\Omega$ is in fact an isomorphism between these two spaces that preserves the natural norms for the second order elliptic problem.

Theorem 1.2 Let $\omega \subset \Omega$ be the union of elements in $\mathcal{T}_h$. Then for all $\widehat{p} \in W_h(\omega) \times (\Lambda_h(\omega) \cap \Lambda_h^0(\Omega))$,

$$d_\omega(\widehat{p}, \widehat{p}) \simeq a_\omega(\mathcal{I}^\omega \widehat{p}, \mathcal{I}^\omega \widehat{p}). \tag{1.38}$$

Proof It is easy to show (see, e.g., [22]) that for $\phi \in U_h(\omega)$,

$$a_\omega(\mathcal{I}^\omega \widehat{p}, \mathcal{I}^\omega \widehat{p}) \simeq \sum_{\substack{\tau \in \widehat{\mathcal{T}}, \\ \tau \subset \omega}} \alpha_{|\tau}|\tau|^{1-2/n} \sum_{\substack{\textbf{vertices}: \\ v_i, v_j \in \tau}} (\phi(v_i) - \phi(v_j))^2. \tag{1.39}$$

By virtue of Theorem 1.1 and (1.39), it is enough to show that

$$\sum_{\substack{\tau \in \mathcal{T}_h, \\ \tau \subset \omega}} |\tau|^{1-2/n} \sum_{\substack{\textbf{nodes}: \\ n_i, n_j \in \tau}} (\widehat{p}(n_i) - \widehat{p}(n_j))^2 \simeq$$
$$\sum_{\substack{\tau \in \widehat{\mathcal{T}}, \\ \tau \subset \omega}} |\tau|^{1-2/n} \sum_{\substack{\textbf{vertices}: \\ v_i, v_j \in \tau}} ((\mathcal{I}^\omega \widehat{p})(v_i) - (\mathcal{I}^\omega \widehat{p})(v_j))^2. \tag{1.40}$$

Since vertices of $F_\tau(\widehat{\mathcal{T}}_{\overline{\tau}})$ contain the nodal points of τ and $\widehat{p} = \mathcal{I}^\omega \widehat{p}$ at these points, we have

$$\sum_{\substack{\textbf{nodes}: \\ n_i, n_j \in \overline{\tau}}} (\widehat{p}(n_i) - \widehat{p}(n_j))^2 \le C \sum_{\widehat{\tau} \in \widehat{\mathcal{T}}_\tau} \sum_{\substack{\textbf{vertices}: \\ v_i, v_j \in \widehat{\tau}}} \left((\mathcal{I}^\omega \widehat{p})(v_i) - (\mathcal{I}^\omega \widehat{p})(v_j)\right)^2,$$

where the constant is controlled by the regularity of the subtriangulation. Hence, by summing over the elements of $\mathcal{T}_h$ in ω, we conclude that the right hand side of (1.40) dominates the left hand side.

To prove that the left hand side dominates the right, we note that the differences in the right hand side are of three types: the difference at two primary vertices, the difference at two secondary vertices, and the difference at a primary and a secondary vertex. Since $\widehat{p}$ and $\mathcal{I}^\omega\widehat{p}$ agree at primary vertices of $\widehat{\mathcal{T}}$, the difference at two primary vertices occurs as a term in the left hand side. For two secondary vertices v_1, v_2 in an element $\tau \in \widehat{\mathcal{T}}$ containing a primary vertex v_3, we see that

$$\begin{aligned}\left(\left(\mathcal{I}^\omega\widehat{p}\right)(v_1) - \left(\mathcal{I}^\omega\widehat{p}\right)(v_2)\right)^2 &\leq 2\left(\left(\mathcal{I}^\omega\widehat{p}\right)(v_1) - \left(\mathcal{I}^\omega\widehat{p}\right)(v_3)\right)^2 \\ &\quad +2\left(\left(\mathcal{I}^\omega\widehat{p}\right)(v_2) - \left(\mathcal{I}^\omega\widehat{p}\right)(v_3)\right)^2.\end{aligned}$$

Hence, it is enough to bound the difference at a secondary and primary vertex by terms in the left hand side of (1.40).

Let v_{n+1} be a secondary vertex with adjacent primary vertices $v_1, \ldots, v_n$, and let $\widehat{p}_j = \widehat{p}(v_j)$. Noting that for $j = 1, \ldots, n$

$$\left(\mathcal{I}^\omega\widehat{p}\right)(v_j) = \widehat{p}_j, \qquad \left(\mathcal{I}^\omega\widehat{p}\right)(v_{n+1}) = \frac{1}{n}\sum_{j=1}^{n}\left(\mathcal{I}^\omega\widehat{p}\right)(v_j) = \frac{1}{n}\sum_{j=1}^{n}\widehat{p}_j,$$

we see that

$$\left(\left(\mathcal{I}^\omega\widehat{p}\right)(v_{n+1}) - \left(\mathcal{I}^\omega\widehat{p}\right)(v_i)\right)^2 = \frac{1}{n^2}\left(\sum_{j=1}^{n}(\widehat{p}_j - \widehat{p}_i)\right)^2 \leq \frac{n}{n^2}\sum_{j=1}^{n}(\widehat{p}_j - \widehat{p}_i)^2,$$

by the Cauchy-Schwarz inequality. The proof is completed by summing over all triangles of $\widehat{\mathcal{T}}$. The number of such terms, and hence the constant in the bound, is controlled since the regularity of the mesh implies that there is an a priori maximum number of adjacent elements that can share a secondary point. □

We have constructed the isomorphism above using piecewise linear functions on a refined mesh composed of triangles and tetrahedra. For the rectangular and right-angled hexahedral elements, a more natural construction can be made using continuous piecewise bilinear and trilinear functions defined on a refined mesh of rectangular and right-angled hexahedral elements. Clearly, Theorem 1.2 holds in this case as well. This variant is used in the multigrid methods of Chapter 3.

Using the techniques in the proof of Theorem 1.2, the following lemma is also easy to prove.

Lemma 1.3 There exists a constant $C > 0$ independent of h, A, and α, but depending on $\widehat{A}$, the regularity of the triangulation $\mathcal{T}_h$ and the choice of mixed finite element space such that

$$a_\omega(\mathcal{I}^\omega\phi, \mathcal{I}^\omega\phi) \le C a_\omega(\phi,\phi) \quad \forall\phi \in U_h(\Omega). \tag{1.41}$$

Additionally, there exists a constant $\widetilde{C}$ also independent of A such that

$$|\mathcal{I}^\omega\phi|_{k,\omega} \le \widetilde{C}|\phi|_{k,\omega} \quad \forall\phi \in U_h(\Omega),\ k = 0, 1. \tag{1.42}$$

Chapter 2

Schwarz Domain Decomposition Methods for Hybrid Mixed Finite Elements

2.1 Other Domain Decomposition Methods for Mixed Finite Elements

In this section, we summarize previous work on domain decomposition methods for mixed finite elements. For a more comprehensive survey of domain decomposition methods and their theory, we refer the interested reader to the recent review articles of Chan and Mathew [20] and Le Tallec [48].

The first domain decomposition methods for the solution of the linear systems arising from mixed finite element discretizations were two substructuring methods introduced by Glowinski and Wheeler [42]. The first method of Glowinski and Wheeler involves solving a reduced problem defined in terms of the primal variables on the interface between subdomains; the second involves an interface problem defined solely in terms of the dual variables. Since these two methods are in fact Schur complement methods involving neither a coarse space nor further preconditioning, they result in operators with condition numbers that depend strongly on the size of the mesh and subdomains. In particular, if we set a length scale by choosing the diameter of Ω to be 1, and let H and h denote the characteristic sizes of the subdomains and mesh spacing, respectively, then the operators defined by Glowinski and Wheeler have condition numbers that are $O((hH)^{-1})$ under the standard assumptions regarding the regularity of the mesh and the decomposition of Ω into subdomains made explicit in the next section.

Multigrid strategies were applied to the primal and dual substructuring operators to accelerate convergence in [41] and [27]. While there currently exists no convergence theory for the multigrid accelerated substructuring methods for the mixed finite elements, a comparison with a parallel implementation of semi-coarsening multigrid reported in [26] showed the multigrid accelerated substructuring method to yield a competitive algorithm. The method has also been used in [7].

In [73, 75], Rusten and Winther constructed two preconditioners for the saddle-point problem (1.9)–(1.10) in two dimensions that use as key components an iteration of the primal substructuring method and a modification of the dual substructuring method of Glowinski and Wheeler for Laplace's equation. Since the spectrum of the saddle point problem includes both positive and negative eigenvalues, the rate of convergence of the Krylov method they chose to use, the minimal residual algorithm, with their preconditioner is not explicitly available (see [74]). Their theory and experiments indicate that their primal variable preconditioner results in an algorithm whose convergence rate depends strongly on the mesh spacing h, and their dual variable preconditioner results in an algorithm whose rate of convergence is independent of h. However, if used with many subdomains, both methods of Rusten and Winther would depend strongly on H since neither has a coarse problem, cf. [90].

Overlapping Schwarz domain decomposition methods were first applied to mixed finite element discretizations of (1.1)–(1.2) in two dimensions by Mathew [55]. The saddle point problem (1.9)–(1.10) is first reduced to one that is symmetric and positive definite defined in terms of a subset the primal unknowns, namely, those fluxes in $\mathbf{V}_h(\Omega)$ that are divergence free. A description of Mathew's method is also given in [56]. A uniform bound on the asymptotic convergence rate for the Schwarz method for sufficiently large overlap was derived by Ewing and Wang [38] and Mathew [57]. The key to their analysis was the ability to represent the subspace of divergence free vector fields of mixed finite element spaces as the curl of continuous scalar polynomial stream functions. Using this representation, the convergence properties for the algorithm were reduced to those of a standard conforming discretization. The stream function representation was valid only in two dimensions; and hence, the analysis presented in [38] and [57] was limited to two dimensional problems. A bound that is $O(1/H^2)$ for the natural extension of the algorithm in three dimensions can be deduced from Lemma 2.4 of this chapter and the techniques in [55], but a uniform bound is at this time apparently an open question.

In this chapter, we apply Schwarz methods to the dual variable problem (1.19) involving only the mixed finite element approximations to the scalar variable p and its trace on the boundary of elements. We consider the standard overlapping additive Schwarz domain decomposition method due to Dryja and Widlund [33] and Nepomnyaschikh [61] applied to this reduced problem in both two and three dimensions. We also consider the Schwarz method applied to the dual variable substructuring method of Glowinski and Wheeler using the subspaces proposed by Smith [78].

In both the standard and substructuring Schwarz methods, results are derived in terms of a variable amount of overlap using the results of Dryja and Widlund [34]. Let H denote the diameter of the subdomains and δ the amount of the overlap between subdomains. For the standard overlapping Schwarz method, it is shown that the condition number of the dual-variable additive Schwarz method grows at worst as $O(1 + H/\delta)$ in both two and three dimensions, for all the standard mixed finite element families, and for elements of any fixed order. For the substructuring Schwarz method, it is shown that the condition number grows at worst like $O(1 + \log(H/\delta)^2)$. If the overlap is "generous", i.e. δ is some fixed fraction of H, the condition numbers in both cases are bounded by a constant that is independent of both the subdomain size and the mesh size. To the author's knowledge, these are the first asymptotically optimal domain decomposition results for mixed finite elements in three dimensions.

For both the standard and substructuring Schwarz methods, these are the same bounds derived in [34] for the standard Galerkin approximation using conforming linear finite elements. In fact, our central technique is to exploit the isomorphism between the mixed finite element space and a conforming piecewise linear finite element space introduced in Section 1.4. Using this isomorphism, we can inherit the existing conforming theory with only modest modification. As mentioned previously, this tool was first used in [25] to analyze the rate of convergence of Mandel's Balancing Domain Decomposition method proposed in [52].

The remainder of the chapter is divided into three sections. The standard additive Schwarz method is recalled in Section 2.2 along with its abstract convergence theory. In Section 2.3, we formulate and analyze the standard Schwarz method applied to the dual variable problem. The formulation and analysis of the substructuring Schwarz method of Smith applied to the dual problem are the subjects of Section 2.4.

2.2 Abstract Schwarz Theory

Following the presentation in [34], we recall the additive Schwarz method of Dryja and Widlund [33] and Nepomnyaschikh [61]. Let $d(\cdot,\cdot)$ be a positive definite, symmetric bilinear form on a finite dimensional Hilbert space $\mathcal{V}$, and let f be an element in the dual space of $\mathcal{V}$. The additive Schwarz method (with exact solves) for the variational problem of finding $p^* \in \mathcal{V}$ satisfying

$$d(p^*, q) = f(q) \quad \forall q \in \mathcal{V}, \tag{2.1}$$

is completely determined by a subspace decomposition of $\mathcal{V}$ in the following way. Let $\mathcal{V}_i$ be subspaces of $\mathcal{V}$ such that $\mathcal{V} = \mathcal{V}_0 + \ldots + \mathcal{V}_M$. For each subspace $\mathcal{V}_i$, define an energy orthogonal projection operator $\mathcal{P}_i : \mathcal{V} \to \mathcal{V}_i$ by

$$d(\mathcal{P}_i p, q) = d(p, q) \quad \forall q \in \mathcal{V}_i. \tag{2.2}$$

The additive Schwarz method for (2.1) involves the solution of

$$\mathcal{P} p^* = \hat{f}, \tag{2.3}$$

where

$$\mathcal{P} \equiv \sum_{i=0}^{M} \mathcal{P}_i, \qquad \hat{f} \equiv \sum_{i=0}^{M} f_i, \tag{2.4}$$

and $f_i \in \mathcal{V}_i$ is defined by

$$d(f_i, q) = f(q) \quad \forall q \in \mathcal{V}_i. \tag{2.5}$$

It is easy to see that the solutions to (2.1) and (2.3) are the same since

$$\mathcal{P}_i p^* = f_i.$$

The operator $\mathcal{P}$ is symmetric and positive definite with respect to the d-inner product, so conjugate gradients can be applied. Moreover, for well chosen subspaces $\mathcal{V}_i$, the condition number of $\mathcal{P}$ is much smaller than the one corresponding to (2.1) and so no further preconditioning is needed. Recall that the rate of convergence of the conjugate gradient algorithm can be bounded in terms of the condition number of $\mathcal{P}$, the ratio of the largest and smallest eigenvalues of $\mathcal{P}$; see, e.g., [45, 43].

Abstract bounds on the condition number of $\mathcal{P}$ have been derived in terms of two quantities that we now define. Let $C_0 > 0$ be a constant such that for every $q \in \mathcal{V}$ there exists a representation $q = \sum_{i=0}^{M} q_i$ with $q_i \in \mathcal{V}_i$ satisfying

$$\sum_{i=0}^{M} d(q_i, q_i) \leq C_0 d(q, q). \tag{2.6}$$

Let $\rho(\mathcal{E})$ denote the spectral radius of $\mathcal{E} = \{\epsilon_{ij}\}$, the matrix of strengthened Cauchy-Schwarz constants; that is, ϵ_{ij} is the smallest constant for which

$$|d(q_i, q_j)| \leq \epsilon_{ij} d(q_i, q_i)^{\frac{1}{2}} d(q_j, q_j)^{\frac{1}{2}} \quad \forall q_i \in \mathcal{V}_i,\ \forall q_j \in \mathcal{V}_j,\ i, j \geq 1. \tag{2.7}$$

The following theorem, due to Dryja and Widlund [35], bounds the condition number of the additive Schwarz methods in terms of the two quantities given above. It represents the continued refinement of results given by Nepomnyaschikh [61] and Lions [51].

Theorem 2.1 The eigenvalues of $\mathcal{P}$ lie in the interval $[C_0^{-1}, \rho(\mathcal{E}) + 1]$. Hence, the condition number of $\mathcal{P}$ is bounded by $C_0(\rho(\mathcal{E}) + 1)$.

In this thesis, we restrict our attention to the additive algorithm with exact solves. Analogous statements for the multiplicative Schwarz method and various hybrid methods can easily be deduced from the results contained herein since their rates of convergence can be bounded in terms of the same constants; cf. [11, 93, 32, 53]. Likewise, the use of inexact solves in (2.2), i.e. the replacement of the projections $\mathcal{P}_i$ by approximate projections, can be handled in the standard way; see, e.g. [35].

2.3 The Overlapping Schwarz Method for the Dual Problem

In this section, we first define the set of subspaces that we use in the Schwarz algorithm applied to (1.19). We then derive a bound on the condition number of the induced operator by estimating the constants in the abstract theory using the results for the existing conforming finite element theory.

Without loss of generality, we assume that Ω has unit diameter. We introduce a two-level decomposition of Ω, a coarse triangulation of Ω into nonoverlapping subdomains $\{\Omega_i\}_{i=1}^{M}$, and a further refinement into elements $\mathcal{T}_h$. We assume that the triangulation into subdomains is quasi-regular with characteristic length H and that $\mathcal{T}_h$ is quasi-regular with characteristic length h. For convenience, we consider subdomains that are triangular, rectangular, tetrahedral or rectangular solids as appropriate for the dimension of the problem and the mixed finite element family considered. We extend each subdomain Ω_i to a enclose a larger region Ω_i' that is also the union of elements of $\mathcal{T}_h$. The extended domains $\{\Omega_i'\}_{i=1}^{M}$ form an overlapping covering of Ω, and we characterize the extent of the overlap of by

$$\delta = \min_{i=1,\dots,M} \mathrm{dist}(\partial\Omega_i \setminus \partial\Omega, \partial\Omega_i' \setminus \partial\Omega).$$

Let $n_{\max}$ denote the maximum number of subdomains for $1 \leq i \leq M$ for which $\mathbf{meas}(\Omega_i' \cap \Omega_j') > 0$, $1 \leq j \leq M$. For δ small enough (for instance $\delta \leq CH$), $n_{\max}$ is uniformly bounded by a constant that only depends on the regularity of the triangulation of Ω into nonoverlapping subdomains.

For $i = 1, \ldots, M$, let $W_h(\Omega_i') \times \Lambda_h(\Omega_i') \subset W_h(\Omega) \times \Lambda_h^0(\Omega)$ denote the set of functions that vanish at all nodes on the boundary and outside of Ω_i'. Let $U_H(\Omega)$ denote the "coarse space" of continuous functions that are linear, bilinear, or trilinear as appropriate on each Ω_i. Let $\mathcal{V}_0$ denote the subspace of functions in $W_h(\Omega) \times \Lambda_h^0(\Omega)$ that have the same value as functions in $U_H(\Omega)$ at the primary vertices (i.e. the space of nodal interpolants).

In the following lemma we recall the crux of the proof due to Dryja and Widlund (Theorem 3 of [34]) that the Schwarz method applied to the conforming Galerkin discretization has a condition number that is $O(1 + H/\delta)$.

Lemma 2.1 There exists a constant C independent of h, H, and δ such that for every $\phi \in U_h(\Omega)$, there exists a decomposition $\phi = \sum_{i=0}^{M} \phi_i$ with $\phi_0 \in U_H(\Omega)$, $\phi_i \in U_h(\Omega) \cap H_0^1(\Omega_i')$, $1 \leq i \leq M$ satisfying

$$\sum_{i=0}^{M} |\phi_i|_{1,\Omega}^2 \leq C \left(1 + \frac{H}{\delta}\right) |\phi|_{1,\Omega}^2. \tag{2.8}$$

We now show that the application of the Schwarz method to the dual variable mixed finite element discretization converges at the same rate. Note that the isomorphism is used only as a tool in the analysis and does not enter the algorithm.

Theorem 2.2 The condition number $\kappa(\mathcal{P}_\Omega)$ of the additive Schwarz operator $\mathcal{P}_\Omega$ defined by (2.4) induced by the decomposition

$$W_h(\Omega) \times \Lambda_h^0(\Omega) = \mathcal{V}_0 + \sum_{i=1}^{M} (W_h(\Omega_i') \times \Lambda_h(\Omega_i'))$$

of the hybrid mixed finite finite element space satisfies

$$\kappa(\mathcal{P}_\Omega) \leq C(1 + n_{\max}) \left(1 + \frac{H}{\delta}\right).$$

The constant C is independent of h, δ, and H.

Proof The verification that the largest eigenvalue of $\mathcal{P}_\Omega$ is bounded by $(1 + n_{\max})$ is standard. Since $d_\Omega(\hat{p}_i, \hat{p}_j) \equiv 0$ for $\hat{p}_i \in W_h(\Omega_i') \times \Lambda_h(\Omega_i')$, $\hat{p}_j \in W_h(\Omega_j') \times \Lambda_h(\Omega_j')$ with $\Omega_i' \cap \Omega_j' = \emptyset$, there are at most $n_{\max}$ nonzero entries, each no larger than 1, in the i-th row of the matrix of strengthened Cauchy-Schwarz inequalities. By the Gershgorin Circle Theorem (see [43]), $\rho(\mathcal{E})$ is bounded by $n_{\max}$.

Let $\mathcal{N}_\Omega$ denote the nodal interpolation map at the primary vertices into $W_h(\Omega) \times \Lambda_h^0(\Omega)$. For $\widehat{p} \in W_h(\Omega) \times \Lambda_h^0(\Omega)$, let $(\mathcal{I}^\Omega \widehat{p})_i$ denote the decomposition of $\mathcal{I}^\Omega \widehat{p} \in U_h(\Omega)$ arising in Lemma 2.1, and set $\widehat{p}_i = \mathcal{N}_\Omega((\mathcal{I}^\Omega \widehat{p})_i)$. Since

$$\mathcal{V}_0 = \mathcal{N}_\Omega(U_H(\Omega)), \qquad W_h(\Omega_i') \times \Lambda_h(\Omega_i') = \mathcal{N}_\Omega(U_h(\Omega) \cap H_0^1(\Omega_i')),$$

it is easy to check that $\widehat{p}_0 \in \mathcal{V}_0$, $\widehat{p}_i \in W_h(\Omega_i') \times \Lambda_h(\Omega_i')$ and $\widehat{p} = \sum_{i=0}^M \widehat{p}_i$. Using the equivalence

$$a_\Omega(\phi, \phi) \simeq |\phi|_{1,\Omega}^2,$$

with Theorem 1.2 and Lemma 1.3, we see that for $i = 0, \ldots, M$,

$$d_\Omega(\widehat{p}_i, \widehat{p}_i) \leq C |\mathcal{I}^\Omega((\mathcal{I}^\Omega \widehat{p})_i)|_{1,\Omega}^2 \leq C |(\mathcal{I}^\Omega \widehat{p})_i|_{1,\Omega}^2.$$

Summing and applying Lemma 2.1 and Theorem 1.2, we conclude that

$$\sum_{i=0}^M d_\Omega(\widehat{p}_i, \widehat{p}_i) \leq C \sum_{i=0}^M |(\mathcal{I}^\Omega \widehat{p})_i|_{1,\Omega}^2 \leq C\left(1 + \frac{H}{\delta}\right) |\mathcal{I}^\Omega \widehat{p}|_{1,\Omega}^2 \leq C\left(1 + \frac{H}{\delta}\right) d_\Omega(\widehat{p}, \widehat{p}).$$

Hence, C_0 in (2.6) is bounded by $C\,(1 + H/\delta)$. An application of Theorem 2.1 completes the proof. □

In [68], Pavarino and Ramé give results for a parallel implementation of a very closely related method that verify the bound given above.

2.4 Smith's Substructuring Schwarz Method

2.4.1 The Dual-Variable Substructuring Problem

Following the presentation in [25], we recall the reduction of (1.19) to a problem involving only the interface unknowns first given by Glowinski and Wheeler [42].

Let Ω be partitioned into a quasi-uniform triangulation of nonoverlapping subdomains $\{\Omega_i\}_{i=1}^M$ with diameters that are $O(H)$. Denote by Γ the set of internal interfaces

$$\Gamma = \bigcup_{i=1,\ldots,M} \partial\Omega_i \setminus \partial\Omega.$$

Let $\Lambda_h(\Gamma)$ denote the subspace of functions in $\Lambda_h^0(\Omega)$ with support contained on Γ, and let $\Lambda_h(\partial\Omega_i) \subset \Lambda_h(\Gamma)$ denote those functions with support on $\partial\Omega_i \setminus \partial\Omega$.

Define the bilinear form $s : \Lambda_h(\Gamma) \times \Lambda_h(\Gamma) \to \mathbb{R}$ by

$$s(\lambda, \mu) = \sum_{i=1}^M s_i(\lambda, \mu),$$

where

$$s_i(\lambda, \mu) = d_{\Omega_i}(\widehat{p}_i(\lambda), \widehat{p}_i(\mu)),$$

and $\widehat{p}_i(\lambda) \in W_h(\Omega_i) \times \Lambda_h(\Omega_i)$ satisfies

$$d_{\Omega_i}(\widehat{p}_i(\lambda), \widehat{q}_i) = 0 \qquad \forall \widehat{q}_i \in W_h(\Omega_i) \times \Lambda_h^0(\Omega_i), \tag{2.9}$$

$$\widehat{p}_i(\lambda) = \lambda \qquad \text{on } \partial\Omega_i. \tag{2.10}$$

In practice, the bilinear form can be more easily evaluated as

$$s_i(\lambda, \mu) = -\int_{\partial\Omega_i} \mu \mathbf{u}_i(\lambda) \cdot \nu_{\Omega_i}\, ds,$$

where $[\mathbf{u}_i(\lambda), p_i(\lambda)] \in \mathbf{V}_h(\Omega_i) \times W_h(\Omega_i)$ solves

$$\int_{\Omega_i} \nabla\cdot \mathbf{u}_i(\lambda) q\, dx = 0 \qquad \forall q \in W_h(\Omega_i),$$

$$\int_{\Omega_i} A^{-1}\mathbf{u}_i(\lambda) \cdot \mathbf{v}\, dx - \int_{\Omega_i} p_i(\lambda) \nabla\cdot \mathbf{v}\, dx = -\int_{\partial\Omega_i} \lambda \mathbf{v} \cdot \nu_{\Omega_i}\, ds \qquad \forall \mathbf{v} \in \mathbf{V}_h(\Omega_i).$$

Let $\overline{p}_i \in W_h(\Omega_i) \times \Lambda_h(\Omega_i)$ satisfy

$$d_{\Omega_i}(\overline{p}_i, [q, \mu]) = \int_{\Omega_i} f q\, dx \qquad \forall [q, \mu] \in W_h(\Omega_i) \times \Lambda_h^0(\Omega_i), \tag{2.11}$$

$$\overline{p}_i = 0 \qquad \text{on } \partial\Omega_i. \tag{2.12}$$

Then, as shown in [42], (1.19) is equivalent to finding $\lambda \in \Lambda_h(\Gamma)$ such that

$$s(\lambda, \mu) = \sum_{i=1}^{M} d_{\Omega_i}(\overline{p}_i, \widehat{p}_i(\mu)) \quad \forall \mu \in \Lambda_h(\Gamma). \tag{2.13}$$

The solution to (1.19) is recovered subdomain by subdomain as

$$\overline{p}_i + \widehat{p}_i(\lambda) \in W_h(\Omega_i) \times \Lambda_h(\Omega_i).$$

The process described above is nothing more than the standard reduction of (1.19), or equivalently (1.11)–(1.13), to a "Schur complement" system involving only the unknowns $\Lambda_h(\Gamma)$ on the interfaces of subdomains.

2.4.2 A Substructuring Isomorphism

Analogous to $\widetilde{U}_h(\Omega)$ defined in Section 1.4, we construct a space of continuous, piecewise linear functions that are isomorphic to $\Lambda_h(\Gamma)$ with respect to the natural norms induced by the interface problem. Let $U_h(\Gamma)$ denote the space of restrictions of the piecewise linear functions in $U_h(\Omega)$ to $\bigcup_{i=1}^{M} \partial\Omega_i$. For each subdomain Ω_i define a map

$$\mathcal{I}^{\partial\Omega_i} : \Lambda_h(\partial\Omega_i) \to U_h(\Gamma)$$

by

$$\mathcal{I}^{\partial\Omega_i}\lambda(x) = \begin{cases} 0, \text{ if } x \in \partial\Omega; \\ \\ \lambda(x), \text{ if } x \text{ is a primary vertex on } \partial\Omega_i; \\ \\ \text{The average of all adjacent primary vertices on } \partial\Omega_i, \\ \quad \text{if } x \text{ is a secondary vertex on } \partial\Omega_i; \\ \\ \text{The continuous piecewise linear interpolant of the} \\ \quad \text{above vertex values, if } x \text{ is not a vertex of } \widehat{\mathcal{T}}. \end{cases} \tag{2.14}$$

We denote its range by $\widetilde{U}_h(\partial\Omega_i)$. Equivalently,

$$\mathcal{I}^{\partial\Omega_i}\lambda(x) = \begin{cases} (\mathcal{I}^{\Omega_i}\widehat{p}_\lambda)(x), & x \in \partial\Omega_i; \\ 0, & \text{otherwise}; \end{cases} \tag{2.15}$$

where $\mathcal{I}^{\Omega_i}$ is defined by (1.37) and $\widehat{p}_\lambda$ is any element in $W_h(\Omega) \times \Lambda_h^0(\Omega)$ such that $(\widehat{p}_\lambda)_{|\partial\Omega_i} = \lambda$. Since the value of $\mathcal{I}^{\Omega_i}\widehat{p}_\lambda$ on $\partial\Omega_i$ depends only on the value of λ at the primary vertices (the nodal degrees of freedom) on $\partial\Omega_i$, $\mathcal{I}^{\partial\Omega_i}\lambda$ is well defined. Again since $\mathcal{I}^{\partial\Omega_i}$ is defined for any function defined at the primary vertices, by an abuse of notation, we will also consider $\mathcal{I}^{\partial\Omega_i}$ as a map from $U_h(\partial\Omega_i)$ into $U_h(\partial\Omega_i)$.

Following [10] and [31], define the following scaled Sobolev norms:

$$\|u\|_{1,\Omega_i}^2 = |u|_{1,\Omega_i}^2 + \frac{1}{H^2}|u|_{0,\Omega_i}^2, \quad \|u\|_{1/2,\partial\Omega_i}^2 = |u|_{1/2,\partial\Omega_i}^2 + \frac{1}{H}|u|_{0,\partial\Omega_i}^2. \tag{2.16}$$

In [54], the following lemma is proven.

Lemma 2.2 Let B be the linear operator on $H^{\frac{1}{2}}(\partial\Omega_i)$ defined by

$$B(\phi) = \int_{\partial\Omega_i} \phi\, ds$$

if the $(n-1)$-dimensional measure of $\partial\Omega_i \cap \partial\Omega$ is zero, and

$$B(\phi) = \phi_{|\partial\Omega_i \cap \partial\Omega}$$

otherwise. Then for all $\phi \in \mathrm{Ker}(B)$

$$\|\phi\|_{1/2,\partial\Omega_i} \simeq |\phi|_{1/2,\partial\Omega_i}, \qquad \|\phi\|_{1,\Omega_i} \simeq |\phi|_{1,\Omega_i}$$

with equivalence constants independent of h and H.

The following theorem is a result of the standard Trace Theorem (see, e.g. [59]), Widlund's Discrete Extension Theorem [89] for continuous piecewise linear finite elements, and Lemma 2.2.

Theorem 2.3 The following equivalence holds uniformly for $\widehat{\phi} \in U_h(\partial\Omega_i)$ with equivalence constants independent of h and H:

$$|\widehat{\phi}|_{1/2,\partial\Omega_i} \simeq \inf_{\substack{\phi \in U_h(\Omega_i) \\ \phi_{|\partial\Omega_i} = \widehat{\phi}}} |\phi|_{1,\Omega_i}. \tag{2.17}$$

The next three lemmas provide the analogue to Theorem 1.2 and Lemma 1.3 for the substructuring forms and are recalled from [25] with proof for completeness.

Lemma 2.3 *(Lemma 6.2 of [25])* The following equivalence holds uniformly for $\widehat{\phi} \in \widetilde{U}_h(\partial\Omega_i)$ with constants independent of h and H:

$$|\widehat{\phi}|_{1/2,\partial\Omega_i} \simeq \inf_{\substack{\phi \in \widetilde{U}_h(\Omega_i) \\ \phi_{|\partial\Omega_i} = \widehat{\phi}}} |\phi|_{1,\Omega_i}. \tag{2.18}$$

Consequently, for $\lambda \in \Lambda_h(\partial\Omega_i)$

$$|\mathcal{I}^{\partial\Omega_i}\lambda|_{1/2,\partial\Omega_i} \simeq \inf_{\substack{\widehat{p}_\lambda \in W_h(\Omega_i) \times \Lambda_h(\Omega_i) \\ (\widehat{p}_\lambda)_{|\partial\Omega_i} = \lambda}} |\mathcal{I}^{\Omega_i}\widehat{p}_\lambda|_{1,\Omega_i}. \tag{2.19}$$

Proof As a consequence of Lemma 1.3 and the inclusion $\widetilde{U}_h(\Omega_i) \subset U_h(\Omega_i)$, we have

$$\inf_{\substack{\phi \in U_h(\Omega_i) \\ \phi_{|\partial\Omega_i} = \widehat{\phi}}} |\phi|_{1,\Omega_i} \simeq \inf_{\substack{\phi \in \widetilde{U}_h(\Omega_i) \\ \phi_{|\partial\Omega_i} = \widehat{\phi}}} |\phi|_{1,\Omega_i} \quad \forall \widehat{\phi} \in \widetilde{U}_h(\partial\Omega_i).$$

Equation (2.18) follows by an application of Theorem 2.3. Equation (2.19) follows from the fact that

$$\mathcal{I}^{\partial\Omega_i}(\Lambda_h(\partial\Omega_i)) = \widetilde{U}_h(\partial\Omega_i), \qquad \mathcal{I}^{\Omega_i}(W_h(\Omega_i) \times (\Lambda_h(\Omega_i) \cap \Lambda_h^0(\Omega))) = \widetilde{U}_h(\Omega_i).$$

$\square$

Lemma 2.4 *(Theorem 6.5 of [25])* For all $\lambda \in \Lambda_h(\partial\Omega_i)$,

$$s_i(\lambda, \lambda) \simeq |\mathcal{I}^{\partial\Omega_i}\lambda|^2_{1/2,\partial\Omega_i}. \tag{2.20}$$

Proof By a direct computation using (2.9)–(2.10),

$$s_i(\lambda_i, \lambda_i) = \inf_{\substack{\widehat{p} \in W_h(\Omega_i) \times \Lambda_h(\Omega_i) \\ \widehat{p}_{|\partial\Omega_i} = \lambda_i}} d_{\Omega_i}(\widehat{p}, \widehat{p}). \tag{2.21}$$

The lemma follows by taking the infimum of (1.38) over $\widehat{p}_{|\partial\Omega_i} = \lambda_i$ and using Lemma 2.3 since

$$\begin{aligned} s_i(\lambda_i, \lambda_i) &= \inf_{\substack{\widehat{p} \in W_h(\Omega_i) \times \Lambda_h(\Omega_i) \\ \widehat{p}_{|\partial\Omega_i} = \lambda_i}} d_{\Omega_i}(\widehat{p}, \widehat{p}) \\ &\simeq \inf_{\substack{\widehat{p} \in W_h(\Omega_i) \times \Lambda_h(\Omega_i) \\ \widehat{p}_{|\partial\Omega_i} = \lambda_i}} |\mathcal{I}^{\Omega_i}\widehat{p}|^2_{1,\Omega_i} \\ &\simeq |\mathcal{I}^{\partial\Omega_i}\lambda_i|^2_{1/2,\partial\Omega}. \end{aligned} \tag{2.22}$$

□

Lemma 2.5 *(Lemma 6.3 of [25])* There exists a constant C independent of h and H such that

$$|\mathcal{I}^{\partial\Omega_i}\widehat{\phi}|_{1/2,\partial\Omega} \leq C|\widehat{\phi}|_{1/2,\partial\Omega} \quad \forall\widehat{\phi} \in U_h(\partial\Omega_i). \tag{2.23}$$

Proof Let $\phi \in U_h(\Omega_i)$ such that $\phi_{|\partial\Omega_i} = \widehat{\phi}$. Since $(\mathcal{I}^{\Omega_i}\phi)_{|\partial\Omega_i} = \mathcal{I}^{\partial\Omega_i}\widehat{\phi}$, we have by Lemma 2.3 and Lemma 1.3

$$|\mathcal{I}^{\partial\Omega_i}\widehat{\phi}|_{1/2,\partial\Omega} \leq C|\mathcal{I}^{\Omega_i}\phi|_{1,\Omega} \leq C|\phi|_{1,\Omega}.$$

The lemma follows by taking the infimum over all $\phi \in U_h(\Omega_i)$ such that $\phi_{|\partial\Omega_i} = \widehat{\phi}$, and using Theorem 2.3. □

Note that $\mathcal{I}^{\partial\Omega_i}\lambda$ does not in general equal $\mathcal{I}^{\partial\Omega_j}\lambda$ on $\partial\Omega_i \cap \partial\Omega_j$ because they can take different values at the secondary vertices. In the next section, we need a pseudo-interpolant of $\lambda \in \Lambda_h(\Gamma)$ in $U_h(\Gamma)$ that is defined on all of Γ. Hence, we define a pseudo-interpolant $\mathcal{I}^\Gamma : \Lambda_h(\Gamma) \to U_h(\Gamma)$ by (2.14) with $\partial\Omega_i$ replaced by the set of interfaces Γ. Equivalently,

$$\mathcal{I}^\Gamma\lambda = (\mathcal{I}^{\Omega\setminus\Gamma}\widehat{p}_\lambda)_{|\Gamma},$$

where $\hat{p}_\lambda$ is any element in $W_h(\Omega) \times \Lambda_h^0(\Omega)$ such that $(\hat{p}_\lambda)_{|\Gamma} = \lambda$, and $\mathcal{I}^{\Omega\setminus\Gamma} : W_h(\Omega) \times \Lambda_h^0(\Omega) \to U_h(\Omega)$ is defined by (1.37) noting that the boundary of $\Omega \setminus \Gamma$ is $\partial\Omega \cup \Gamma$. The next lemma relates the pseudo-interpolant defined on all of Γ to the sum of pseudo-interpolants defined subdomain by subdomain.

Lemma 2.6 There exists a constant $C > 0$ independent of h, H and δ such that

$$\sum_{i=1}^{M} |\mathcal{I}^{\Gamma}\lambda|_{1/2,\partial\Omega_i}^2 \le C \sum_{i=1}^{M} |\mathcal{I}^{\partial\Omega_i}\lambda|_{1/2,\partial\Omega_i}^2 \quad \forall\lambda \in \Lambda_h(\Gamma). \tag{2.24}$$

Proof By using the techniques in the proof of Theorem 1.2, it is easy to show that there exists a constant $C > 0$ depending only on the regularity of the mesh and the choice of mixed finite element space such that

$$\sum_{i=1}^{M} |\mathcal{I}^{\Omega\setminus\Gamma}\hat{p}|_{1,\Omega_i}^2 \le C \sum_{i=1}^{M} |\mathcal{I}^{\Omega_i}\hat{p}|_{1,\Omega_i}^2 \quad \forall\hat{p} \in W_h(\Omega) \times \Lambda_h(\Omega). \tag{2.25}$$

By Lemma 2.3, there exists a constant C independent of h, H such that for each $\lambda \in \Lambda_h(\Gamma)$ there exists an extension $E(\lambda) \in W_h(\Omega) \times \Lambda_h^0(\Omega)$ that agrees with λ on Γ and satisfies uniformly

$$|\mathcal{I}^{\Omega_i} E(\lambda)|_{1,\Omega_i}^2 \le C |\mathcal{I}^{\partial\Omega_i}\lambda|_{1/2,\partial\Omega_i}^2 \quad i = 1, \ldots, M.$$

Since $\mathcal{I}^{\Gamma}\lambda \in U_h(\partial\Omega_i)$ and $\mathcal{I}^{\Omega\setminus\Gamma}E(\lambda) \in U_h(\Omega)$ agree on $\partial\Omega_i$, by Theorem 2.3 we see that

$$|\mathcal{I}^{\Gamma}\lambda|_{1/2,\partial\Omega_i} \le C |\mathcal{I}^{\Omega\setminus\Gamma}E(\lambda)|_{1,\Omega_i} \quad \forall\lambda \in \Lambda_h(\Gamma).$$

Combining these results, we conclude that

$$\sum_{i=1}^{M} |\mathcal{I}^{\Gamma}\lambda|_{1/2,\partial\Omega_i}^2 \;\le\; C\sum_{i=1}^{M} |\mathcal{I}^{\Omega\setminus\Gamma}E(\lambda)|_{1,\Omega_i}^2 \le C\sum_{i=1}^{M} |\mathcal{I}^{\Omega_i}E(\lambda)|_{1,\Omega_i}^2 \le C\sum_{i=1}^{M} |\mathcal{I}^{\partial\Omega_i}\lambda|_{1/2,\partial\Omega_i}^2,$$

which proves the lemma. □

2.4.3 Smith's Vertex Space Substructuring Method

In this section, we analyze the application of the Schwarz method applied to (2.13) using a subspace decomposition of the interface unknowns $\Lambda_h(\Gamma)$ analogous to the one suggested by Smith [78] for conforming elements. Following the presentation given in [34], we construct the decomposition of $\Lambda_h(\Gamma)$ slightly differently in two and three dimensions. In both cases, we first partition Γ into overlapping subsets based on its decomposition as the boundary of subdomains. In two dimensions, for each vertex V_j of Γ, let $\Gamma_\delta^{V_j}$ denote the set of points on Γ that are less than a distance δ from V_j. For each edge E_i of Γ, let $\Gamma_\delta^{E_i}$ denote the interior of the edge E_i. In three dimensions, for each vertex V_j, each edge E_i, and each face F_k of Γ, define $\Gamma_\delta^{V_j}$ as above, let $\Gamma_\delta^{F_k}$ denote the interior of the face F_k, and let $\Gamma_\delta^{E_i}$ denote the set of all points in strips of width δ on all faces which share the common edge E_i. As in the standard overlapping case, let $n_{\max}$ bound the maximum number of subsets Γ_δ^G, $G \in \{E_i, V_j, F_k\}$ intersecting any given subset $\Gamma_\delta^{G^*}$, $G^* \in \{E_i, V_j, F_k\}$. Again, for δ sufficiently small, $n_{\max}$ is uniformly bounded.

Understanding the set of faces to be empty in two dimensions, the decomposition of Γ into subsets induces a decomposition of $\Lambda_h(\Gamma)$ by considering

$$\Lambda_h(\Gamma) = \sum_{G \in \{H, E_i, V_j, F_k\}} \Lambda_h(\Gamma_\delta^G),$$

where for $G \in \{E_i, V_j, F_k\}$, $\Lambda_h(\Gamma_\delta^G) \subset \Lambda_h(\Gamma)$ are those functions that vanish at all nodal points on Γ outside the set Γ_δ^G, and $\Lambda_h(\Gamma_\delta^H) \subset \Lambda_h(\Gamma)$ are the nodal interpolants of the restriction to Γ of continuous functions that are linear on each subdomain Ω_i and vanish on $\partial\Omega$. Examples of $\Lambda_h(\Gamma_\delta^{E_i})$, $\Lambda_h(\Gamma_\delta^{V_j})$, and $\Lambda_h(\Gamma_\delta^{F_k})$ for the lowest order Raviart-Thomas space on rectangles and cubes are depicted in Figure 2.1 and Figure 2.2.

The following lemma is the crux of the analysis of Smith's method in [78] for conforming elements. We recall the small overlap refinement from the proof of Theorem 4 in [34].

Lemma 2.7 For every $\phi \in U_h(\Gamma)$, there exists a decomposition

$$\phi = \sum_{G \in \{H, E_i, V_j, F_k\}} \phi_G$$

with $\phi_H \in U_H(\Gamma)$, $\phi_G \in U_h(\Gamma_\delta^G) = U_h(\Gamma) \cap H_0^1(\Gamma_\delta^G)$ for $G \in \{E_i, V_j, F_k\}$ such that

$$\sum_{G \in \{H, E_i, V_j, F_k\}} \sum_{i=1}^{M} |\phi_G|^2_{1/2,\partial\Omega_i} \le C \left(1 + \log(H/\delta)\right)^2 \sum_{i=1}^{M} |\phi|^2_{1/2,\partial\Omega_i}. \qquad (2.26)$$

The constant C is independent of the choice of ϕ, and the mesh parameters h, H, and δ.

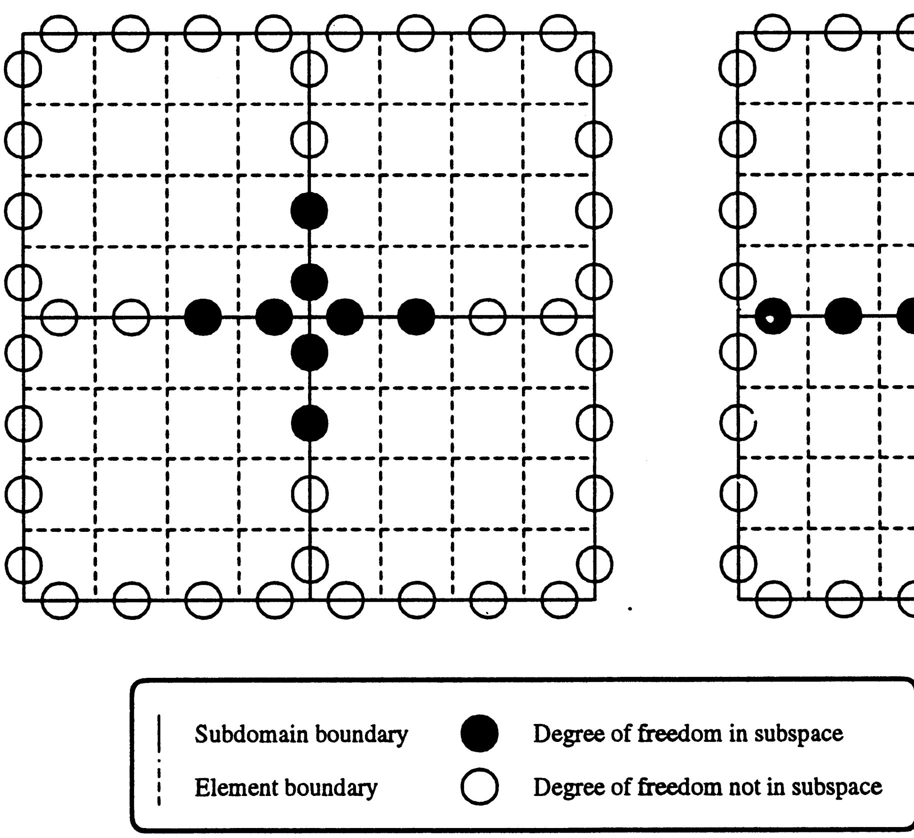

Figure 2.1 Support of a vertex and an edge subspace for rectangular subdomains ($H = 4h$, $\delta = 2h$).

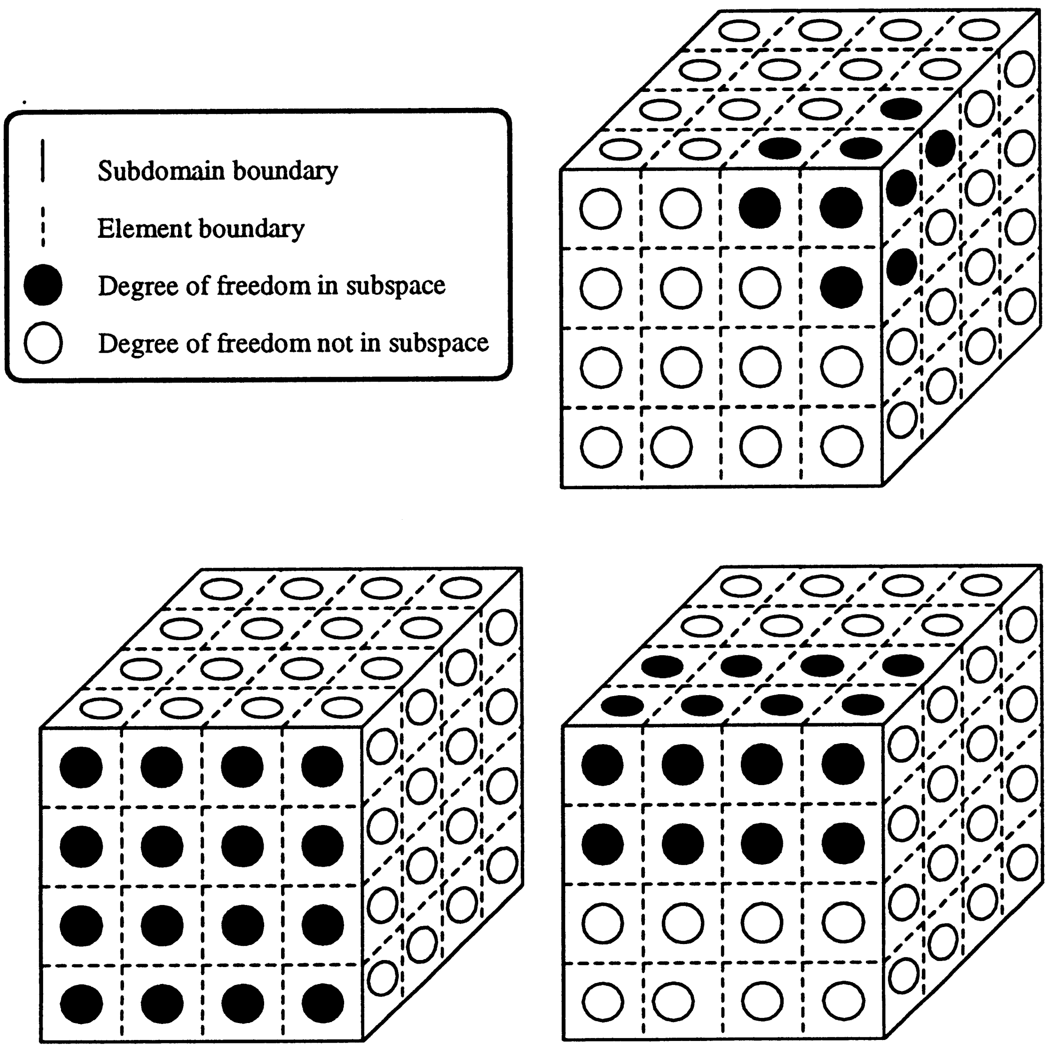

Figure 2.2 Intersection of the support of a vertex, a face, and an edge subspace with a typical cubic subdomain ($H = 4h$, $\delta = 2h$).

Let $\mathcal{P}_\Gamma : \Lambda_h(\Gamma) \rightarrow \Lambda_h(\Gamma)$ denote the additive Schwarz operator defined by (2.4) with the bilinear form $d(\cdot,\cdot)$ replaced by the interface form $s(\cdot,\cdot)$ and the decomposition of $\mathcal{V}$ replaced by the decomposition of $\Lambda_h(\Gamma)$ described above. We now prove that the condition number of $\mathcal{P}_\Gamma$ has the same bound given in [34] for the similar operator for the conforming finite element space.

Theorem 2.4 The condition number of the additive Schwarz operator $\mathcal{P}_\Gamma$ induced by Smith's decomposition for the hybrid mixed finite element discretization satisfies

$$\kappa(\mathcal{P}_\Gamma) \leq C(1+n_{\max})\left(1+\log\left(H/\delta\right)\right)^2. \tag{2.27}$$

The constant C is independent of the mesh parameters h, H, and δ.

Proof As in the proof of Theorem 2.2, the number of nonzero entries per row in the matrix of strengthen Cauchy-Schwarz inequalities can be bounded by $n_{\max}$; hence, the largest eigenvalue of $\mathcal{P}_\Gamma$ is bounded by $1+n_{\max}$.

To bound the smallest eigenvalue, we also proceed as in the proof of Theorem 2.2. Let $\mathcal{N}_\Gamma : U_h(\Gamma) \rightarrow \Lambda_h(\Gamma)$ denote the nodal interpolation into $\Lambda_h(\Gamma)$ for functions defined at primary vertices. For $\lambda \in \Lambda_h(\Gamma)$, set $\lambda_G = \mathcal{N}_\Gamma((\mathcal{I}^\Gamma\lambda)_G)$, $G \in \{H, E_i, V_j, F_k\}$, where $(\mathcal{I}^\Gamma\lambda)_G$ is the decomposition of $\mathcal{I}^\Gamma\lambda \in U_h(\Gamma)$ arising in Lemma 2.7. Since $\mathcal{I}^\Gamma\lambda$ and λ agree at the nodal degrees of freedom of $\Lambda_h(\Gamma)$, and

$$\Lambda_h(\Gamma_\delta^H) = \mathcal{N}_\Gamma(U_H(\Gamma)), \qquad \Lambda_h(\Gamma_\delta^G) = \mathcal{N}_\Gamma(U_h(\Gamma_\delta^G)) \quad \forall G \in \{E_i, V_j, F_k\},$$

it is easy to check that

$$\lambda = \sum_{G\in\{H,E_i,V_j,F_k\}} \lambda_G.$$

Working one subdomain at a time and using Lemma 2.4 and Lemma 2.5, we see that for $G = H$ and for $G \in \{E_i, V_j, F_k\}$ such that $\Gamma_\delta^G \cap \partial\Omega_i \neq \emptyset$ we have

$$\begin{aligned} s_i(\lambda_G,\lambda_G) &\leq C|\mathcal{I}^{\partial\Omega_i}\lambda_G|^2_{1/2,\partial\Omega_i} = C|\mathcal{I}^{\partial\Omega_i}((\mathcal{I}^\Gamma\lambda)_G)|^2_{1/2,\partial\Omega_i} \\ &\leq C|(\mathcal{I}^\Gamma\lambda)_G|^2_{1/2,\partial\Omega_i}. \end{aligned} \tag{2.28}$$

By summing (2.28) over subdomains and subspaces, noting that $s_i(\lambda_G,\lambda_G) = 0$ if $\Gamma_\delta^G \cap \partial\Omega_i = \emptyset$, and applying Lemma 2.7, Lemma 2.6, and Lemma 2.4, we see that

$$\sum_{G\in\{H,E_i,V_j,F_k\}} s(\lambda_G,\lambda_G) \quad = \quad \sum_{G\in\{H,E_i,V_j,F_k\}} \sum_{i=1}^{M} s_i(\lambda_G,\lambda_G)$$

$$
\begin{aligned}
&\leq C \sum_{G\in\{H,E_i,V_j,F_k\}} \sum_{i=1}^{M} |(\mathcal{I}^{\Gamma}\lambda)_G|^2_{1/2,\partial\Omega_i} \\
&\leq C\,(1+\log(H/\delta))^2 \sum_{i=1}^{M} |\mathcal{I}^{\Gamma}\lambda|^2_{1/2,\partial\Omega_i} \\
&\leq C\,(1+\log(H/\delta))^2 \sum_{i=1}^{M} |\mathcal{I}^{\partial\Omega_i}\lambda|^2_{1/2,\partial\Omega_i} \\
&\leq C\,(1+\log(H/\delta))^2 \sum_{i=1}^{M} s_i(\lambda,\lambda) \\
&\leq C\,(1+\log(H/\delta))^2\, s(\lambda,\lambda).
\end{aligned}
$$

An application of Theorem 2.1 completes the proof. $\square$

Chapter 3

Preconditioners for Hybrid Mixed Finite Elements using Conforming Discretizations

3.1 Wohlmuth-Hoppe Type Preconditioners

In this chapter, we construct preconditioners for the dual variable problem (1.19) from preconditioners for conforming discretizations of (1.1)–(1.2). We first follow ideas of Wohlmuth and Hoppe [92] and construct preconditioners for the dual variable problem from preconditioners for (1.1)–(1.2) discretized in $U_h(\Omega)$, the continuous piecewise linear functions on $\widehat{\mathcal{T}}$ defined in Section 1.4. In the next section, two V-cycle multigrid preconditioners for the dual variable problem are formulated. The techniques used in this chapter are also similar to ones employed by Oswald [66]. The results of some numerical experiments that compare the preconditioners constructed in this chapter are reported in Section 3.3.

Motivated by an isomorphism between nonconforming finite elements and a conforming space of functions presented by the author in [23] and analogous to the one given in Section 1.4, Wohlmuth and Hoppe [92] construct a preconditioner for a piecewise linear nonconforming discretization of (1.1)–(1.2) from a hierarchical basis preconditioner for a continuous piecewise linear discretization on a finer grid. Their numerical experiments suggest that the condition number of the resulting preconditioned operator is serendipitously bounded by a constant, a result not explained by current theory (cf., [64]). In this section, we apply their ideas to hybrid mixed finite elements.

3.1.1 A Preconditioner Using a Discretization on $\widehat{\mathcal{T}}$

We introduce some operator notation to simplify our presentation. Define an L^2-innerproduct on $W_h(\Omega) \times \Lambda_h^0(\Omega)$ by

$$(([p,\lambda],[q,\mu])) = \sum_{\tau\in\mathcal{T}_h} \int_\tau pq\,dx + \sum_{e\in\partial\mathcal{T}_h} \int_e \lambda\mu\,ds.$$

Let $D_h : W_h(\Omega) \times \Lambda_h^0(\Omega) \to W_h(\Omega) \times \Lambda_h^0(\Omega)$ be the operator defined by

$$((D_h[p,\lambda],[q,\mu])) = d([p,\lambda],[q,\mu]) \quad \forall [q,\mu] \in W_h(\Omega) \times \Lambda_h^0(\Omega).$$

In this notation, the dual variable problem (1.19) becomes

$$D_h[p,\lambda] = [P_{W_h(\Omega)} f, 0], \tag{3.1}$$

where $P_{W_h(\Omega)}$ is $L^2(\Omega)$-projection onto $W_h(\Omega)$.

Let $(\cdot,\cdot)_h$ be an innerproduct on $U_h(\Omega)$ uniformly equivalent with respect to h to the natural $L^2(\Omega)$-innerproduct, and define the operator $A_h : U_h(\Omega) \to U_h(\Omega)$ by

$$(A_h\phi,\psi)_h = \int_\Omega A\nabla\phi\cdot\nabla\psi\,dx \quad \forall \phi,\psi \in U_h(\Omega).$$

A natural choice for the $(\cdot,\cdot)_h$-innerproduct is

$$(\phi,\psi)_h = \sum_{\tau\in\widehat{\mathcal{T}}} |\tau| \sum_{\substack{\text{vertices}:\\ v_i\,\in\,\tau}} \phi(v_i)\psi(v_i).$$

Let $\mathcal{N} : U_h(\Omega) \to W_h(\Omega)\times\Lambda_h^0(\Omega)$ denote the nodal interpolation map into $W_h(\Omega)\times\Lambda_h^0(\Omega)$ at the primary vertices, and let $\mathcal{N}^* : W_h(\Omega)\times\Lambda_h^0(\Omega) \to U_h(\Omega)$ denote its L^2-adjoint satisfying

$$(\mathcal{N}^*[q,\mu],\phi)_h = (([q,\mu],\mathcal{N}\phi)).$$

No computation is involved in applying $\mathcal{N}$ since it is an identification of nodal degrees of freedom. Likewise, applying $\mathcal{N}^*$ is trivial since the "mass matrix" corresponding to the $(\cdot,\cdot)_h$-innerproduct is diagonal. The "mass matrix" corresponding to the $((\cdot,\cdot))$-innerproduct is also diagonal (or at least edge-wise and element-wise block diagonal depending on the choice of nodal degrees of freedom for the mixed finite element spaces).

The operator $\mathcal{N}A_h^{-1}\mathcal{N}^*$ is clearly symmetric and positive definite on $W_h(\Omega)\times\Lambda_h^0(\Omega)$ in the $((\cdot,\cdot))$-innerproduct and can be used as an effective preconditioner for D_h. We show that the condition number of the operator resulting from preconditioning D_h by $\mathcal{N}A_h^{-1}\mathcal{N}^*$ is bounded by a constant independent of h, a consequence of the following lemma due to Nepomnyaschikh [62].

Lemma 3.1 *(Lemma 2.2 of [62])* Let $\widehat{H}$ and H be Hilbert spaces with the scalar products $((\cdot,\cdot))$ and $(\cdot,\cdot)$, respectively. Let $\widehat{A} : \widehat{H} \to \widehat{H}$ and

$A : H \to H$ be self-adjoint, positive definite and continuous operators. Let $R : H \to \widehat{H}$ be a linear operator such that

$$((\hat{A}R\phi, R\phi)) \leq C_R(A\phi, \phi) \quad \forall \phi \in H.$$

If there exists an operator $T : \widehat{H} \to H$ such that

$$RT\hat{\phi} = \hat{\phi} \quad \forall \hat{\phi} \in \widehat{H},$$

and

$$C_T(AT\hat{\phi}, T\hat{\phi}) \leq ((\hat{A}\hat{\phi}, \hat{\phi})) \quad \forall \hat{\phi} \in \widehat{H},$$

then

$$C_T((\hat{A}^{-1}\hat{\phi}, \hat{\phi})) \leq ((RA^{-1}R^*\hat{\phi}, \hat{\phi})) \leq C_R((\hat{A}^{-1}\hat{\phi}, \hat{\phi})) \quad \forall \hat{\phi} \in \widehat{H},$$

where $R^* : \widehat{H} \to H$ is the adjoint of R satisfying

$$(R^*\hat{\phi}, \phi) = ((\hat{\phi}, R\phi)) \quad \forall \phi \in H, \hat{\phi} \in \widehat{H}.$$

Theorem 3.1 If $\mathcal{N}A_h^{-1}\mathcal{N}^*$ is used as a preconditioner for D_h, then the condition number of the preconditioned operator satisfies

$$\kappa(\mathcal{N}A_h^{-1}\mathcal{N}^*D_h) \leq C$$

where C is a constant independent of h and α; equivalently,

$$((\mathcal{N}A_h^{-1}\mathcal{N}^*[q, \mu], [q, \mu])) \simeq ((D_h^{-1}[q, \mu], [q, \mu])) \quad \forall [q, \mu] \in W_h(\Omega) \times \Lambda_h^0(\Omega).$$

Proof Since we have previously understood $\mathcal{I}^\Omega$ both as a map from $W_h(\Omega) \times \Lambda_h^0(\Omega)$ into $U_h(\Omega)$ and a map from $U_h(\Omega)$ into itself, to avoid possible confusion we denote the two maps by

$$\mathcal{I}_1^\Omega : W_h(\Omega) \times \Lambda_h^0(\Omega) \to U_h(\Omega),$$
$$\mathcal{I}_2^\Omega : U_h(\Omega) \to U_h(\Omega),$$

and note that

$$\mathcal{I}_2^\Omega \equiv \mathcal{I}_1^\Omega \mathcal{N} : U_h(\Omega) \to U_h(\Omega).$$

By Theorem 1.2 and Lemma 1.3, we have for $\phi \in U_h(\Omega)$ that

$$\begin{aligned} ((D_h\mathcal{N}\phi, \mathcal{N}\phi)) &= d_\Omega(\mathcal{N}\phi, \mathcal{N}\phi) \\ &\leq Ca_\Omega(\mathcal{I}_1^\Omega\mathcal{N}\phi, \mathcal{I}_1^\Omega\mathcal{N}\phi) = Ca_\Omega(\mathcal{I}_2^\Omega\phi, \mathcal{I}_2^\Omega\phi) \\ &\leq Ca_\Omega(\phi, \phi) = C(A_h\phi, \phi)_h. \end{aligned}$$

Also by Theorem 1.2, we see that for $[q,\mu] \in W_h(\Omega) \times \Lambda_h^0(\Omega)$

$$\begin{aligned}(A_h \mathcal{I}_1^{\Omega}[q,\mu], \mathcal{I}_1^{\Omega}[q,\mu])_h &= a(\mathcal{I}_1^{\Omega}[q,\mu], \mathcal{I}_1^{\Omega}[q,\mu]) \\ &\leq Cd([q,\mu],[q,\mu]) = C((D_h[q,\mu],[q,\mu])).\end{aligned}$$

An application of Lemma 3.1 with $\widehat{H} = W_h(\Omega) \times \Lambda_h^0(\Omega)$, $H = U_h(\Omega)$, $\hat{A} = D_h$, $A = A_h$, $R = \mathcal{N}$, and $T = \mathcal{I}_1^{\Omega}$ completes the proof. □

3.1.2 Using Preconditioners for A_h

The use of $\mathcal{N} A_h^{-1} \mathcal{N}^*$ as a preconditioner may be prohibitively expensive since it requires the solution of a conforming discretization of the elliptic problem at every iteration. We can replace A_h^{-1} with any operator B_h approximating A_h^{-1}. If B_h satisfies

$$c_1(A_h^{-1}\phi,\phi)_h \leq (B_h\phi,\phi)_h \leq c_2(A_h^{-1}\phi,\phi)_h \quad \forall \phi \in U_h(\Omega), \tag{3.2}$$

then

$$((D_h^{-1}[q,\mu],[q,\mu])) \leq C_1(A_h^{-1}\mathcal{N}^*[q,\mu],\mathcal{N}^*[q,\mu])_h \leq \frac{C_1}{c_1}(B_h\mathcal{N}^*[q,\mu],\mathcal{N}^*[q,\mu])_h,$$

$$((D_h^{-1}[q,\mu],[q,\mu])) \geq C_2(A_h^{-1}\mathcal{N}^*[q,\mu],\mathcal{N}^*[q,\mu])_h \geq \frac{C_2}{c_2}(B_h\mathcal{N}^*[q,\mu],\mathcal{N}^*[q,\mu])_h,$$

and the following corollary to Theorem 3.1 is immediate.

Corollary 3.1 If B_h is a symmetric positive definite operator on $U_h(\Omega)$ satisfying (3.2), then

$$\kappa(\mathcal{N} B_h \mathcal{N}^* D_h) \leq \frac{c_2}{c_1}\kappa(\mathcal{N} A_h^{-1}\mathcal{N}^* D) \leq C\frac{c_2}{c_1}.$$

Many preconditioners for the hybrid mixed formulation can be constructed in this manner by using existing preconditioners for the conforming case. One may then take advantage of the growing supply of efficient, well-crafted implementations of preconditioners for conforming finite elements. A likely candidate for B_h is a multigrid or multilevel preconditioner for A_h. We recall the construction of one possible V-cycle preconditioner for A_h below.

Assume that the triangulation $\mathcal{T}_h$ is the result of $J-1$ successive refinements $\mathcal{T}_2, \mathcal{T}_3, \ldots, \mathcal{T}_J = \mathcal{T}_h$ of a coarse triangulation $\mathcal{T}_1$. We set $\mathcal{T}_{J+1} = \widehat{\mathcal{T}}$, the refined triangulation of $\mathcal{T}_h$ constructed in Section 1.4. For $i = 1, \ldots, J+1$, denote by U_i the space of continuous functions that are piecewise linear (or bilinear, or trilinear, as appropriate) with respect to $\mathcal{T}_i$ and vanish on $\partial\Omega$. Note that

$$U_1 \subset U_2 \subset \ldots \subset U_J \subset U_{J+1} = U_h(\Omega).$$

For $i = 1, \ldots, J+1$, define operators $A_i : U_i \to U_i$ by

$$(A_i\phi, \psi) = a(\phi, \psi) \quad \forall \phi, \psi \in U_i,$$

and $L^2(\Omega)$-projections $Q_i : U_{J+1} \to U_i$ by

$$(Q_i\phi, \psi) = (\phi, \psi) \quad \forall \phi, \psi \in U_i.$$

For $i = 2, \ldots, J+1$, let $R_i : U_i \to U_i$ denote one sweep of the point Gauss-Seidel iterative method for the operator A_i, and let R_i^T denote its adjoint (a sweep using the reverse ordering).

Algorithm 3.1 Define the multiplicative V-cycle operator

$$B_i^m : U_i \to U_i$$

for $r \in U_i$ by

$$B_1^m r = A_1^{-1} r \tag{3.3}$$

if $i = 1$; otherwise, let

$$\begin{aligned} \phi &= R_i^T r, && (3.4)\\ \hat{\phi} &= \phi + B_{i-1}^m Q_{i-1}(r - A_i\phi), && (3.5)\\ B_i^m r &= \hat{\phi} + R_i(r - A_i\hat{\phi}). && (3.6) \end{aligned}$$

An optimal rate of convergence for the V-cycle has been proven under increasingly less restrictive conditions (see, e.g., [8, 6, 91]). We recall one such result due to Bramble and Pasciak [9].

Theorem 3.2 *(Theorem 4.2 of [9])* If the entries in the coefficient of A are in the Sobolev space $W_q^\gamma(\Omega)$ for $\gamma \in (0, 1/2)$ and $q > n/\gamma$, and if the solution p to Laplace's equation ((1.1)–(1.2) with $A = I$) satisfies for some $\delta \in (0, 1]$

$$\|p\|_{1+\delta,\Omega} \leq C\|f\|_{-1+\delta,\Omega},$$

then there exists a positive constant c_1 independent of h and the number of levels J such that

$$c_1(A_h^{-1}\phi,\phi) \le (B_{J+1}^m\phi,\phi) \le (A_h^{-1}\phi,\phi).$$

A uniform rate of convergence for the resulting preconditioner for D_h follows directly from Corollary 3.1, Theorem 3.2 and the assumption that the $(\cdot,\cdot)_h$-innerproduct is equivalent to the $L^2(\Omega)$-innerproduct.

Corollary 3.2 There exists a constant $C > 0$ independent of h and J such that

$$\kappa(\mathcal{N}B_{J+1}^m\mathcal{N}^*D_h) \le C.$$

3.2 Two V-Cycle Preconditioners for D_h

In this section, we construct of two V-cycle preconditioners for the dual variable problem. The first is obtained by adding a pre-smoothing and post-smoothing step to $\mathcal{N}B_{J+1}^m\mathcal{N}^*$. The second method is a natural extension to hybrid mixed finite elements of a method due to Bramble, Pasciak and Xu [12] for the dual variable problem arising from the mixed finite element discretization without hybridization. To avoid possible confusion, we note that this is not the popular BPX-multilevel method [13] also due to the same authors.

3.2.1 A V-cycle Preconditioner using $\mathcal{N}B_{J+1}^m\mathcal{N}^*$

With respect to the mixed finite element space $W_h(\Omega)\times\Lambda_h^0(\Omega)$, we can consider $U_h(\Omega)$ as a "coarse space" even though $U_h(\Omega)$ is defined on $\hat{\mathcal{T}}$, a finer grid than $\mathcal{T}_h$. The operators $\mathcal{N}A_h^{-1}\mathcal{N}^*$ and $\mathcal{N}B_{J+1}^m\mathcal{N}^*$ can likewise be thought of as coarse grid correction operators. Let R_m denote one sweep of point Gauss-Seidel on the hybrid mixed space $W_h(\Omega)\times\Lambda_h^0(\Omega)$ dual problem D_h, and let R_m^T denote its adjoint. By adding a pre-smoothing and a post-smoothing step to $\mathcal{N}A_h^{-1}\mathcal{N}^*$ and $\mathcal{N}B_{J+1}^m\mathcal{N}^*$, we arrive at a two-level and a multiple level V-cycle preconditioner for D_h.

Algorithm 3.2 Define a two-level preconditioner

$$\hat{A}_2^m : W_h(\Omega)\times\Lambda_h^0(\Omega) \to W_h(\Omega)\times\Lambda_h^0(\Omega)$$

for $[r,\rho] \in W_h(\Omega) \times \Lambda_h^0(\Omega)$ by

$$\begin{aligned} [q,\mu] &= R_m^T[r,\rho], & (3.7)\\ [\hat{q},\hat{\mu}] &= [q,\mu] + \mathcal{N}A_h^{-1}\mathcal{N}^*([r,\rho] - D_h[q,\mu]), & (3.8)\\ \hat{A}_2^m[r,\rho] &= [\hat{q},\hat{\mu}] + R_m([r,\rho] - D_h[\hat{q},\hat{\mu}]). & (3.9)\end{aligned}$$

Define a multiple level V-cycle preconditioner

$$\hat{B}_{J+1}^m : W_h(\Omega) \times \Lambda_h^0(\Omega) \to W_h(\Omega) \times \Lambda_h^0(\Omega)$$

by replacing A_h^{-1} in (3.8) by B_{J+1}^m.

Extensions to multilevel (additive) formulations, other cycles and multiple smoothings are straightforward. The resulting multigrid and multilevel preconditioners constructed for mixed finite elements in this manner are significantly different from the multigrid and multilevel preconditioners previously constructed for mixed finite elements in [87, 88, 37, 14, 86, 2]. This construction is closely related to the construction of a hierarchical basis preconditioner in [92] for piecewise linear nonconforming finite elements. To the extent that $\hat{A}_2^m$ and $\hat{B}_{J+1}^m$ use conforming finite elements for the coarser spaces, they bear some resemblance to a multigrid method proposed for mixed finite elements in [12] and to some nonconforming finite element preconditioners in [64, 66].

3.2.2 A Second V-cycle Preconditioner

In [12], Bramble, Pasciak and Xu proposed a multigrid algorithm for the dual variable problem arising from the mixed finite element method without hybridization that use conforming finite elements for the coarser spaces. We extend their algorithm in a natural way to the dual problem for hybrid mixed finite element spaces. Because of the close relationship between hybrid mixed finite elements and nonconforming elements (see [4, 2]), the spaces used in our extension are also closely related to the spaces used in the multilevel methods due to Bramble and Oswald [64, 66].

Using notation from the previous section, recall that U_J is the space of continuous functions that vanish on $\partial\Omega$ and are piecewise linear, bilinear, or trilinear as appropriate for the dimension and shape of the element with respect to $\mathcal{T}_h$. Define a prolongation map

$$I_c^m : U_J \to W_h(\Omega) \times \Lambda_h^0(\Omega)$$

by

$$I_c^m \phi = [P_{W_h(\Omega)}\phi, P_{\Lambda_h^0(\Omega)}\phi],$$

where $P_{W_h(\Omega)}$ and $P_{\Lambda_h^0(\Omega)}$ are $L^2(\Omega)$-projection onto $W_h(\Omega)$ and $\Lambda_h^0(\Omega)$, respectively. Note that for all but the lowest order mixed finite element spaces, the conditions above are natural injection since $U_J \subset W_h(\Omega)$ and $(U_J)_{|\partial\mathcal{T}_h} \subset \Lambda_h^0(\Omega)$. We use $(I_c^m)^T$, the adjoint of prolongation for the restriction operator.

Algorithm 3.3 Define a V-cycle preconditioner

$$C_J^m : W_h(\Omega) \times \Lambda_h^0(\Omega) \to W_h(\Omega) \times \Lambda_h^0(\Omega)$$

for $[r, \rho] \in W_h(\Omega) \times \Lambda_h^0(\Omega)$ by

$$\begin{aligned} [q, \mu] &= R_m^T[r, \rho], && (3.10)\\ [\hat{q}, \hat{\mu}] &= [q, \mu] + I_c^m B_J^m (I_c^m)^T([r, \rho] - D_h[q, \mu]), && (3.11)\\ C_J^m[r, \rho] &= [\hat{q}, \hat{\mu}] + R_m([r, \rho] - D_h[\hat{q}, \hat{\mu}]). && (3.12) \end{aligned}$$

Extensions to multilevel (additive) formulations, other cycles and multiple smoothings are straightforward.

3.3 Numerical Experiments

In this section we report on numerical experiments using several preconditioners for D_h. The preconditioners used in the numerical experiments are:

- $\mathcal{N} A_h^{-1} \mathcal{N}^*$: defined in Section 3.1.1,
- $\mathcal{N} B_{J+1}^m \mathcal{N}^*$: defined in Section 3.1.2,
- $\widehat{A}_2^m$: defined in Section 3.2.1,
- $\widehat{B}_{J+1}^m$: defined in Section 3.2.1,
- C_J^m: defined in Section 3.2.2,
- line-$\widehat{B}_{J+1}^m$: $\widehat{B}_{J+1}^m$ with line Gauss-Seidel smoothings on the conforming spaces,
- line-C_J^m: C_J^m with line Gauss-Seidel smoothings on the conforming spaces,
- diag(D_h): Jacobi preconditioning,

- IC(0): Incomplete Cholesky preconditioner with zero additional fill-in,
- IC(1): Incomplete Cholesky preconditioner with one level of fill-in.

The implementation of the Incomplete Cholesky and Jacobi preconditioners are from the **NSPCG** package [63]. The multigrid method for the conforming spaces was implemented using Dendy's Black Box Multigrid [28, 29] which uses operator based prolongation and restriction operators instead of the projections introduced in the Algorithm 3.1 to yield a more robust algorithm in the case of highly varying coefficients.

The dual variable problems solved used the lowest order hybridized Raviart-Thomas space on rectangles. The elliptic problems considered were defined on rectangular domains

$$\Omega_a = (0, a) \times (0, 1),$$

and were of the form

$$\begin{aligned} -\nabla \cdot A \nabla p = 0 \quad & \text{in } \Omega_a, \\ p = 0 \quad & \text{on } \{y = 0\} \cap \partial\Omega_a, \\ p - 0.1\,(A\nabla p) \cdot \nu_{\Omega_a} = g \quad & \text{on } \{y = 1\} \cap \partial\Omega_a. \end{aligned}$$

Periodic boundary conditions in the first coordinate direction were imposed. The problems were discretized on a rectangular grid consisting of $n1 \times n2$ cells uniformly spaced in each coordinate direction. The following instances were considered:

- Test Problem I: $a = 1$, $A = I$, $g = \sin(2\pi x)$, and the mesh spacing h was refined to observe the relationship with the condition number;
- Test Problem II: $a = 2$, $n1 = 160$, $n2 = 80$, $g = -1$, and the coefficient A was taken from an elliptic problem arising in one step of a miscible displacement simulation (see [72]) and varies from 1.87 to 335370 as depicted in Figure 3.1;
- Test Problem III: $n1 = 128$, $n2 = 128$, $A = I$, $g = -1$, and the aspect ratio a was varied.

For the multigrid methods, refinement was carried out by subdividing each rectangular element into four elements by connecting the center with the midpoint of the sides. Continuous bilinear elements were used on the coarser grids.

All experiments were carried out on an otherwise unloaded IBM RS6000 Model 550 with 192 MB of memory in double precision arithmetic. The initial guess for the solution was zero, and preconditioned conjugate gradient iteration was continued until the error in the energy norm was reduced by a factor of $(10^{-6}/\sqrt{\kappa})$ where κ is an estimate of the condition number of the preconditioned operator. The estimate of the condition number was dynamically calculated by exploiting the similarity between conjugate gradients and Lanczos' method for finding eigenvalues using the code of Ashby, Manteuffel and Joubert [5]. The run times reported reflect the composite time to construct the preconditioner and to perform preconditioned conjugate gradient iterations until convergence. The formation of D_h is not included since it is the same for each preconditioner.

We see in Table 3.1 that the condition numbers for $\mathcal{N}A_h^{-1}\mathcal{N}^*$ and $\mathcal{N}B_{J+1}^m\mathcal{N}^*$ are uniformly bounded as predicted by Theorem 3.1 and Corollary 3.2. For the larger problems, the two most competitive methods are the two V-cycle preconditioners $\widehat{B}_{J+1}^m$ and C_J^m. In both Test Problem I and II, the use of $\widehat{B}_{J+1}^m$ leads to a lower conditioner number, but as implemented C_J^m is slightly faster as we see in Table 3.2 and Table 3.3.

Since the constants in Theorem 1.1 depend on the aspect ratio of the elements, the bounds on the condition numbers in Theorem 3.1 and Corollary 3.2 also depend on the aspect ratio. In Table 3.4, we see that the dependence on aspect ratio is strong. If line Gauss-Seidel smoothing is used instead of point Gauss-Seidel smoothing on the coarser grids, then the effect is mitigated.

Method of Preconditioning	Mesh Spacing 1/8	1/16	1/32	1/64	1/128	1/256
diag(D_h)	185.81	748.04	2997.09	11993.20	47962.30	182344.00
IC(0)	47.41	189.40	757.38	3029.31	12115.30	48437.70
IC(1)	3.58	9.71	29.13	101.64	394.22	1576.96
$\mathcal{N}A_h^{-1}\mathcal{N}^*$	2.93	2.96	2.97	2.98	2.98	2.98
$\mathcal{N}B_{J+1}^m\mathcal{N}^*$	2.97	2.99	3.00	2.99	2.99	2.99
$\widehat{A}_2^m$	1.33	1.33	1.33	1.33	1.33	1.33
$\widehat{B}_{J+1}^m$	1.36	1.50	1.54	1.54	1.54	1.53
C_J^m	2.07	2.09	2.19	2.21	2.22	2.22

Table 3.1 Condition Numbers for Test Problem I

Method of Preconditioning	Mesh Spacing 1/8	1/16	1/32	1/64	1/128	1/256
$\mathrm{diag}(D_h)$	0.02	0.09	0.61	5.38	59.46	DNC*
IC(0)	0.02	0.13	1.52	15.49	205.50	1636.36
IC(1)	0.02	0.08	0.95	7.22	109.65	781.94
$\mathcal{N}A_h^{-1}\mathcal{N}^*$	0.81	1.28	3.16	10.34	39.03	152.12
$\mathcal{N}B_{J+1}^m\mathcal{N}^*$	0.80	1.28	2.84	5.36	16.25	61.83
$\widehat{A}_2^m$	0.52	0.85	2.11	7.09	27.23	113.63
$\widehat{B}_{J+1}^m$	0.46	0.67	1.29	3.77	14.10	56.32
C_J^m	0.55	0.76	1.31	3.66	13.58	52.87

* Did not converge in 1000 iterations

Table 3.2 Run Time in CPU Seconds for Test Problem I

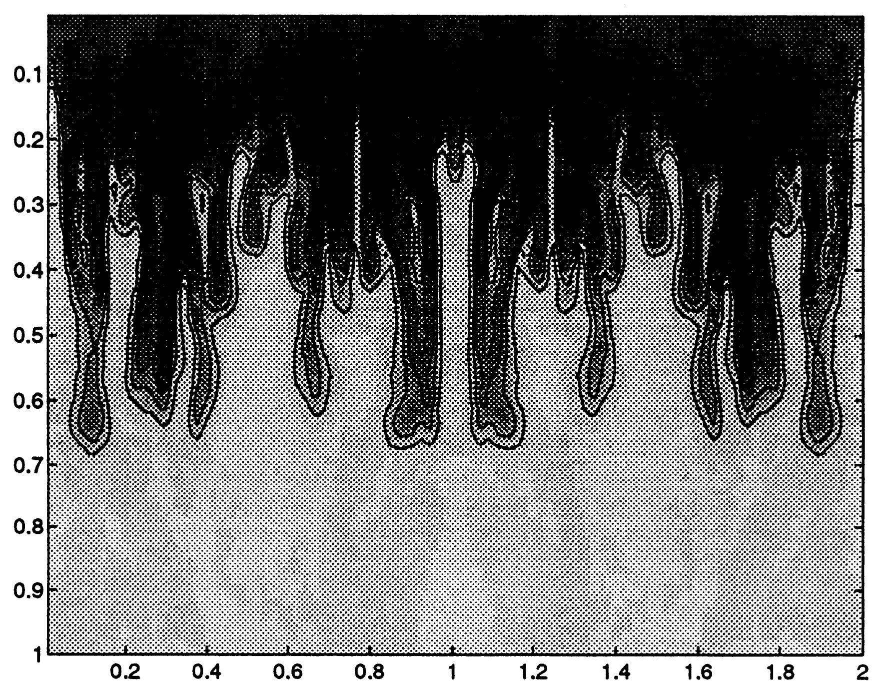

Figure 3.1 Coefficient in Test Problem II

Method of Preconditioning	Condition Number	Itera-tions	Run Time (secs)
diag(D_h)	117073.00	834	40.34
IC(0)	33565.60	425	80.42
IC(1)	3912.56	137	29.84
$\mathcal{N}A_h^{-1}\mathcal{N}^*$	4.99	13	55.47
$\widehat{A}_2^m$	1.62	7	318.60
$\widehat{B}_{J+1}^m$	3.23	11	10.19
C_J^m	3.89	13	7.86

Table 3.3 Results for Test Problem II

Method of	Aspect Ratio			
Preconditioning	1:1	2:1	4:1	8:1
diag(D_h)	47962.31	21170.49	10053.22	8466.91
IC(0)	12115.34	5021.93	1503.36	396.01
$\widehat{A}_2^m$	1.33	1.13	1.65	2.57
$\widehat{B}_{J+1}^m$	1.54	1.47	13.75	24.75
C_J^m	2.22	2.44	17.69	32.71
Line-$\widehat{B}_{J+1}^m$	1.77	1.76	2.10	3.16
Line-C_J^m	2.44	2.58	2.49	2.56

Table 3.4 Condition Numbers for Test Problem III

Chapter 4

Applications to Other Discretizations

The analysis of the domain decomposition and multigrid preconditioners presented in the previous two chapters followed almost directly from the isomorphism between the hybrid mixed finite element space and the conforming space of functions constructed in Section 1.4. Consequently, the properties of the hybrid mixed finite element discretization that were required for the analysis were limited to the existence of a nodal basis that permitted the construction of a refined triangulation $\widehat{\mathcal{T}}$ and the equivalence of the of the bilinear form expressed in Theorem 1.2 which followed immediately from the representation in Theorem 1.1. For the domain decomposition methods, the fact that the support of the discrete flux operator was contained in the support of its arguments was also used.

Many other discretizations of second order elliptic problems possess these properties. For instance, nonconforming finite element discretizations of (1.1)–(1.2) give rise to the quadratic form

$$d(\phi, \phi) = \sum_{\tau \in \mathcal{T}_h} \int_\tau A \nabla \phi \cdot \nabla \phi \, dx.$$

If the finite element space admits a nodal basis (i.e., it is of Lagrange-type), then it satisfies the equivalence in Theorem 1.1, a fact proven in [23]. The proof consists of identifying the local kernel as the constant functions and proving Lemma 1.2 for nonconforming elements by mapping back to a reference element. It is straightforward to check that other discretizations of (1.1)–(1.2) that possess these properties include most point-centered finite difference methods, the hybridized form of the expanded mixed finite element scheme in [47, 3], mixed finite element schemes in which quadrature is used to derive finite difference schemes (see Section 7 of [24]), collocation methods, and nonconforming finite element spaces of Lagrange-type (see, e.g., [44] in which many such spaces are constructed). Cell-centered finite difference methods including the "box method" [85] are added to the list by considering them as point centered methods on the dual mesh. Consequently, the constructions and analysis presented in the previous chapters can be applied to these methods with only minor modifications.

In the case of nonconforming spaces of Lagrange-type, the analysis of three domain decomposition methods, including the two Schwarz methods considered in Chapter 2, has been carried out in detail by the author in [23]. The results in [23] for the standard Schwarz method applied to the discretization by piecewise linear nonconforming finite elements were subsequently recovered by Brenner in [15] using a more general frame work. The framework of Brenner allows for the use of an arguably more natural coarse space that is also nonconforming. Brenner's theory also extends in a natural way to fourth order problems.

The isomorphism between piecewise linear nonconforming finite elements and a space of piecewise linear conforming elements used in [23] has also been used by Wohlmuth and Hoppe in [92]. They constructed a hierarchical basis preconditioner for piecewise linear nonconforming finite elements from one for conforming finite elements. In Chapter 3, we constructed preconditioners for the dual problem (1.19) from conforming spaces in an analogous manner. It should be noted that Wohlmuth and Hoppe observed an independence of the condition number for this new hierarchical basis preconditioner. Existing theory would have predicted a growth in the condition number proportional to the number of levels which was observed for the hierarchical basis preconditioner due to Oswald [65].

Chapter 5

An Application to Sediment Transport

5.1 Model Formulation

In this section, we make a formal derivation of a model for the erosion of sediment under quasi-steady potential flow in a closed channel. The derivation is formal because even under the simplifying assumptions made herein, it is not known whether solutions to the equations exist and are sufficiently regular. The model is a highly simplified version of a model used in [79, 80].

We consider a two dimensional, x-periodic, time-dependent channel $\Omega(t)$ which we assume can be written as

$$\Omega(t) = \{(x,z) \mid \eta_b(t,x) < z < \eta_a(x), x \in I_{\mathrm{p}}\},$$

where I_{p} is the periodic unit interval. Denote the nonperiodic boundary of $\Omega(t)$ by $\Gamma_a \cup \Gamma_b(t)$ where

$$\Gamma_a = \{(x,z) \mid z = \eta_a(x), x \in I_{\mathrm{p}}\},$$
$$\Gamma_b(t) = \{(x,z) \mid z = \eta_b(t,x), x \in I_{\mathrm{p}}\}.$$

We assume that the channel is full of an incompressible, inviscid fluid. The top of the channel Γ_a is fixed and independent of time, while the bottom of the channel Γ_b is composed of an erodible sediment. We are interested in the evolution of the bedform η_b that determines Γ_b.

The motion of sediment particles is usually divided into three regimes (see, e.g., [83, 84, 77, 81]):

- suspended particle motion in which sediment is entrained into the fluid proper and buoyed by turbulent forces;
- rolling or sliding along the sediment bed;
- saltation in which particles bounce along the bed loosing contact with the bed by only a few grain diameters.

The latter two types of motion are typically lumped together under the name bed-load transport. We assume that no particles are suspended in the flow so that bed-load transport is the only mechanism of sediment motion. Moreover, we assume that all particles are of a single size and composition. Since bed-load transport takes place on or very near the sediment bed, we consider all bed-load transport to be tangential to Γ_b.

Let

$$\nu(t) = \{\nu_x(t), \nu_z(t)\}$$

denote the outward pointing unit normal to $\Omega(t)$, and

$$\tau(t) = \{-\nu_z(t), \nu_x(t)\}$$

the unit tangent. We make the nonstandard choice of sign for the unit tangent so that $\tau(t)$ on $\Gamma_b(t)$ will point in the streamwise direction for our simulations. Let $Q(t, x, z)$ for $(x, z) \in \Gamma_b(t)$ denote the tangential volumetric sediment flux per unit channel width measured in bed volumes. By bed volumes we mean the sediment volume divided by one minus the porosity of sediment bed. By conservation of sediment mass, we have

$$\frac{\partial(x, \eta_b)}{\partial t} \cdot \nu = \frac{\partial \eta_b}{\partial t} \cdot \nu_z = -\frac{\partial Q}{\partial \tau} \quad \text{on } \Gamma_b, \tag{5.1}$$

since no sediment in entrained into the fluid.

We make the following assumptions regarding the fluid flow field in $\Omega(t)$:

(A1) the flow is divergence free;

(A2) the flow is irrotational;

(A3) the flow is essentially time independent on the time scale of the bed evolution.

Assumption (A2) is a very strong assumption and will lead to a simple model for the flow. Assumption (A3) is also somewhat restrictive in that it implies that the flow reaches a steady state in a time much smaller than the relevant time scale for the evolution of Γ_b. The fact that the flow achieves a steady state at all can be an inappropriate assumption for turbulent flows.

Since the flow is two dimensional and divergence free, we can write the fluid velocity $\mathbf{u}$ as the curl of a scalar stream function ϕ; that is,

$$\mathbf{u}(t) = \nabla\times\phi(t) = \left(-\frac{\partial\phi(t)}{\partial z}, \frac{\partial\phi(t)}{\partial x}\right)^T = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \nabla\phi(t). \tag{5.2}$$

Since the flow is irrotational, we have that

$$0 = \nabla\times\mathbf{u}(t) = \frac{\partial \mathbf{u}_x(t)}{\partial z} - \frac{\partial \mathbf{u}_z(t)}{\partial x} = -\nabla\times(\nabla\times\phi(t)) = -\Delta\phi(t) \quad \text{in } \Omega(t). \tag{5.3}$$

To obtain a well-posed problem, we cannot impose all of the physically relevant boundary conditions. The two we specify are that there is no fluid flux across the top and bottom of the channel; that is,

$$\mathbf{u}\cdot\nu = 0 \quad \text{on } \Gamma_a, \tag{5.4}$$

$$\mathbf{u}\cdot\nu = \frac{\partial \eta_b}{\partial t}\cdot\nu_z = -\frac{\partial Q}{\partial \tau} \quad \text{on } \Gamma_b. \tag{5.5}$$

The last equality follows from (5.1). Since $\partial\phi/\partial\tau = -\mathbf{u}\cdot\nu$, we find by integrating (5.4) and (5.5) over Γ_a and Γ_b, respectively, that

$$\phi(t) = g_a(t) \quad \text{on } \Gamma_a, \tag{5.6}$$

$$\phi(t) = Q(t) + g_b(t) \quad \text{on } \Gamma_b, \tag{5.7}$$

for some functions $g_a(t)$, $g_b(t)$ that are spatially constant for every time t. Since the stream function ϕ is defined only up to a constant, we choose to normalize it by setting $g_a(t) = 0$.

The constant $g_b(t)$ is closely related to the flux of fluid through $\Omega(t)$ in the following way. Let $V(t,x)$ be the total discharge of fluid in the horizontal direction at time t; that is,

$$V(t,x) = \int_{\eta_b(t,x)}^{\eta_a(x)} \mathbf{u}(t)\cdot(1,0)\,dz = \phi(t,x,\eta_b(t,x)) - \phi(t,x,\eta_a(t,x)) = \phi(t,x,\eta_b(t,x)). \tag{5.8}$$

Integrating $V(t,x)$ over I_{p}, we see that the average horizontal discharge of fluid is given by

$$\overline{V}(t) = \int_{I_{\mathrm{p}}} V(t,x)\,dx = \int_{I_{\mathrm{p}}} \phi(t,x,\eta_b(t,x))\,dx.$$

Integrating (5.7), we have

$$\int_{I_{\mathrm{p}}} \phi(t,x,\eta_b(t,x)) - Q(t,x,\eta_b(t,x))\,dx = g_b(t);$$

hence,

$$\phi(t) - Q(t) + \int_{I_{\mathrm{p}}} Q(t,x,\eta_b(t,x))\,dx = \overline{V}(t) \quad \text{on } \Gamma_b. \tag{5.9}$$

Since the fluid is incompressible, a change in horizontal fluid flux must represent a vertical fluid flux and, hence, a change in elevation of the bed. By conservation of the fluid mass, we see that

$$\frac{\partial \eta_b(x,t)}{\partial t} = \frac{\partial V(x,t)}{\partial x} = \frac{\partial \phi(x,\eta_b(x),t)}{\partial x} \quad \text{on } I_{\mathrm{p}}. \tag{5.10}$$

The only component of the model left unspecified is the relationship between sediment flux Q and the fluid flow $\mathbf{u}$. In the numerical experiments presented in the next section, the simplest possible model is used, a linear relationship between sediment flux and the tangential fluid velocity

$$Q = \alpha \mathbf{u} \cdot \tau \quad \text{on } \Gamma_b. \tag{5.11}$$

Many other relationships have been proposed; see, for instance, Table 6.1 of [77].

To summarize, the following quantities are specified as input to the model:

- the initial channel geometry η_a and $\eta_b(0)$,
- the average horizontal discharge of fluid $\overline{V}(t)$, and
- the proportionality constant α between fluid velocity $\mathbf{u}$ and tangential sediment flux Q.

The evolution of the channel is determined by

- **A Fluid Flow Model:**

$$\mathbf{u}(t) - \nabla \times \phi(t) = 0 \quad \text{in } \Omega(t), \tag{5.12}$$

$$-\Delta \phi(t) = 0 \quad \text{in } \Omega(t), \tag{5.13}$$

$$\phi(t) = 0 \quad \text{on } \Gamma_a, \tag{5.14}$$

$$\phi(t) - Q(t) + \int_{I_{\mathrm{p}}} Q(t,x,\eta_b(x))\, dx = \overline{V}(t) \quad \text{on } \Gamma_b; \tag{5.15}$$

- **An Erosion Model:**

$$\frac{\partial \eta_b(x,t)}{\partial t} - \frac{\partial \phi(x,\eta_b(x),t)}{\partial x} = 0 \quad \text{on } I_{\mathrm{p}}; \tag{5.16}$$

- **A Sediment Flux Model:**

$$Q(t) = \alpha \mathbf{u}(t) \cdot \tau = \alpha \frac{\partial \phi(t)}{\partial \nu} \quad \text{on } \Gamma_b. \tag{5.17}$$

5.2 Discretization

In this and subsequent sections, we describe an algorithm for calculating a mixed finite element approximation to (5.12)–(5.17). We refrain from introducing additional notation to differentiate between the functions in (5.12)–(5.17) and their finite element approximations since all functions in the subsequent sections will be from appropriate finite element spaces.

Let

$$\delta_{\widehat{x}} = \{\widehat{x}_0, \widehat{x}_1, \ldots, \widehat{x}_N\}$$

with

$$0 = \widehat{x}_0 < \widehat{x}_1 < \ldots < \widehat{x}_{N-1} < \widehat{x}_N = 1$$

denote a partition of the unit periodic interval I_p. For each t, let $\eta_b(t) : I_p \to \mathbb{R}$ be a continuous function that is piecewise linear with respect to $\delta_{\widehat{x}}$.

A three step time-splitting technique is used to advance the approximation of the bedform η_b through a sequence of discrete time levels $\{t^n\}$.

Algorithm 5.1 Given a current time t^n and a current bedform $\eta_b(t^n)$, compute an advance time t^{n+1} and bedform $\eta_b(t^{n+1})$ as follows.

1. Compute a mixed finite element approximation to the flow equations (5.12)–(5.15) and (5.17) in $\Omega(t^n)$.
2. Compute an acceptable advanced time t^{n+1} subject to a "CFL-like" constraint.
3. Update $\eta_b(t^{n+1})$ using the Erosion Model (5.16).

A detailed description of these three steps follows.

5.2.1 Discretization of the Flow Equations in $\Omega(t)$

Let I denote the (non-periodic) unit interval. Denote the unit square, periodic in the first coordinate direction by $\widehat{\Omega} = I_p \times I$, and its outward normal by $\widehat{\nu}$. Let

$$\delta_{\widehat{z}} = \{\widehat{z}_0, \widehat{z}_1, \ldots, \widehat{z}_M\}$$

with

$$0 = \widehat{z}_0 < \widehat{z}_1 < \ldots < \widehat{z}_{M-1} < \widehat{z}_M = 1$$

denote a partition of I. Let $\widehat{\mathcal{T}}$ denote the tensor product partitioning $\delta_{\widehat{x}} \otimes \delta_{\widehat{z}}$ of $\widehat{\Omega}$ into rectangles, and $\partial\widehat{\mathcal{T}}$ the union of the boundaries of cells in $\widehat{\mathcal{T}}$. For convenience, we use the same partitioning of I_{p} for the discretization of η_b above and $\widehat{\Omega}$ here.

If $\eta_a(x) > \eta_b(t,x)$ for $x \in I_{\mathrm{p}}$, then

$$F(t,\widehat{x},\widehat{z}) = (\widehat{x}, \eta_b(t,\widehat{x}) + \widehat{z}(\eta_a(\widehat{x}) - \eta_b(t,\widehat{x})))$$

is a bijection between $\widehat{\Omega}$ and $\Omega(t)$. Let DF denote its Jacobian matrix, and set

$$J(\widehat{x},\widehat{z}) = |\det(DF(\widehat{x},\widehat{z}))| \quad (\widehat{x},\widehat{z}) \in \widehat{\Omega},$$

$$J_{\widehat{\nu}}(\widehat{x},\widehat{z}) = J(\widehat{x},\widehat{z})|DF^{-1}(\widehat{x},\widehat{z})\widehat{\nu}| \quad (\widehat{x},\widehat{z}) \in \partial\widehat{\Omega}.$$

Here and throughout the rest of this section, we suppress the time variable when it causes no confusion, noting that the flow variables depend on time only through the domain $\Omega(t)$ (equivalently, F) and the boundary condition involving $\overline{V}(t)$.

Denote by $\widetilde{\mathbf{V}}_h(\widehat{\Omega}) \times W_h(\widehat{\Omega}) \times \Lambda_h(\widehat{\Omega})$ the lowest order Raviart-Thomas hybrid mixed finite elements on $\widehat{\mathcal{T}}_h$ (see Section 1.2) with periodicity imposed in first coordinate direction. Define the hybrid mixed finite element spaces $\widetilde{\mathbf{V}}_h(\Omega) \times W_h(\Omega) \times \Lambda_h(\Omega)$ on

$$\mathcal{T}_h = \bigcup_{\widehat{\tau} \in \widehat{\mathcal{T}}_h} F(\tau),$$

the "triangulation" of Ω induced by $\widehat{\mathcal{T}}_h$ and F, as follows (see [82]). For $\widehat{\mathbf{v}} \in \widetilde{\mathbf{V}}_h(\widehat{\Omega})$, $\widehat{w} \in W_h(\widehat{\Omega})$, define $\mathbf{v} \in \widetilde{\mathbf{V}}_h(\Omega)$ and $w \in W_h(\Omega)$ at $(x,z) = F(\widehat{x},\widehat{z})$ by

$$\mathbf{v}(x,z) = \frac{1}{J(\widehat{x},\widehat{z})} DF(\widehat{x},\widehat{z})\widehat{\mathbf{v}}(\widehat{x},\widehat{z}),$$

$$w(x,z) = \widehat{w}(\widehat{x},\widehat{z}).$$

For $\widehat{\mu} \in \Lambda_h(\widehat{\Omega})$, define $\mu \in \Lambda_h(\Omega)$ at $(x,z) = F(\widehat{x},\widehat{z})$ for $(\widehat{x},\widehat{z}) \in \partial\widehat{\mathcal{T}}_h$ by

$$\mu(x,z) = \widehat{\mu}(\widehat{x},\widehat{z}).$$

Let $\Lambda_h^0(\widehat{\Omega}) \subset \Lambda_h(\widehat{\Omega})$ and $\Lambda_h^0(\Omega) \subset \Lambda_h(\Omega)$ denote those functions that vanish on $\{(\widehat{x},\widehat{z}) \in \partial\widehat{\Omega} \mid \widehat{x} = 1\}$ and Γ_a, respectively. By the hybrid mixed finite element approximation to the flow equations (5.12)–(5.15), we mean the triple

$$\left\{ \mathbf{u} \equiv -\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \widetilde{\mathbf{u}}, \phi, \lambda \right\}$$

where

$$\{\widetilde{\mathbf{u}}, \phi, \lambda\} \in \widetilde{\mathbf{V}}_h(\Omega) \times W_h(\Omega) \times \Lambda_h^0(\Omega)$$

satisfies

$$\sum_{\widehat{\tau}\in\widehat{\mathcal{T}}_h} \left(\int_{F(\widehat{\tau})} \widetilde{\mathbf{u}} \cdot \mathbf{v}\, dx - \int_{F(\widehat{\tau})} \phi \nabla\cdot \mathbf{v}\, dx + \int_{F(\partial\widehat{\tau})} \lambda \mathbf{v}\cdot\nu\, ds \right) = 0 \quad \forall \mathbf{v} \in \widetilde{\mathbf{V}}_h(\Omega), \tag{5.18}$$

$$-\sum_{\widehat{\tau}\in\widehat{\mathcal{T}}_h} \int_{F(\widehat{\tau})} w \nabla\cdot \widetilde{\mathbf{u}}\, dx = 0 \quad \forall w \in W_h(\Omega), \tag{5.19}$$

$$\sum_{\widehat{\tau}\in\widehat{\mathcal{T}}_h} \int_{F(\partial\widehat{\tau})\setminus\Gamma_b} \mu \widetilde{\mathbf{u}}\cdot\nu\, ds = 0 \quad \forall \mu \in \Lambda_h^0(\Omega), \tag{5.20}$$

$$\sum_{\widehat{\tau}\in\widehat{\mathcal{T}}_h} \int_{F(\partial\widehat{\tau})\cap\Gamma_b} \mu \left(\frac{1}{\alpha}\lambda + \widetilde{\mathbf{u}}\cdot\nu + I_1 \right) ds = \sum_{\widehat{\tau}\in\widehat{\mathcal{T}}_h} \int_{F(\partial\widehat{\tau})\cap\Gamma_b} \mu \frac{1}{\alpha} \overline{V}(t)\, ds \quad \forall \mu \in \Lambda_h^0(\Omega), \tag{5.21}$$

where

$$I_1 = -\int_{I_p} \widetilde{\mathbf{u}}(\overline{x}, \eta_b(t, \overline{x}))\cdot\nu\, d\overline{x}.$$

Equations (5.18)–(5.20) are discussed in Section 1.2, and (5.21) is a weak form of (5.15) with (5.17). The rotation applied to $\widetilde{\mathbf{u}}$ above is to correct for the fact that $-\widetilde{\mathbf{u}}$ is an approximation to the gradient of ϕ, but $\mathbf{u}$ is an approximation to the curl.

We will do most of our calculations on the reference domain $\widehat{\Omega}$. Additionally, we can avoid the difficulties introduced by the nonlocal, nonsymmetric term I_1 in (5.21) by taking advantage of the linearity of the sediment flux function in the following way. Let

$$\{\widehat{\mathbf{u}}, \widehat{\phi}, \widehat{\lambda}\} \in \widetilde{\mathbf{V}}_h(\widehat{\Omega}) \times W_h(\widehat{\Omega}) \times \Lambda_h^0(\widehat{\Omega})$$

satisfy

$$\sum_{\widehat{\tau}\in\widehat{\mathcal{T}}_h} \left(\int_{\widehat{\tau}} \widehat{A}\widehat{\mathbf{u}} \cdot \widehat{\mathbf{v}}\, d\widehat{x} - \int_{\widehat{\tau}} \widehat{\phi} \nabla\cdot \widehat{\mathbf{v}}\, d\widehat{x} + \int_{\partial\widehat{\tau}} \widehat{\lambda} \widehat{\mathbf{v}}\cdot\widehat{\nu}\, d\widehat{s} \right) = 0 \quad \forall \widehat{\mathbf{v}} \in \widetilde{\mathbf{V}}_h(\widehat{\Omega}), \tag{5.22}$$

$$-\sum_{\widehat{\tau}\in\widehat{\mathcal{T}}_h} \int_{\widehat{\tau}} \widehat{w} \nabla\cdot \widehat{\mathbf{u}}\, d\widehat{x} = 0 \quad \forall \widehat{w} \in W_h(\widehat{\Omega}), \tag{5.23}$$

$$\sum_{\widehat{\tau}\in\widehat{\mathcal{T}}_h} \int_{\partial\widehat{\tau}\setminus F^{-1}(\Gamma_b)} \widehat{\mu} \widehat{\mathbf{u}}\cdot\widehat{\nu}\, d\widehat{s} = 0 \quad \forall \widehat{\mu} \in \Lambda_h^0(\widehat{\Omega}), \tag{5.24}$$

$$\sum_{\widehat{\tau}\in\widehat{\mathcal{T}}_h} \int_{\partial\widehat{\tau}\cap F^{-1}(\Gamma_b)} \widehat{\mu} \left(\frac{1}{\alpha} \widehat{\lambda} J_{\widehat{\nu}} + \widehat{\mathbf{u}}\cdot\widehat{\nu} \right) d\widehat{s} = \sum_{\widehat{\tau}\in\widehat{\mathcal{T}}_h} \int_{\partial\widehat{\tau}\cap F^{-1}(\Gamma_b)} \widehat{\mu} J_{\widehat{\nu}}\, d\widehat{s} \quad \forall \widehat{\mu} \in \Lambda_h^0(\widehat{\Omega}), \tag{5.25}$$

where $\widehat{A} = \frac{1}{J}(DF)^T(DF)$. Let

$$I_2 = -\int_{I_p} \widehat{\mathbf{u}}(\overline{\widehat{x}}, 0) \cdot (0, -1)\, d\overline{\widehat{x}},$$

and set

$$\beta = \frac{\overline{V}(t)}{\alpha}(1 + I_2)^{-1}.$$

Scaling $\widehat{\mathbf{u}}$, $\widehat{\phi}$ and $\widehat{\lambda}$ by β leaves equations to (5.22)–(5.24) unchanged. By adding $I_2\widehat{\mu}J_{\widehat{\nu}}$ to both integrands of (5.25) and multiplying by β, we see that

$$\begin{aligned}\sum_{\widehat{\tau}\in\widehat{\mathcal{T}}_h}\int_{\partial\widehat{\tau}\cap F^{-1}(\Gamma_b)}\widehat{\mu}\left(\frac{\beta}{\alpha}\widehat{\lambda}J_{\widehat{\nu}} + \beta\widehat{\mathbf{u}}\cdot\widehat{\nu} + \beta I_2 J_{\widehat{\nu}}\right)d\widehat{s} &= \sum_{\widehat{\tau}\in\widehat{\mathcal{T}}_h}\int_{\partial\widehat{\tau}\cap F^{-1}(\Gamma_b)}\widehat{\mu}\beta(1+I_2)J_{\widehat{\nu}}\,d\widehat{s} \\ &= \sum_{\widehat{\tau}\in\widehat{\mathcal{T}}_h}\int_{\partial\widehat{\tau}\cap F^{-1}(\Gamma_b)}\widehat{\mu}\frac{1}{\alpha}\overline{V}(t)J_{\widehat{\nu}}\,d\widehat{s},\end{aligned}$$

an appropriate weak form of (5.21). Hence, we can calculate the hybrid mixed finite element approximation to (5.12)–(5.15) by taking

$$\begin{aligned}\mathbf{u} &= -\beta\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\left(\frac{1}{J}DF\widehat{\mathbf{u}}\right)\circ F^{-1}, \\ \phi &= \beta\widehat{\phi}\circ F^{-1}, \\ \lambda &= \beta\widehat{\lambda}\circ F^{-1}.\end{aligned}$$

Again, the rotation applied to $\widehat{\mathbf{u}}$ corrects for the fact that $\mathbf{u}$ approximates the curl of ϕ, but $-\widehat{\mathbf{u}}$ approximates the gradient of $\widehat{\phi}$.

5.2.2 Time Step Calculation

Recognizing that the Erosion Model (5.16) is formally hyperbolic, we expect the explicit time stepping in Algorithm 5.1 to be only conditionally stable (see, e.g., [49]). The quantity

$$u_{\max} = \sup_{x\in I_p}|\mathbf{u}(t^n, x, \eta_b(x))\cdot\tau|$$

represents a measure of the speed at which information is propagated along the boundary Γ_b. We enforce a "CFL-type" condition by taking

$$t^{n+1} = t^n + (1-\epsilon)(u_{\max})^{-1}\min_{i=1,\dots,N}(\widehat{x}_i - \widehat{x}_{i-1}), \tag{5.26}$$

where ϵ is some quantity strictly between 0 and 1. In the numerical experiments under steady flow, we take $\epsilon = 0.01$. A similar condition was imposed by Chan Hong, et al. [21] for modeling the interface between fresh and salt water in a saturated porous media.

If the forcing function $\overline{V}(t)$ is unsteady, the time step can be further reduced to capture the dynamics of the boundary condition.

5.2.3 Discretization of the Erosion Model

We discretize (5.16) using techniques appropriate for hyperbolic equations. Since $\lambda(t, x, \eta_b(t, x))$ is an approximation to the trace of $\phi(t)$ on Γ_b, we treat it as the evaluation of the flux function at $\eta_b(t)$. We note that $\lambda(t, x, \eta_b(t, x))$ depends in a highly nonlinear and nonlocal manner on $\eta_b(t)$ through the domain $\Omega(t)$ in (5.12)–(5.15); hence, our use of standard methods for hyperbolic conservation laws is formal at best.

Since $\lambda(t, x, \eta_b(t, x))$ is discontinuous at the nodes in the partition $\delta_{\widehat{x}}$ of I_p, we define a point centered value by averaging. For $i = 0, \ldots, N-1$, let $x_{i+\frac{1}{2}} = \frac{1}{2}(x_{i+1} + x_i)$, and set

$$\lambda^n_{i+\frac{1}{2}} = \lambda(t^n, x_{i+\frac{1}{2}}, \eta_b(t^n, x_{i+\frac{1}{2}})) = \widehat{\lambda}(t^n, x_{i+\frac{1}{2}}, 0).$$

For $i = 0, \ldots, N$, set

$$\lambda^n_i = \frac{1}{2}(\lambda^n_{i-\frac{1}{2}} + \lambda^n_{i+\frac{1}{2}}),$$

where we take $\lambda^n_{N+\frac{1}{2}} = \lambda^n_{\frac{1}{2}}$ and $\lambda^n_{-\frac{1}{2}} = \lambda^n_{N-\frac{1}{2}}$ by periodicity.

Assuming that the flow is from left to right, i.e. $\mathbf{u} \cdot (1,0) \geq 0$, we update the bedform by the upwind formula (see, e.g., [49])

$$\eta_b(t^{n+1}, x_i) = \eta_b(t^n, x_i) + \frac{(t^{n+1} - t^n)}{(x_i - x_{i-1})}(\lambda^n_i - \lambda^n_{i-1}) \tag{5.27}$$

for $i = 0, \ldots, N$ with $\lambda^n_{-1} = \lambda^n_{N-1}$.

5.3 Two Numerical Simulations

Not surprisingly, the most computationally taxing portion of Algorithm 5.1 is the solution of (5.22)–(5.25) arising the in flow calculations in Step 1. At each time step, a new linear system must be formed since the domain $\Omega(t)$ and $\widehat{A}$ have changed from the previous time step. As in Section 1.3, we reduce (5.22)–(5.25) to a dual variable problem involving only the unknowns $\widehat{\lambda}$ and $\widehat{\phi}$. Preconditioned conjugate gradients is then applied to the dual problem using one of the preconditioners constructed in Chapter 3. Because the dual problems at consecutive time steps may have only changed by a small amount, considerable savings in computational time may be achieved by using the same preconditioner for multiple time steps. In this way, the cost of forming the preconditioner can be amortized over multiple time steps. For the simulations presented here, the multigrid V-cycle preconditioner $\widehat{B}^m_{J+1}$ described in Section 3.2.1 was the optimal choice of the preconditioners studied in Chapter 3 since the time to apply the preconditioner was approximately the same as C_J^m, but the number of iterations needed was smaller since the resulting condition number is better.

In the first experiment, we consider the erosion of a sinusoidal bed-form under steady flow. The input parameters for the model are

$$\eta_a(x) = \frac{1}{2}, \qquad \eta_b(0,x) = \frac{1}{20}\sin(2\pi x),$$
$$\overline{V}(t) = 1/2, \qquad \alpha = 0.1.$$

The periodic interval I_p was divided into 128 intervals of uniform length. The flow equations were discretized using a uniformly spaced 128x64 grid for $\widehat{\Omega}$. The preconditioner was only reinitialized when the CPU time required to do so was less than the time spent taking "extra iterations" in subsequent time steps due to using the old preconditioner. The number of "extra iterations" per time step was assumed to be the number of iterations take at the current time step less the number of iterations taken at the time step when the preconditioner was last reinitialized.

Since the bedforms tend to be advected down stream, to further forestall the need to reinitialize the preconditioner, a correlation length was computed between the current bedform and the bedform when the preconditioner was first computed. Because the solution to the fluid flow equations are periodic in the streamwise direction, the preconditioner for the current bedform was evaluated by first rotating the residual by the correlation length, applying the old preconditioner and shifting the preconditioned residual back. Using this procedure, the preconditioner was only initialized six times in the 1250 time steps of the simulation.

Figure 5.1 displays the bedforms at several early times in the simulation. Note that the bedform elevation is exaggerated in Figure 5.1 and in some subsequent figures. We see that the bedform is both advected in the streamwise direction and eroded. This phenomenon is also shown clearly in Figure 5.3 where contours corresponding to extreme elevations are pinched out. Furthermore, as the bedform decays it continues to be advected at approximately the same speed. The decay in the extremal elevations of the bedform is shown in Figure 5.2 and is not initially monotone.

In the second experiment, the top of the channel has a semi-circular constriction. The fluid moves at a higher velocity through the constriction and scours a hole in an initially flat bed. In this simulation, the model parameters are

$$\eta_b(0,x) = 0, \qquad \overline{V}(t) = 1/8, \qquad \alpha = 0.1,$$

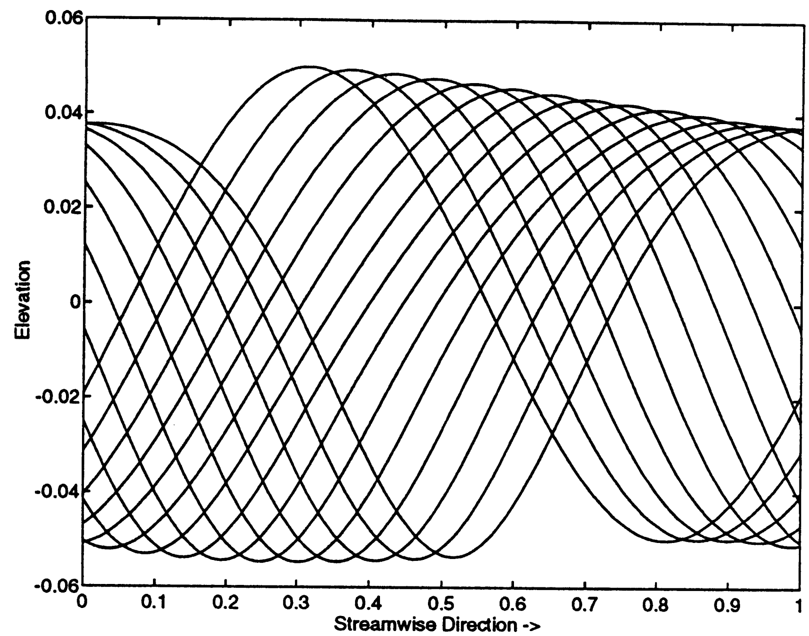

Figure 5.1 Early Evolution of Initially Sinusoidal Bedform

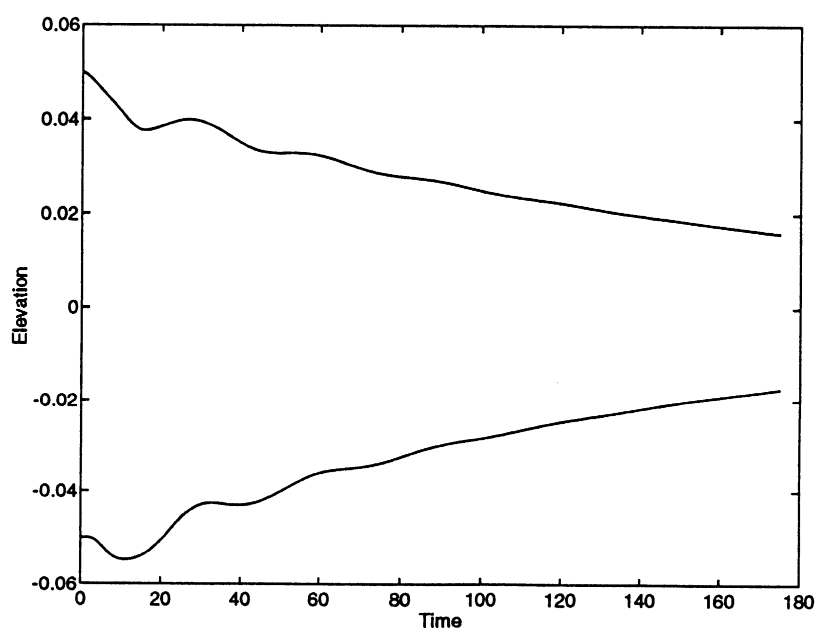

Figure 5.2 Extremal Elevations for Initially Sinusoidal Bedform

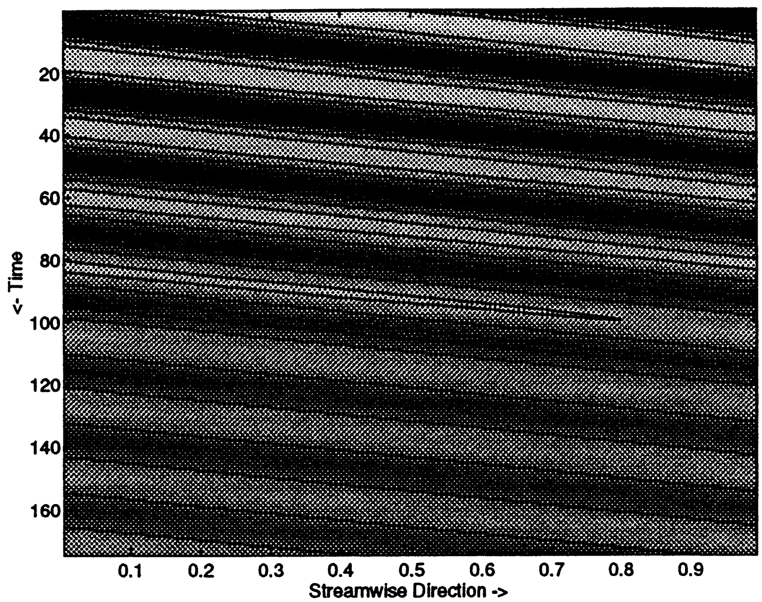

Figure 5.3 Elevation Contours of the Initially Sinusoidal Bedform

and the top of the channel is at a constant height of 1/8 except for a semi-circular indentation centered at x=0.53125 of radius 1/32, i.e.

$$\eta_a(x) = \begin{cases} 0.125 & \text{if } 0 \leq x \leq 0.5 \\ 0.125 - \sqrt{(0.03125)^2 - (x - 0.53125)^2} & \text{if } 0.5 \leq x \leq 0.5625 \\ 0.125 & \text{if } 0.5625 \leq x \leq 1 \end{cases}$$

The periodic interval I_{p} was divided into 512 uniform intervals. The flow equations were discretized on a uniformly spaced 512x64 grid on $\hat{\Omega}$.

As shown in Figure 5.4, a scoured region develops slightly before the constriction in the channel and the plug of scoured sediment is advected down stream. Since the channel is periodic, the sediment plug is eventually advected through the scoured region and the dynamics become more complex. The dynamics can also be seen in Figures 5.5–5.7.

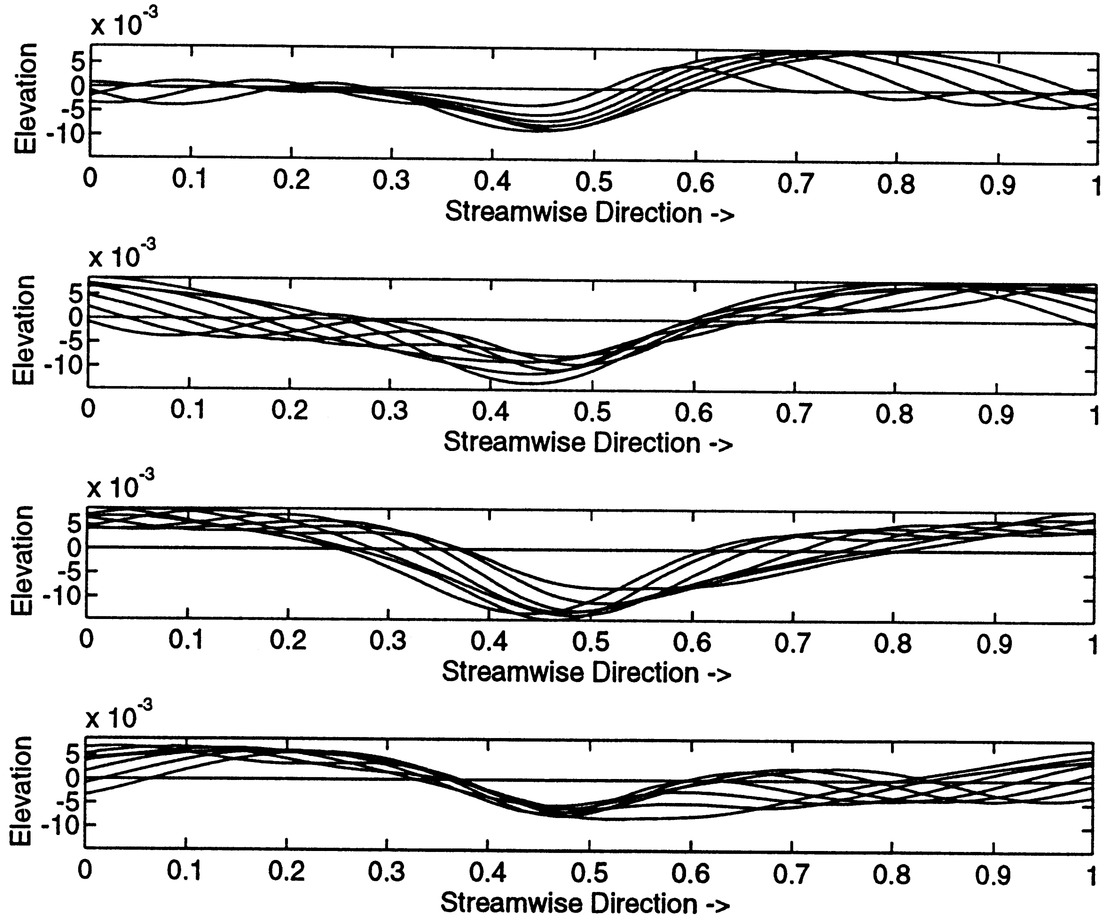

Figure 5.4 Evolution of Scour Hole

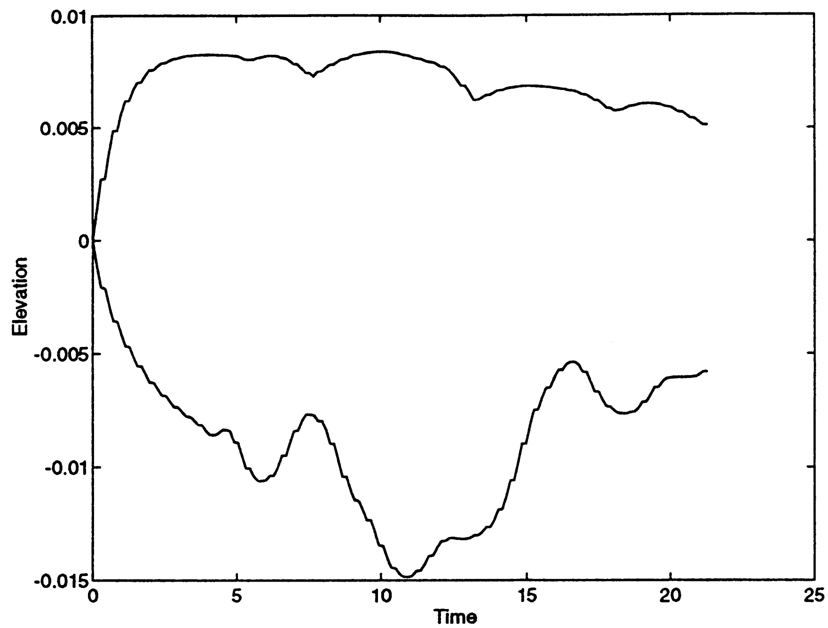

Figure 5.5 Extremal Elevations in Scour Experiment

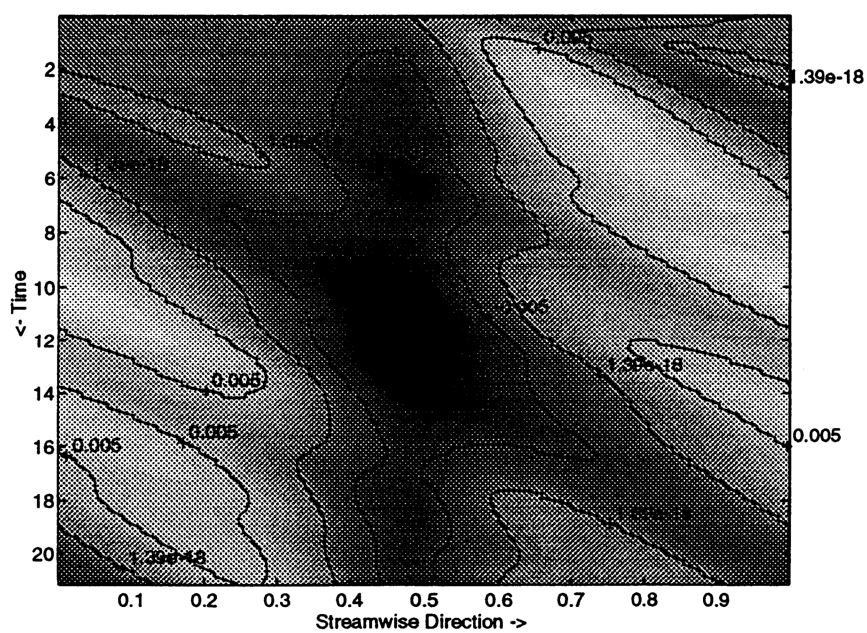

Figure 5.6 Elevation Contours in Scour Experiment

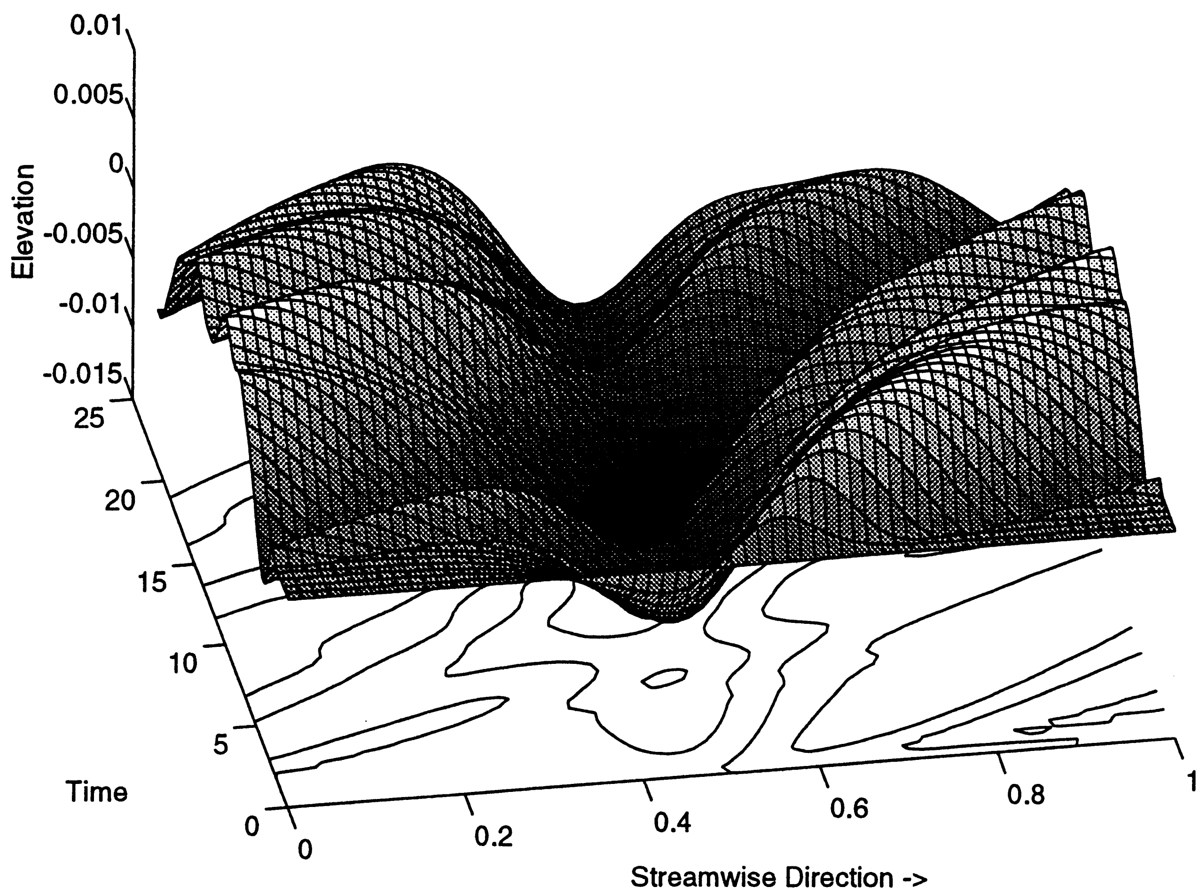

Figure 5.7 Elevation Surface in Scour Experiment

Bibliography

[1] R. ADAMS, *Sobolev Spaces*, Academic Press, 1975.

[2] T. ARBOGAST AND Z. CHEN, *On the implementation of mixed methods as nonconforming methods for second order elliptic problems*, Dept. Comp. Appl. Math. TR93-31, Rice University, Aug. 1993.

[3] T. ARBOGAST, M. F. WHEELER, AND I. YOTOV, *Mixed finite elements for elliptic problems with tensor coefficients as finite differences*, Dept. Comp. Appl. Math. TR94-02, Rice University, Houston, TX 77251, Jan. 1994.

[4] D. N. ARNOLD AND F. BREZZI, *Mixed and nonconforming finite element methods: Implementation, postprocessing and error estimates*, Math. Model. and Numer. Anal., 19 (1985), pp. 7–32.

[5] S. F. ASHBY, T. A. MANTEUFFEL, AND P. E. SAYLOR, *A taxonomy for conjugate gradient methods*, SIAM J. Numer. Anal., 27 (1990), pp. 1542–1568.

[6] R. E. BANK AND C. C. DOUGLAS, *Sharp estimates for multigrid rates of convergence with general smoothing and acceleration*, SIAM J. Numer. Anal., 22 (1985), pp. 617–633.

[7] R. BECKIE, E. F. WOOD, AND A. A. ALDAMA, *A multigrid acceleration for domain decomposition mixed finite element simulation of groundwater flow*, in Computational methods in water resources, IX, Vol. 1, Comput. Mech., Southampton, 1992.

[8] D. BRAESS AND W. HACKBUSCH, *A new convergence proof for the multigrid method including the V cycle*, SIAM J. Numer. Anal., 20 (1983), pp. 967–975.

[9] J. H. BRAMBLE AND J. E. PASCIAK, *New estimates for multilevel algorithms including the V-cycle*, Math. of Comp., 60 (1993), pp. 447–471.

[10] J. H. Bramble, J. E. Pasciak, and A. H. Schatz, *An iterative method for elliptic problems on regions partitioned into substructures*, Math. of Comp., 46 (1986), pp. 361–369.

[11] J. H. Bramble, J. E. Pasciak, J. Wang, and J. Xu, *Convergence estimates for product iterative methods with applications to domain decomposition*, Math. of Comp., 57 (1991), pp. 1–21.

[12] J. H. Bramble, J. E. Pasciak, and J. Xu, *The analysis of multigrid algorithms with nonnested spaces or noninherited quadratic forms*, Math. Comp., 56 (1991), pp. 1–34.

[13] ——, *Parallel multilevel preconditioners correction: A unifying approach*, Math. Comp., (1991).

[14] S. C. Brenner, *A multigrid alorithm for the lowest order Raviart-Thomas mixed triangular finite element method*, SIAM J. Numer. Anal., 29 (1992), pp. 647–678.

[15] ——, *Two-level additive Schwaz preconditioners for nonconfomring fintie element methods.* To appear in the *Proceedings of the Seventh International Symposium on Domain Decomposition Methods for Partial Differential Equations*, Oct. 1993.

[16] F. Brezzi, J. Douglas, Jr, R. Durán, and M. Fortin, *Mixed finite elements for second order elliptic problems in three variables*, Numer. Math., 51 (1987), pp. 237–250.

[17] F. Brezzi, J. Douglas, Jr, M. Fortin, and L. D. Marini, *Efficient rectangular mixed finite elements in two and three space variables*, Math. Modelling and Num. Ana., 21 (1987), pp. 581–604.

[18] F. Brezzi, J. Douglas, Jr, and L. D. Marini, *Two families of mixed finite elements for second order elliptic problems*, Numer. Math., 47 (1985), pp. 217–235.

[19] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, vol. 15 of Springer Series in Computational Mathematics, Springer-Verlag, New York, 1991.

[20] T. F. Chan and T. P. Mathew, *Domain decomposition algorithms*, in Acta Numerica, Cambridge University Press, 199, pp. 61–143.

[21] J. R. Chan Hong, C. J. van Duijn, D. Hilhorst, and J. van Kester, *The interface between fresh and salt groundwater: A numerical study*, IMA J. Appl. Math., 42 (1989), pp. 209–240.

[22] P. G. Ciarlet, *The finite element method for elliptic problems*, vol. 4 of Studies in mathematics and its applications, North-Holland Pub. Co., New York, 1978.

[23] L. C. Cowsar, *Domain decomposition methods for nonconforming finite element spaces of Lagrange-type*, in Melson et al. [58], pp. 93–109.

[24] ——, *Dual-variable Schwarz methods for mixed finite elements*, Dept. Comp. and Appl. Math. TR93-09, Rice University, Mar. 1993.

[25] L. C. Cowsar, J. Mandel, and M. F. Wheeler, *Balancing domain decomposition for mixed finite elements*, Dept. Comp. and Appl. Math. TR93-08, Rice University, Mar. 1993. Submitted.

[26] L. C. Cowsar, A. Weiser, and M. F. Wheeler, *Parallel multigrid and domain decomposition algorithms for elliptic equations*, in Proceedings of the Fifth International Symposium on Domain Decomposition Methods for Partial Differential Equations, D. E. Keyes et al., eds., SIAM, 1992, pp. 376–385. Also appeared as CRPC-TR91154.

[27] L. C. Cowsar and M. F. Wheeler, *Parallel domain decomposition method for mixed finite elements for elliptic partial differential equations*, in Proceedings of the Fourth International Symposium on Domain Decomposition Methods for Partial Differential Equations, R. Glowinski et al., eds., SIAM, 1991.

[28] J. E. Dendy, *Black box multigrid*, J. Comput. Phys., 48 (1982), pp. 366–386.

[29] ——, *Black box multigrid for periodic and singular problems*, Appl. Math. Comp., 25 (1988), pp. 1–10.

[30] J. Douglas and J. Roberts, *Global estimates for mixed methods for second order elliptic equations*, Math. of Comp., 44 (1985), pp. 39–52.

[31] M. DRYJA, *A method of domain decomposition for three-dimensional finite element elliptic problems*, in Glowinski et al. [40], pp. 43–61.

[32] M. DRYJA, B. F. SMITH, AND O. B. WIDLUND, *Schwarz analysis of iterative substructuring algorithms for elliptic problems in three dimensions*, tech. rep., Department of Computer Science, Courant Institute, June 1993.

[33] M. DRYJA AND O. B. WIDLUND, *An additive variant of the Schwarz alternating method for the case of many subregions*, Tech. Rep. 339, Courant Institute of Mathematical Sciences, 1987.

[34] ——, *Domain decomposition algorithms with small overlap*, Tech. Rep. 606, Department of Computer Science, Courant Institute, May 1992. To appear in SIAM J. Sci. Stat. Comput., Vol. 15 (3).

[35] ——, *The Neumann-Neumann method as an additive Schwarz method for finite element elliptic problems*, Tech. Rep. TR-626, Department of Computer Science, Courant Institute, March 1993.

[36] L. J. DURLOFSKY AND M. C. H. CHIEN, *Development of a mixed finite-element based compositional reservoir simulator*, in Proceedings of 12th SPE Symposium on Reservoir Simulation, SPE, Inc., 1993, Society of Petroleum Engineers, pp. 221–231.

[37] R. E. EWING AND J. WANG, *Analysis of multilevel decomposition iterative methods for mixed finite element methods.* Submitted to R.A.I.R.O., Modél. Math. Anal. Numér., 1991.

[38] ——, *Analysis of the Schwarz algorithm for mixed finite element methods*, R.A.I.R.O., Modél. Math. Anal. Numér., 26 (1992), pp. 739–756.

[39] B. FRAEJIS DE VEUBEKE, *Displacement and equilibrium models in the finite element method*, in Stress Analysis, O. C. Zienkiewicz and G. Holister, eds., Wiley, New York, 1965.

[40] R. GLOWINSKI ET AL., eds., *Proceedings of the First International Symposium on Domain Decomposition Methods for Partial Differential Equations*, SIAM, Jan. 1987.

[41] R. Glowinski, W. Kinton, and M. F. Wheeler, *Acceleration of domain decomposition algorithms for mixed finite elements by multi-level methods*, in Proceedings of the Third International Symposium on Domain Decomposition Methods for Partial Differential Equations, T. Chan et al., eds., SIAM, Mar. 1989, pp. 263–290.

[42] R. Glowinski and M. F. Wheeler, *Domain decomposition and mixed finite element methods for elliptic problems*, in Glowinski et al. [40], pp. 144–172.

[43] G. H. Golub and C. F. Van Loan, *Matix Computations*, The Johns Hopkins University Press, 2nd ed., 1990.

[44] J.-P. Hennart, J. Jaffre, and J. E. Roberts, *A constructive method for deriving finite elements of nodal type*, Numer. Math., 53 (1988), pp. 701–738.

[45] M. R. Hestenes, *Conjugate Direction Methods in Optimization*, Springer-Verlag, 1980.

[46] M. R. Hestenes and E. Stiefel, *Methods of conjugate gradients for solving linear systems*, J. Res. Nat. Bur. Stand., 49 (1952), pp. 409–436.

[47] J. Koebbe, *A computationally efficient modification of mixed finite element methods for flow problems with full transmissivity tensors*, Numer. Meth. for PDE's, 9 (1993), pp. 339–355.

[48] P. Le Tallec, *Domain decomposition methods in computational mechanics*, Comp. Mech. Adv., 1 (1994), pp. 121–220.

[49] R. J. LeVeque, *Numerical Methods for Conservation Laws*, Lectures in Mathematics, Birkhäuser, 1992.

[50] J. L. Lions and E. Magenes, *Non-homogeneous boundary value problems and applications*, Springer-Verlag, 1972.

[51] P. L. Lions, *On the Schwarz alternation method, I*, in Glowinski et al. [40], pp. 1–42.

[52] J. Mandel, *Balancing domain decomposition.* To appear in Communications on Applied Numerical Methods.

[53] ——, *Hybrid domain decomposition with unstructured subdomains.* To appear in the *Proceedings of the Sixth International Symposium on Domain Decomposition Methods for Partial Differential Equations*, 1993.

[54] J. MANDEL AND M. BREZINA, *Balancing domain decomposition: theory and performance in two and three dimensions.* Submitted.

[55] T. P. MATHEW, *Domain Decomposition and Iterative Refinement Methods for Mixed Finite Element Discretisations of Elliptic Problems*, PhD thesis, Courant Institute of Mathematical Sciences, Sept. 1989. Also Computer Science Tech. Report 463.

[56] ——, *Schwarz alternating and iterative refinement methods for mixed formulations of elliptic problems, part I: Algorithms and numerical results*, Numer. Math., 65 (1993), pp. 445–468.

[57] ——, *Schwarz alternating and iterative refinement methods for mixed formulations of elliptic problems, part II: Convergence theory*, Numer. Math., 65 (1993), pp. 469–492.

[58] N. D. MELSON, T. A. MANTEUFFEL, AND S. F. MCCORMICK, eds., *Sixth Copper Mountain Conference on Multigrid Methods, NASA CP-3224*, NASA, 1993.

[59] J. NECAS, *Les méthodes directes en théorie des équations elliptiques*, Academia, Prague, 1967.

[60] J. NEDELEC, *Mixed finite elements in* $\mathbb{R}^3$, Numer. Math., 35 (1980), pp. 315–341.

[61] 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, USSR, 1986.

[62] ——, *Mesh theorems on traces, normalizations of function traces and their inversion*, Sov. J. Numer. Anal. Math. Modelling, 6 (1991), pp. 223–242.

[63] T. C. OPPE, W. D. JOUBERT, AND D. R. KINCAID, *NSPCG user's guide, Version 1.0*, tech. rep., University of Texas at Austin, 1988.

[64] P. OSWALD, *On a BPX-perconditioner for P1 elements*, Tech. Rep. Forsch.-Erg. Math./91/2, FSU Jena, 1991. Submitted to Computing.

[65] ——, *On a hierarchical basis multilevel method with nonconfomring p1 elements.*, Numer. Math., 62 (1992), pp. 189–212.

[66] ——, *Preconditioners for nonconforming elements*, Tech. Rep. Math/93/3, Institut für Angewandte Mathematik, Friedrick-Schiller-Universität Jena, June 1993.

[67] J. E. PASCIAK, *Private communication.* Jan. 1993.

[68] L. PAVARINO AND M. RAMÉ, *Numerical experiments with an overalpping additive schwaz solver for 3-D parallel reservoir simulation*, Dept. Comp. Appl. Math. TR93-24, Rice University, June 1993.

[69] P. A. RAVIART AND J. M. THOMAS, *A mixed finite element method for second order elliptic problems*, in Mathematical Aspects of Finite Element Methods: Lecture Notes in Mathematics 606, I. Galligani and E. Magenes, eds., Berlin, 1977, Springer-Verlag, pp. 292–315.

[70] J. K. REID, *On the method of conjugate gradients for the solution of large sparse systems of linear equations*, in Large Sparse Sets of Linear Equations, J. K. Reid, ed., Academic Press, 1971, pp. 231–254.

[71] J. E. ROBERTRS AND J.-M. THOMAS, *Mixed and hybrid methods*, in Handbook of Numerical Analysis, Vol. II, P. G. Ciarlet and J. Lions, eds., Elsevier Science Publishers B.V., 1991, pp. 523–639.

[72] T. F. RUSSELL AND M. F. WHEELER, *Finite element and finite difference methods for continuous flows in porous media*, in Mathematics of reservoir simulation, R. E. Ewing, ed., SIAM, Philadelphia, 1983, ch. II, pp. 35–106.

[73] T. RUSTEN, *Iterative Methods for Mixed Finite Element Systems*, PhD thesis, Department of Informtics, University of Oslo, Oslo, Norway, July 1991. Tech. Rep. 156.

[74] T. RUSTEN AND R. WINTHER, *A preconditioned iterative method for saddle point problems*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 887–904.

[75] ——, *Substructure preconditioners for elliptic saddle point problems*, Math. Comp., 60 (1993), pp. 23–48.

[76] M. SARKIS, *Two-level Schwarz methods for nonconforming finite elements and discontinuous coefficients*, in Melson et al. [58], pp. 543–566.

[77] J. F. A. SLEATH, *Sea Bed Mechanics*, John Wiley & Sons, 1984.

[78] B. F. SMITH, *An optimal domain decomposition preconditioner for the finite element solution of linear elasticity problems*, SIAM J. of Sci. Stat. Comput., 13 (1992), pp. 364–378.

[79] C. SRINIVAS, *Two-dimensional modeling of flow and sedimentation*, Master's thesis, Rice University, Houston, Texas, 1992.

[80] C. SRINIVAS, B. RAMASWAMY, L. C. COWSAR, M. WHEELER, AND D. TETZLAFF, *Two-dimensional modeling of flow and sedimentation.* To appear in Proceedings of the VIII International Conference on Finite Elements in Fluids, 1993.

[81] D. M. TETZLAFF AND J. W. HARBAUGH, *Simulating Clastic Sedimentation*, Computer Methods in the Geosciences, Van Nostrand Reinhold, New York, 1989.

[82] J. M. THOMAS, *Sur l'analyse numérique des méthodes d'éléments finis hybrides et mixtes*, thèse d'etat, Université Pierre et Marie Curie, Paris, 1977.

[83] L. C. VAN RIJN, *Sediment transport, part I: Bed load transport*, Journal of Hydraulic Engineering, 110 (1984), pp. 1431–1456.

[84] ——, *Sediment transport, part II: Suspended load transport*, Journal of Hydraulic Engineering, 110 (1984), pp. 1613–1642.

[85] R. S. VARGA, *Matrix Iterative Analysis*, Prentice-Hall, Inc., 1962.

[86] P. S. VASSILEVSKI AND J. WANG, *Multilevel iterative methods for mixed finite element discretizations of elliptic problems*, Numer. Math., 63 (1992), pp. 503–520.

[87] R. VERFÜRTH, *A multilevel algorithm for mixed problems*, SIAM J. Numer. Anal., 21 (1984), pp. 264–271.

[88] ——, *Multilevel algorithms for mixed problems. II. Treatment of the mini-element*, SIAM J. Numer. Anal., 25 (1988), pp. 285–293.

[89] 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., GAMM, 1987, pp. 110–122.

[90] ——, *Iterative substructuring methods: Algorithms and theory for problems in the plane*, in Glowinski et al. [40], pp. 113–128.

[91] G. WITTUM, *Linear iterations as smoothers in multigrid methods: theory with applications to incomplete decompositions*, Impact of Computing in Sci. Eng., 1 (1989), pp. 180–215.

[92] B. WOHLMUTH AND R. H. HOPPE, *Multilevel approaches to nonconforming finite element discretizations of linear second order elliptic boundary value problems*, Tech. Rep. TUM-M9320, Technische Universität München, Oct. 1993.

[93] J. XU, *Iterative methods by space decomposition and subspace correction*, SIAM Review, 34 (1992), pp. 581–613.