PRECONDITIONING FOR THE MIXED FORMULATION OF LINEAR PLANE ELASTICITY

A Dissertation

by

YANQIU WANG

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2004

Major Subject: Mathematics

PRECONDITIONING FOR THE MIXED FORMULATION OF LINEAR PLANE ELASTICITY

A Dissertation

by

YANQIU WANG

Submitted to Texas A&M University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Approved as to style and content by:

Joseph E. Pasciak (Chair of Committee)	
James H. Bramble (Member)	Raytcho Lazarov (Member)
Vivek Sarin (Member)	Albert Boggess (Head of Department)

August 2004

Major Subject: Mathematics

ABSTRACT

Preconditioning for the Mixed Formulation of Linear Plane Elasticity. (August 2004)

Yanqiu Wang, B.S., Fudan University, China;

M.S., Fudan University, China

Chair of Advisory Committee: Dr. Joseph E. Pasciak

In this dissertation, we study the mixed finite element method for the linear plane elasticity problem and iterative solvers for the resulting discrete system. We use the Arnold-Winther Element in the mixed finite element discretization. An overlapping Schwarz preconditioner and a multigrid preconditioner for the discrete system are developed and analyzed.

We start by introducing the mixed formulation (stress-displacement formulation) for the linear plane elasticity problem and its discretization. A detailed analysis of the Arnold-Winther Element is given. The finite element discretization of the mixed formulation leads to a symmetric indefinite linear system.

Next, we study efficient iterative solvers for the symmetric indefinite linear system which arises from the mixed finite element discretization of the linear plane elasticity problem. The preconditioned Minimum Residual Method is considered. It is shown that the problem of constructing a preconditioner for the indefinite linear system can be reduced to the problem of constructing a preconditioner for the $\boldsymbol{H}(\mathbf{div})$ problem in the Arnold-Winther finite element space. Our main work involves developing an overlapping Schwarz preconditioner and a multigrid preconditioner for the $\boldsymbol{H}(\mathbf{div})$ problem. We give condition number estimates for the preconditioned systems together with supporting numerical results.

To my parents

ACKNOWLEDGMENTS

I am very grateful to my advisor, Dr. Joseph E. Pasciak, for his guidance, support, patience, and encouragement throughout my graduate studies. His technical and editorial advice was essential to the completion of this dissertation. I thank him for teaching me innumerable lessons and insights on academic research.

I express my gratitude to all the other members of my advisory committee for their consistent support and valuable comments on my research. I thank Dr. James H. Bramble for his inspiring lectures on finite element methods and preconditioning skills and for his valuable ideas. I thank Dr. Raytcho Lazarov for his valuable advice and constant encouragement. I also thank Dr. Vivek Sarin for his willingness to serve as a member of my advisory committee and for his interest in the subject of my research.

I thank Dr. Yalchin Efendiev for sharing with me many interesting ideas and insights. I thank my officemates, Taejong Kim and Tzanio Kolev, for innumerable useful discussions. I thank all people in the numerical analysis group for the wonderful academic atmosphere.

I thank Dr. Jay Walton, Dr. Thomas Schlumprecht, Dr. Michael Pilant and Ms. Monique Stewart and other members in the Department of Mathematics and the Institute of Scientific Computation at Texas A&M University for all the help they offered during my graduate study.

Finally, my parents receive my deepest love for their love, encouragement and support.

TABLE OF CONTENTS

CONTENTS Page

I INTRODUCTION 1

II PRELIMINARIES 5

- A. Sobolev spaces 5
- B. The linearized theory of elasticity 14
- C. Variational principles 18
- D. The Airy operator 24

III MIXED FINITE ELEMENT DISCRETIZATION 31

- A. Mixed finite element method 31
- B. Arnold-Winther elements 34
- C. Airy operator on the discrete level 47
- D. Other mixed elements 53

IV ITERATIVE SOLVERS FOR SADDLE POINT PROBLEMS . 57

- A. The preconditioned MINRES method 57
- B. Preconditioning the saddle-point problem using norm equivalence 62
- C. Other iterative solvers 67

V THE OVERLAPPING SCHWARZ PRECONDITIONER 73

- A. Framework of an overlapping Schwarz preconditioner . . . 73
- B. An overlapping Schwarz preconditioner for the $\boldsymbol{H}(\mathbf{div})$ problem 76
- C. The condition number estimate 83

VI THE MULTIGRID PRECONDITIONER 91

- A. Framework of a multigrid preconditioner 91
- B. A multigrid preconditioner for the $\boldsymbol{H}(\mathbf{div})$ problem 93
- C. The condition number estimate 97

VII NUMERICAL RESULTS 103

VIII CONCLUSIONS 110

Page

REFERENCES 112
APPENDIX A 118
VITA 122

LIST OF TABLES

TABLE		Page
1	Approximation behavior of the Arnold-Winther element.	104
2	Condition number estimates with $H = 1/2$, $\delta = 1/8$ and $K = 4$.	106
3	Condition number estimates for Problem (4.11) with $H = 1/4$ and $\delta = 1/8$.	107
4	Condition number estimates for Problem (4.11) with $K = 4$ and $H = 1/2$.	108
5	Condition number estimates for $\mathbf{\Lambda}_k$, $\mathbf{B}_k^R\mathbf{\Lambda}_k$, $\mathbf{B}_k\mathbf{\Lambda}_k$ and $\mathbf{B}_k^m\mathbf{\Lambda}_k$.	109
6	Condition number estimates for $\mathbf{B}_k^V\mathbf{\Lambda}_k$.	109

LIST OF FIGURES

FIGURE Page

1 Typical plane strain model and plane stress model. 17

2 Finite elements $\boldsymbol{\Sigma}_T$ and $\boldsymbol{V}_T$. 35

3 Clusters of triangles. 37

4 An example of T and $\boldsymbol{x} \in V_c$. 42

5 The Argyris finite element. 48

6 The Johnson-Mercier finite element. 54

7 The Arnold-Douglas-Gupta finite element with $k = 2$. 55

8 Sub-domains $\tilde{\Omega}_i$ and Ω_i. 77

9 The coarse mesh and sub-domains with $H = 1/2$, $\delta = 1/8$ and $K = 4$. 106

CHAPTER I

INTRODUCTION

The purpose of this thesis is to study the mixed formulation (stress-displacement formulation) of the planar linear elasticity problem and to develop preconditioners for the resulting linear system.

Mixed finite element methods [2, 22] have been widely used in solving partial differential equations. Compared to the primal-based methods, mixed finite element methods have some well-known advantages. For example, the dual variable, which is usually the variable of primary interest, is computed directly as a fundamental unknown. Mixed methods also have some obvious disadvantages, such as the necessity of constructing stable pairs of finite element spaces and the fact that the resulting discrete system is indefinite. Therefore, the construction of stable pairs of finite element spaces and the development of efficient iterative solvers for the resulting discrete system remain two of the most important issues in the applications of mixed finite element methods.

For decades, extensive research has been carried out to explore the mixed formulation of the plane elasticity problem (also known as the weak formulation of the Hellinger-Reissner Principle). Most of this research focused on developing stable pairs of mixed finite element spaces and several different solutions have been proposed [3, 4, 8, 44]. As stated in those papers, the crux of the difficulty is that the stress tensor in the Hellinger-Reissner Principle has to be symmetric. Indeed, this symmetry condition is so hard to satisfy that the authors of [3, 4, 44] resorted to composite elements. Only recently did Arnold and Winther construct a new stable

---

This dissertation follows the style and format of *Mathematics of Computation*.

pair of mixed finite elements [8] which does not use composite elements. We will base our research on the lowest order Arnold-Winther finite element.

We mention some alternative ways to circumvent the difficulty of constructing stable pairs of finite elements. One way is to reformulate the saddle-point problem by using Lagrangian functionals so that it does not require symmetric tensors [3, 5]. Another way is to use the least-square formulation so that the classical discrete inf-sup condition is no longer needed [14, 23, 24]. Also, other authors resort to the use of stabilizing techniques (see [34] and the references therein).

The Arnold-Winther finite element spaces consist of piecewise polynomials. It has been proved in [8] that the Arnold-Winther finite element spaces are stable for the pure displacement boundary problem. In Chapter III, we will generalize their proof for stability to problems with more general boundary conditions under certain regularity assumptions.

The discretization of the mixed formulation leads to a symmetric indefinite linear system. Generally speaking, there are three main approaches for solving large symmetric indefinite linear systems corresponding to mixed formulations. One can use the well-studied Uzawa-type method [10, 31, 17, 12, 55]. The second choice is the positive definite reformulation proposed by Bramble and Pasciak in [15] and [16]. The third choice is the preconditioned minimum residual method analyzed in [6, 49]. We adopt the idea of the preconditioned minimum residual method. An analysis similar to the one in [6] will show that the problem of constructing a preconditioner for the indefinite linear system derived from the mixed formulation of linear plane elasticity can be reduced to the problem of constructing a preconditioner for the $\boldsymbol{H}(\mathbf{div})$ problem on the Arnold-Winther finite element space on the symmetric tensor field.

We consider an overlapping Schwarz preconditioner and a multigrid preconditioner. Overlapping Schwarz methods provide efficient preconditioners for second order elliptic problems. Two-level additive Schwarz methods were first introduced by Dryja and Widlund [27, 28]. Multiplicative Schwarz methods and a general framework for the overlapping Schwarz methods were developed by Bramble, Pasciak, Wang and Xu [18]. Later, Dryja and Widlund showed that two-level Schwarz methods work well even in the case of small overlapping of sub-domains [29]. For a systematic analysis of the overlapping Schwarz method, see also [52, 53, 54]. The application of the overlapping Schwarz method to vector problems on $H(\mathrm{div})$ was discussed in [43, 6].

Multigrid methods provide another type of efficient preconditioner. A vast amount of research has been done in this area, among which we refer to survey papers [20, 13, 53], the book by Hackbusch [41], the book by Bramble and Zhang [19] and the references therein. In [53], it was pointed out that the multigrid algorithm and the multiplicative Schwarz algorithm can be considered in the general framework of the so-called *Successive Subspace Correction Methods*. Application of the multigrid method to vector problems on $H(\mathrm{div})$ was studied in [42, 6, 7].

Our main work involves developing an overlapping Schwarz preconditioner and a multigrid preconditioner for the $\boldsymbol{H}(\mathbf{div})$ problem on the Arnold-Winther finite element space. The discrete operator which results from the $\boldsymbol{H}(\mathbf{div})$ problem is not uniformly elliptic. This causes the main difficulty in the development and the analysis of our preconditioners since the classical techniques require the operator to have some elliptic regularity. To deal with this difficulty, we follow the idea of using a Helmholtz-like decomposition [6, 32] and decompose the Arnold-Winther finite element space into two orthogonal subspaces: the subspace of divergence free functions and its orthogonal compliment. Then, the analysis of our preconditioners can be done on these two subspaces separately.

Our results state that, for polygonal domains and the pure displacement boundary problem or the pure traction boundary problem, the condition number of the preconditioned system using the overlapping Schwarz preconditioner is uniform with respect to the mesh size and the number of sub-domains; for convex polygonal domains and the pure displacement boundary problem or the pure traction boundary problem, the condition number of the preconditioned system using the variable V-cycle multigrid preconditioner is independent of the number of levels.

Finally, we give an outline of this thesis. In Chapter II, we introduce the basics about Sobolev spaces, the linear plane elasticity problem together with its variational formulations, and the Airy operator which connects the plane elasticity problem to the biharmonic problem. Several important issues like the existence and uniqueness of the weak problem, the regularity of the weak solution, a decomposition of the $\boldsymbol{H}(\mathbf{div}\,, \Omega, \mathbb{S}_2)$ space and the related exact sequence are discussed in this chapter. In Chapter III, we introduce the mixed finite element discretization of plane elasticity and the Arnold-Winther elements. The Arnold-Winther element is related to the Argyris element by the Airy operator. In this chapter, we also briefly introduce several other finite elements for the mixed formulation. In Chapter IV, we discuss iterative solvers for the linear system which results from the saddle-point problem. The preconditioned minimum residual algorithm is given there. We show that the problem of finding a preconditioner for the saddle-point problem can be reduced to the problem of finding a preconditioner for the $\boldsymbol{H}(\mathbf{div})$ problem. Also in this section, we briefly introduce several other iterative solvers for the saddle-point problem. We develop an overlapping Schwarz preconditioner for the $\boldsymbol{H}(\mathbf{div})$ problem in Chapter V and a multigrid preconditioner for the $\boldsymbol{H}(\mathbf{div})$ problem in Chapter VI. The condition numbers of the resulting preconditioned systems are analyzed there. Finally, in Chapter VII, we give the results of numerical experiments.

CHAPTER II

PRELIMINARIES

In this chapter, we introduce the basics about Sobolev space, the linear elasticity problem together with its variational formulations, and the Airy operator. We start with some definitions and properties of Sobolev spaces, especially the symmetric matrix space $\boldsymbol{H}(\mathbf{div}\,, \Omega, \mathbb{S}_2)$ and its properties on the boundary $\partial\Omega$. Then we describe the model problem for planar linear elasticity and discuss its variational formulations. The existence, uniqueness and the regularity results for the weak solution are stated. Finally, we introduce the Airy operator, which connects the divergence free part of $\boldsymbol{H}(\mathbf{div}\,, \Omega, \mathbb{S}_2)$ to the biharmonic problem and yields an orthogonal decomposition of the space $\boldsymbol{H}(\mathbf{div}\,, \Omega, \mathbb{S}_2)$. The most interesting property of this decomposition is that the orthogonal complement of the divergence free part gains $\boldsymbol{H}^s$-regularity. This decomposition will be the key factor for the analysis in Chapter V and Chapter VI.

A. Sobolev spaces

Let Ω be a bounded, open subset of $\mathbb{R}^n$ with a Lipschitz continuous boundary and denote $\partial\Omega$ to be the boundary of Ω. Let $(x_1, \ldots, x_n)$ be a coordinate system for $\mathbb{R}^n$. We restrict our attentions on real-valued functions over Ω. A *scalar function* u on Ω is defined as a mapping from Ω to $\mathbb{R}$. Let $\alpha = (\alpha_1, \ldots, \alpha_n)$ be a multi-index, where α_i, $i = 1, \ldots, n$, are nonnegative integers. The length of α is defined by $|\alpha| = \sum_{i=1}^n \alpha_i$. Denote

$$D^\alpha u = \frac{\partial^{|\alpha|} u}{\partial^{\alpha_1} x_1 \cdots \partial^{\alpha_n} x_n}.$$

Here D^α is considered in the weak sense.

We review some basic concepts of Sobolev spaces. Let $C^k(\Omega)$ be the space of kth order continuously differentiable functions on Ω and $C^\infty(\Omega)$ be the space of infinitely differentiable functions on Ω. Define $C^k(\bar{\Omega})$ to be the space of functions which are kth order differentiable and continuous up to the boundary $\partial\Omega$. Let $C_0^\infty(\Omega)$ be the subset of $C^\infty(\Omega)$ such that every $u \in C_0^\infty(\Omega)$ has compact support in Ω. For $1 \leq p \leq \infty$, let $L^p(\Omega)$ be the Lebesgue space defined on Ω with the norm

$$\|u\|_{L^p(\Omega)} = \begin{cases} \left(\int_\Omega |u|^p \, d\boldsymbol{x}\right)^{1/p}, & 1 \leq p < \infty, \\ \left(\operatorname*{ess\,sup}_{\boldsymbol{x}\in\Omega} |u(\boldsymbol{x})|\right), & p = \infty. \end{cases}$$

Let $W^{s,p}(\Omega)$ be the Sobolev space for $s \geq 0$ and $1 \leq p \leq \infty$ with the norm and the semi-norm defined by:

1. if $s = m$ is an integer:

$$\|u\|_{W^{m,p}(\Omega)} = \begin{cases} \left(\sum_{|\alpha|\leq m} \|D^\alpha u\|^p_{L^p(\Omega)}\right)^{1/p}, & 1 \leq p < \infty, \\ \max_{|\alpha|\leq m} \left(\|D^\alpha u\|_{L^\infty(\Omega)}\right), & p = \infty, \end{cases}$$

$$|u|_{W^{m,p}(\Omega)} = \begin{cases} \left(\sum_{|\alpha|= m} \|D^\alpha u\|^p_{L^p(\Omega)}\right)^{1/p}, & 1 \leq p < \infty, \\ \max_{|\alpha|= m} \left(\|D^\alpha u\|_{L^\infty(\Omega)}\right), & p = \infty, \end{cases}$$

2. if $s = m + t$ where m is an integer and $0 < t < 1$:

$$\|u\|_{W^{s,p}(\Omega)} = \begin{cases} \left(\|u\|^p_{W^{m,p}(\Omega)} + \sum_{|\alpha|=m} \int_\Omega \int_\Omega \frac{|D^\alpha u(\boldsymbol{x}) - D^\alpha u(\boldsymbol{y})|^p}{|\boldsymbol{x}-\boldsymbol{y}|^{n+tp}} \, d\boldsymbol{x} \, d\boldsymbol{y} \right)^{1/p}, \\ \qquad\qquad \text{when } 1 \le p < \infty, \\ \max \left(\|u\|_{W^{m,\infty}(\Omega)}, \max_{|\alpha|=m} \operatorname*{ess\,sup}_{\substack{\boldsymbol{x},\boldsymbol{y}\in\Omega \\ \boldsymbol{x}\neq\boldsymbol{y}}} \frac{|D^\alpha u(\boldsymbol{x}) - D^\alpha u(\boldsymbol{y})|}{|\boldsymbol{x}-\boldsymbol{y}|^t} \right), \\ \qquad\qquad \text{when } p = \infty, \end{cases}$$

$$|u|_{W^{s,p}(\Omega)} = \begin{cases} \left(\sum_{|\alpha|=m} \int_\Omega \int_\Omega \frac{|D^\alpha u(\boldsymbol{x}) - D^\alpha u(\boldsymbol{y})|^p}{|\boldsymbol{x}-\boldsymbol{y}|^{n+tp}} \, d\boldsymbol{x} \, d\boldsymbol{y} \right)^{1/p}, \\ \qquad\qquad \text{when } 1 \le p < \infty, \\ \max_{|\alpha|=m} \left(\operatorname*{ess\,sup}_{\substack{\boldsymbol{x},\boldsymbol{y}\in\Omega \\ \boldsymbol{x}\neq\boldsymbol{y}}} \frac{|D^\alpha u(\boldsymbol{x}) - D^\alpha u(\boldsymbol{y})|}{|\boldsymbol{x}-\boldsymbol{y}|^t} \right), \\ \qquad\qquad \text{when } p = \infty. \end{cases}$$

For $u \in W^{s,p}(\Omega)$, let $\tilde{u}$ be the extension of u to $\mathbb{R}^n$ by zero outside Ω. Define

$$\tilde{W}^{s,p}(\Omega) = \{u \in W^{s,p}(\Omega) \text{ such that } \tilde{u} \in W^{s,p}(\mathbb{R}^n)\}, \tag{2.1}$$

where $W^{s,p}(\mathbb{R}^n)$ is the Sobolev space defined on $\mathbb{R}^n$. Define

$$W_0^{s,p}(\Omega) = \overline{C_0^\infty(\Omega)}^{W^{s,p}(\Omega)},$$

which means the closure of $C_0^\infty(\Omega)$ under $W^{s,p}(\Omega)$ norm.

Let $< \cdot, \cdot >$ be the duality pairing. For $s < 0$ and $1 < p < \infty$, we define the Sobolev space $W^{s,p}(\Omega)$ to be the dual space of $W_0^{-s,q}(\Omega)$, where $\frac{1}{p} + \frac{1}{q} = 1$. The norm on $W^{s,p}(\Omega)$, $s < 0$, is defined by

$$\|u\|_{W^{s,p}(\Omega)} = \sup_{\substack{v \in W_0^{-s,q}(\Omega) \\ v \neq 0}} \frac{< u, v >}{\|v\|_{W^{-s,q}(\Omega)}}.$$

$W^{s,2}(\Omega)$ is a Hilbert space and is commonly denoted by $H^s(\Omega)$. Similarly, $W_0^{s,2}(\Omega)$ is denoted by $H_0^s(\Omega)$.

For properties of Sobolev spaces, see [47, 45, 37, 35, 1]. We only list several important results in the following.

Theorem II.1. *When $s - 1/p$ is not an integer, we have*

$$\tilde{W}^{s,p}(\Omega) = W_0^{s,p}(\Omega).$$

Furthermore, when $0 \leq s < 1/p$, we have

$$\tilde{W}^{s,p}(\Omega) = W_0^{s,p}(\Omega) = W^{s,p}(\Omega).$$

Let Γ be an open subset of $\partial\Omega$, which is Lipschitz continuous. We can define a unit outward normal vector field $\boldsymbol{n}$ almost everywhere on Γ and $\boldsymbol{n} \in (L^\infty(\Gamma))^2$. We say that Γ is of class $C^{k,1}$ if for every $\boldsymbol{x} \in \Gamma$, there exists a neighborhood V of $\boldsymbol{x}$ in $\mathbb{R}^n$ and a map $\phi : \mathbb{R}^{n-1} \to V \cap \Gamma$ which is k times continuously differentiable and its derivatives of order k are Lipschitz continuous. If Γ is of class $C^{k,1}$ for $k \geq 1$, then the unit outward normal vector $\boldsymbol{n}$ is of class $C^{k-1,1}$.

For $u \in C^\infty(\bar{\Omega})$, define $\Gamma u = u|_\Gamma$.

Theorem II.2. *(Trace theorem) If Γ is of class $C^{k,1}$, then the mapping*

$$u \to \{\Gamma u, \Gamma\frac{\partial u}{\partial \boldsymbol{n}}, \ldots, \Gamma\frac{\partial^m u}{\partial \boldsymbol{n}^m}\}$$

which is defined for $u \in C^\infty(\bar{\Omega})$ has a unique continuous extension as an operator *from*

$$W^{s,p}(\Omega) \ \textit{onto} \ \prod_{i=0}^{m} W^{s-i-1/p,p}(\Gamma) \qquad \textit{for } m+1/p < s \le k+1.$$

Corollary II.1. *We have*

$$H_0^1(\Omega) = \{u \in H^1(\Omega) \ \textit{such that} \ u|_{\partial\Omega} = 0\},$$

$$H_0^2(\Omega) = \{u \in H^2(\Omega) \ \textit{such that} \ u|_{\partial\Omega} = 0, \ (D^\alpha u)|_{\partial\Omega} = 0, \ \textit{for} \ |\alpha| = 1\}.$$

We also define the following spaces:

$$H_{0,\Gamma}^1(\Omega) = \{u \in H^1(\Omega) \text{ such that } u|_\Gamma = 0\},$$

$$H_{0,\Gamma}^2(\Omega) = \{u \in H^2(\Omega) \text{ such that } u|_\Gamma = 0, \ (D^\alpha u)|_\Gamma = 0, \ \text{for } |\alpha| = 1\}.$$

Remark II.1. *If Γ is of class $C^{1,1}$, then*

$$H_{0,\Gamma}^2(\Omega) = \{u \in H^2(\Omega) \ \textit{such that} \ u|_\Gamma = 0, \ \frac{\partial u}{\partial \boldsymbol{n}}|_\Gamma = 0\}.$$

However, it is possible to extend slightly the above statement for the case when Ω is a bounded two-dimensional Lipschitz polygon. Notice that a polygon is never of class $C^{1,1}$. Let Γ_i, $1 \le i \le N$, be the boundary edges of Ω. Then $u \to \{\Gamma_1 \frac{\partial u}{\partial \boldsymbol{n}}, \ldots, \Gamma_N \frac{\partial u}{\partial \boldsymbol{n}}\}$ is a linear, continuous operator from $H^2(\Omega)$ onto $\prod_{i=1}^N H^{1/2}(\Gamma_i)$. Furthermore, we have

$$H_{0,\Gamma}^2(\Omega) = \{u \in H^2(\Omega) \ \textit{such that} \ u|_\Gamma = 0, \ \frac{\partial u}{\partial \boldsymbol{n}}|_{\Gamma_i \cap \Gamma} = 0 \ \textit{for} \ 1 \le i \le N\}.$$

Similar to the definition (2.1) for $\tilde{W}^{s,p}(\Omega)$, we define the space $\tilde{W}^{s,p}(\Gamma)$ on an open set $\Gamma \subset \partial\Omega$ by

$$\tilde{W}^{s,p}(\Gamma) = \{u \in W^{s,p}(\Gamma) \text{ such that } \tilde{u} \in W^{s,p}(\partial\Omega)\},$$

where $\tilde{u}$ is the extension of u to $\partial\Omega$ by zero outside Γ. For $u \in H_{0,\Gamma}^1(\Omega)$, notice that $u|_{\partial\Omega} \in H^{1/2}(\partial\Omega)$ and $u|_\Gamma = 0$. Clearly, we have

$$u|_{\partial\Omega\backslash\Gamma} \in \tilde{W}^{1/2,2}(\partial\Omega\backslash\Gamma).$$

Normally, when $n = 2$ and Γ is a 1-D curve, we denote

$$H_{00}^{1/2}(\Gamma) = \tilde{W}^{1/2,2}(\Gamma)$$

with the norm (if Γ is simply connected)

$$\|u\|^2_{H_{00}^{1/2}(\Gamma)} = \|u\|^2_{H^{1/2}(\Gamma)} + \int_\Gamma (\frac{u^2}{s} + \frac{u^2}{|\Gamma| - s})\, ds.$$

The above definition for the $H_{00}^{1/2}(\Gamma)$ norm can easily be generalized to the case when Γ is multiply connected. We have

$$H_{00}^{1/2}(\Gamma) \subsetneq H_0^{1/2}(\Gamma) = H^{1/2}(\Gamma).$$

Another obvious result is that, for each $u \in H_{00}^{1/2}(\partial\Omega\backslash\Gamma)$, there exists a $v \in H^1_{0,\Gamma}(\Omega)$ such that $v|_{\partial\Omega\backslash\Gamma} = u$.

Next, we generalize the above definitions of Sobolev spaces to the cases of vector functions and symmetric matrix functions. We will focus on two-dimensional problems. Therefore, in the rest of this section, Ω is a bounded, open subset in $\mathbb{R}^2$ with a Lipschitz continuous boundary and the coordinate system is set to be (x, y).

Let $\mathbb{R}^2$ be the space of 2-dimensional real vectors and $\mathbb{S}_2$ be the space of symmetric 2×2 real matrices. Define the inner product between vectors and the inner product between matrices by:

$$\begin{aligned}
\boldsymbol{u} \cdot \boldsymbol{v} &= \sum_{i=1}^{2} u_i v_i, && \text{for } \boldsymbol{u} = (u_i)_{1 \le i \le 2},\ \boldsymbol{v} = (v_i)_{1 \le i \le 2} \ \in \mathbb{R}^2,\\
\boldsymbol{\sigma} : \boldsymbol{\tau} &= \sum_{i,j=1}^{2} \sigma_{ij} \tau_{ij}, && \text{for } \boldsymbol{\sigma} = (\sigma_{ij})_{1 \le i,j \le 2},\ \boldsymbol{\tau} = (\tau_{ij})_{1 \le i,j \le 2} \ \in \mathbb{S}_2.
\end{aligned}$$

Define a 2-dimensional *vector function* $\boldsymbol{v} = (v_i)_{1 \le i \le 2}$ to be a mapping from Ω to $\mathbb{R}^2$ and a 2×2 *symmetric matrix function* $\boldsymbol{\tau} = (\tau_{ij})_{1 \le i,j \le 2}$ to be a mapping from Ω to $\mathbb{S}_2$. We adopt the convention that a Latin character in lower case denotes a scalar or a scalar function, a bold Latin character in lower case denotes a vector or a vector function and a bold Greek character denotes a matrix or a matrix function.

Define

$$W^{s,p}(\Omega,\mathbb{R}^2) = (W^{s,p}(\Omega))^2,$$

$$W^{s,p}(\Omega,\mathbb{S}_2) = (W^{s,p}(\Omega))^3,$$

with norms

$$\|\boldsymbol{v}\|_{W^{s,p}(\Omega,\mathbb{R}^2)} = (\|v_1\|^p_{W^{s,p}(\Omega)} + \|v_2\|^p_{W^{s,p}(\Omega)})^{1/p},$$

$$\|\boldsymbol{\tau}\|_{W^{s,p}(\Omega,\mathbb{S}_2)} = (\|\tau_{11}\|^p_{W^{s,p}(\Omega)} + 2\|\tau_{12}\|^p_{W^{s,p}(\Omega)} + \|\tau_{22}\|^p_{W^{s,p}(\Omega)})^{1/p}.$$

We generalize notations for other spaces in the same fashion.

For simplicity, denote $\|\cdot\|_{s,\Omega}$ to be the H^s-norm over scalar, vector or matrix fields, depending on the type of the function. We also use the notation $(\cdot,\cdot)$ for the L^2-inner product over scalar, vector or matrix fields defined on the whole domain Ω.

Define the gradient of a scalar function u and the gradient of a vector function $\boldsymbol{v}$, respectively, by

$$\nabla u = \begin{pmatrix} \frac{\partial}{\partial x} u \\ \frac{\partial}{\partial y} u \end{pmatrix} \quad \text{and} \quad \nabla \boldsymbol{v} = \begin{pmatrix} \frac{\partial}{\partial x} v_1 & \frac{\partial}{\partial y} v_1 \\ \frac{\partial}{\partial x} v_2 & \frac{\partial}{\partial y} v_2 \end{pmatrix}.$$

Also, we define the divergence of a vector function $\boldsymbol{v}$ and the divergence of a matrix function $\boldsymbol{\tau}$, respectively, by

$$\operatorname{div} \boldsymbol{v} = \frac{\partial}{\partial x} v_1 + \frac{\partial}{\partial y} v_2 \quad \text{and} \quad \mathbf{div}\, \boldsymbol{\tau} = \begin{pmatrix} \frac{\partial}{\partial x} \tau_{11} + \frac{\partial}{\partial y} \tau_{12} \\ \frac{\partial}{\partial x} \tau_{21} + \frac{\partial}{\partial y} \tau_{22} \end{pmatrix}.$$

Define

$$\boldsymbol{H}(\mathbf{div}\,, \Omega, \mathbb{S}_2) = \{\boldsymbol{\tau} \in L^2(\Omega, \mathbb{S}_2) \text{ such that } \mathbf{div}\, \boldsymbol{\tau} \in \boldsymbol{L}^2(\Omega, \mathbb{R}^2)\},$$

with the norm

$$\|\boldsymbol{\tau}\|^2_{\boldsymbol{H}(\mathbf{div}\,,\Omega,\mathbb{S}_2)} = \|\boldsymbol{\tau}\|^2_{0,\Omega} + \|\mathbf{div}\,\boldsymbol{\tau}\|^2_{0,\Omega}.$$

$\boldsymbol{H}(\mathbf{div}\,,\Omega,\mathbb{S}_2)$ is a Hilbert space with the inner product

$$(\boldsymbol{\sigma},\boldsymbol{\tau})_{\boldsymbol{H}(\mathbf{div}\,,\Omega,\mathbb{S}_2)} = (\boldsymbol{\sigma},\boldsymbol{\tau}) + (\mathbf{div}\,\boldsymbol{\sigma},\mathbf{div}\,\boldsymbol{\tau}).$$

Similar to Theorem 2.4 and Theorem 2.5 in [35], we have the following results:

Lemma II.1. *$C^\infty(\Omega,\mathbb{S}_2)$ is dense in $\boldsymbol{H}(\mathbf{div}\,,\Omega,\mathbb{S}_2)$*

Lemma II.2. *For $\boldsymbol{\tau}\in\boldsymbol{H}(\mathbf{div}\,,\Omega,\mathbb{S}_2)$, we have $\boldsymbol{\tau n}|_{\partial\Omega}\in\boldsymbol{H}^{-1/2}(\partial\Omega,\mathbb{R}^2)$, where $\boldsymbol{n}$ is the unit outward normal vector on $\partial\Omega$. Furthermore, the following Green's formula is true:*

$$(\mathbf{div}\,\boldsymbol{\tau},\boldsymbol{v}) + (\boldsymbol{\tau},\nabla\boldsymbol{v}) = <\boldsymbol{\tau n},\boldsymbol{v}>_{\partial\Omega} \quad \textit{for all } \boldsymbol{v}\in\boldsymbol{H}^1(\Omega,\mathbb{R}^2).$$

$\boldsymbol{\tau n}|_{\partial\Omega}$ is called the *normal component* of $\boldsymbol{\tau}$ on $\partial\Omega$. Let Γ be an arbitrary open subset of $\partial\Omega$. For each $\boldsymbol{v}\in\boldsymbol{H}^{1/2}_{00}(\Gamma,\mathbb{R}^2)$, let $\tilde{\boldsymbol{v}}$ be the extension of $\boldsymbol{v}$ to $\partial\Omega$ by zero outside Γ. It is not hard to see that $\|\boldsymbol{v}\|_{\boldsymbol{H}^{1/2}_{00}(\Gamma,\mathbb{R}^2)}$ is equivalent to $\|\tilde{\boldsymbol{v}}\|_{\boldsymbol{H}^{1/2}(\partial\Omega,\mathbb{R}^2)}$ up to a constant independent of $\boldsymbol{v}$. Define

$$<\boldsymbol{\tau n},\boldsymbol{v}>_\Gamma = <\boldsymbol{\tau n},\tilde{\boldsymbol{v}}>_{\partial\Omega}. \tag{2.2}$$

Clearly, $<\boldsymbol{\tau n},\cdot>_\Gamma$ is a well defined functional on $\boldsymbol{H}^{1/2}_{00}(\Gamma,\mathbb{R}^2)$ since

$$\begin{aligned} <\boldsymbol{\tau n},\boldsymbol{v}>_\Gamma &= <\boldsymbol{\tau n},\tilde{\boldsymbol{v}}>_{\partial\Omega} \le \|\boldsymbol{\tau n}\|_{\boldsymbol{H}^{-1/2}(\partial\Omega,\mathbb{R}^2)}\|\tilde{\boldsymbol{v}}\|_{\boldsymbol{H}^{1/2}(\partial\Omega,\mathbb{R}^2)} \\ &\le c\|\boldsymbol{\tau n}\|_{\boldsymbol{H}^{-1/2}(\partial\Omega,\mathbb{R}^2)}\|\boldsymbol{v}\|_{\boldsymbol{H}^{1/2}_{00}(\Gamma,\mathbb{R}^2)}, \end{aligned}$$

where c is a positive constant independent of $\boldsymbol{v}$. Therefore $\boldsymbol{\tau n}$ is in $(\boldsymbol{H}^{1/2}_{00}(\Gamma,\mathbb{R}^2))^*$, the dual space of $\boldsymbol{H}^{1/2}_{00}(\Gamma,\mathbb{R}^2)$. Note that $\boldsymbol{\tau n}$ is not necessarily in $\boldsymbol{H}^{-1/2}(\Gamma,\mathbb{R}^2)$.

Define

$$\boldsymbol{H}_{0,\Gamma}(\mathbf{div}\,,\Omega,\mathbb{S}_2) = \{\boldsymbol{\tau} \in \boldsymbol{H}(\mathbf{div}\,,\Omega,\mathbb{S}_2) \text{ such that}$$
$$< \boldsymbol{\tau n}, \boldsymbol{v} >_\Gamma = 0 \text{ for all } \boldsymbol{v} \in \boldsymbol{H}^{1/2}_{00}(\Gamma,\mathbb{R}^2)\},$$
$$\boldsymbol{H}_0(\mathbf{div}\,,\Omega,\mathbb{S}_2) = \boldsymbol{H}_{0,\partial\Omega}(\mathbf{div}\,,\Omega,\mathbb{S}_2).$$

Notice that if $\boldsymbol{\tau n} \in \boldsymbol{L}^2(\partial\Omega,\mathbb{R}^2)$, then the boundary condition of $\boldsymbol{H}_{0,\Gamma}(\mathbf{div}\,,\Omega,\mathbb{S}_2)$ is equivalent to $\boldsymbol{\tau n}|_\Gamma = \mathbf{0}$.

Remark II.2. *Clearly, an equivalent way to define* $\boldsymbol{H}_{0,\Gamma}(\mathbf{div}\,,\Omega,\mathbb{S}_2)$ *is*

$$\boldsymbol{H}_{0,\Gamma}(\mathbf{div}\,,\Omega,\mathbb{S}_2) = \{\boldsymbol{\tau} \in \boldsymbol{H}(\mathbf{div}\,,\Omega,\mathbb{S}_2) \textit{ such that}$$
$$< \boldsymbol{\tau n}, \boldsymbol{v} >_{\partial\Omega} = 0 \textit{ for all } \boldsymbol{v} \in \boldsymbol{H}^1_{0,\partial\Omega\backslash\Gamma}(\Omega,\mathbb{R}^2)\},$$

which is used in [22].

For $\boldsymbol{v} \in \boldsymbol{H}^{1/2}(\partial\Omega\backslash\Gamma,\mathbb{R}_2)$, there exists $\boldsymbol{w} \in \boldsymbol{H}^{1/2}(\partial\Omega,\mathbb{R}^2)$ such that $\boldsymbol{w}|_{\partial\Omega\backslash\Gamma} = \boldsymbol{v}$. For $\boldsymbol{\tau} \in \boldsymbol{H}_{0,\Gamma}(\mathbf{div}\,,\Omega,\mathbb{S}_2)$, we formally define

$$< \boldsymbol{\tau n}, \boldsymbol{v} >_{\partial\Omega\backslash\Gamma} = < \boldsymbol{\tau n}, \boldsymbol{w} >_{\partial\Omega} . \tag{2.3}$$

We claim that $< \boldsymbol{\tau n}, \boldsymbol{v} >_{\partial\Omega\backslash\Gamma}$ is uniquely defined independent of the choice of $\boldsymbol{w}$. Indeed, if $\boldsymbol{w}_1$, $\boldsymbol{w}_2 \in \boldsymbol{H}^{1/2}(\partial\Omega,\mathbb{R}^2)$ are two different extensions of $\boldsymbol{v}$ to $\partial\Omega$, then clearly $(\boldsymbol{w}_1 - \boldsymbol{w}_2)|_\Gamma \in \boldsymbol{H}^{1/2}_{00}(\Gamma,\mathbb{R}^2)$. Hence $< \boldsymbol{\tau n}, \boldsymbol{w}_1 >_{\partial\Omega} = < \boldsymbol{\tau n}, \boldsymbol{w}_2 >_{\partial\Omega}$.

Finally, we shall mention the properties of $\boldsymbol{H}_{0,\Gamma}(\mathbf{div}\,,\Omega,\mathbb{S}_2)$ relative to a partition of Ω. Let $\Omega = \cup_{i=1}^k \Omega_i$ be a partition into non-overlapping sub-domains, each of which has a Lipschitz boundary $\partial\Omega_i$. Then, similar to Proposition 1.2 on Page 95 of [22], we have the following lemma.

Lemma II.3. *The following statement is true:*

$$\boldsymbol{H}_{0,\Gamma}(\mathbf{div}\,, \Omega, \mathbb{S}_2) = \{\boldsymbol{\tau} \text{ satisfying } \boldsymbol{\tau}|_{\Omega_i} \in \boldsymbol{H}(\mathbf{div}\,, \Omega_i, \mathbb{S}_2) \text{ and}$$
$$\sum_{i=1}^{k} < \boldsymbol{\tau}\boldsymbol{n}_i, \boldsymbol{v} >_{\partial\Omega_i} = 0, \text{ for all } \boldsymbol{v} \in \boldsymbol{H}^1_{0,\partial\Omega\backslash\Gamma}(\Omega, \mathbb{R}^2)\},$$

where $\boldsymbol{n}_i$ is the unit outward normal vector on $\partial\Omega_i$.

B. The linearized theory of elasticity

The theory of elasticity deals with the deformation and the internal force of a given body under external forces and boundary conditions. We call the body a *continuum* if the physical quantities distributed over the body can be represented as continuous fields or piecewise continuous fields.

Our analysis will focus on plane elasticity problems. We use the displacement field $\boldsymbol{u} \in \mathbb{R}^2$ to represent the deformation of a body. In the linearized theory of elasticity, the *infinitesimal strain tensor* of the body is defined by

$$\boldsymbol{\varepsilon}(\boldsymbol{u}) = \frac{1}{2}\left[\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^t\right].$$

The strain tensor describes the deformation independent of both translation and rotation. Indeed, if we define the space of *infinitesimal rigid body motion*:

$$RM := span\left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}-y\\x\end{pmatrix}\right\},$$

which represents the translation and infinitesimal rigid body rotation, then RM is exactly the kernel of $\boldsymbol{\varepsilon}(\cdot)$. The strain tensor $\boldsymbol{\varepsilon}$ is symmetric by definition and therefore $\boldsymbol{\varepsilon} \in \mathbb{S}_2$.

When external forces are exerted on a body, the material volumes interact by exerting internal forces on one another. The internal forces can be characterized by the *linearized Cauchy stress tensor* $\boldsymbol{\sigma}$ and the stress tensor field must satisfy the local balance of linear and angular momentum. We only consider the time-independent case, or in other words, the equilibrium status:

$$
\begin{aligned}
-\operatorname{\mathbf{div}} \boldsymbol{\sigma} &= \boldsymbol{f}, && \textit{(balance of linear momentum)} \\
\boldsymbol{\sigma} &= \boldsymbol{\sigma}^t, && \textit{(balance of angular momentum)}
\end{aligned}
$$

where $\boldsymbol{f}$ is the body force per unit volume. The balance of angular momentum is equivalent to the statement that $\boldsymbol{\sigma} \in \mathbb{S}_2$. The balance of linear momentum is normally referred to as the *equilibrium equation*.

The stress and the strain are related to each other by the general Hooke's law. The relation between the stress and the strain, normally known as the *constitutive equation*, is a property of the material that constitutes the body. Assume we have a *linear hyper-elastic material*, which means that the stress $\boldsymbol{\sigma}$ is a linear function of the strain $\boldsymbol{\varepsilon}$ and there exists a well-defined *strain energy density function* $\mathcal{U}(\boldsymbol{\varepsilon})$ such that

$$
\sigma_{ij} = \frac{\partial \mathcal{U}(\boldsymbol{\varepsilon})}{\partial \varepsilon_{ij}}, \quad \text{for } 1 \leq i, j \leq 2.
$$

Then, the constitutive equation for the infinitesimal deformation [51] can be written as

$$
\boldsymbol{\sigma} = \mathbb{C}\boldsymbol{\varepsilon},
$$

where $\mathbb{C}$ is a symmetric fourth-order tensor which is called the *stiffness tensor*. An equivalent way to state the constitutive equation is

$$
\boldsymbol{\varepsilon} = \mathbb{A}\boldsymbol{\sigma},
$$

where $\mathbb{A} = \mathbb{C}^{-1}$ is the *compliance tensor*. Clearly $\mathbb{A}$ is also a symmetric and fourth-order tensor. We say the material is *homogeneous* if $\mathbb{A}$ and $\mathbb{C}$ are spatially independent tensors.

One special type of elasticity material is the so-called *isotropic* material. We say a material is isotropic if the stiffness tensor $\mathbb{C}$ is invariant with respect to all orthogonal transformations of the coordinate system. In this case, the constitutive equation becomes ([51])

$$\boldsymbol{\sigma} = 2\mu\boldsymbol{\varepsilon} + \lambda(tr\boldsymbol{\varepsilon})\mathbf{I}, \tag{2.4}$$

or

$$\boldsymbol{\varepsilon} = \frac{1}{2\mu}\boldsymbol{\sigma} - \frac{\lambda}{4\mu(\lambda+\mu)}(tr\boldsymbol{\sigma})\mathbf{I}, \tag{2.5}$$

where $tr\boldsymbol{\varepsilon}$ is the trace of $\boldsymbol{\varepsilon}$, $\mathbf{I}$ is the 2×2 identity matrix and λ, μ are the *Lamè* coefficients which satisfy $(\mu, \lambda) \in [\mu_1, \mu_2] \times (0, \infty)$ for $0 < \mu_1 < \mu_2$. If a material satisfies

$$tr\boldsymbol{\varepsilon} = div\, \boldsymbol{u} = 0,$$

we say it is *incompressible*. Notice that, by (2.4), the problem of incompressible material can be considered to be the extreme situation of λ tends to ∞. Therefore, we say a material is *nearly incompressible* if λ is large, comparing to μ. In this case, $|\mathbb{C}|$ is large while $|\mathbb{A}|$ is bounded.

Combining all the above, we have the so-called *field equations* for linear elasticity problems:

$$\begin{aligned} \boldsymbol{\varepsilon} &= \frac{1}{2}\left[\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^t\right], \\ \boldsymbol{\sigma} &= \mathbb{C}\boldsymbol{\varepsilon}, \\ -\mathbf{div}\,\boldsymbol{\sigma} &= \boldsymbol{f}. \end{aligned} \tag{2.6}$$

System (2.6) is not a well-posted problem unless we provide it with appropriate boundary conditions. Let Ω be the region occupied by the body we are studying and let $\partial\Omega = \Gamma_D \cup \Gamma_T$ where $\Gamma_D \cap \Gamma_T = \emptyset$. We consider two types of boundary conditions:

$$\begin{aligned} &\textit{displacement boundary condition:} \qquad \boldsymbol{u} = \boldsymbol{u}_0 \quad \text{on } \Gamma_D, \\ &\textit{traction boundary condition:} \qquad \boldsymbol{\sigma n} = \boldsymbol{t}_0 \quad \text{on } \Gamma_T, \end{aligned} \tag{2.7}$$

where $\boldsymbol{n}$ is the outward normal vector on Γ.

Finally, we mention the applications of plane elasticity problems. In the real world, elastic bodies are alway three-dimensional. However, there are cases when the displacement $\boldsymbol{u}$, the strain $\boldsymbol{\varepsilon}$ and the stress $\boldsymbol{\sigma}$ are nearly independent of one spatial coordinate. In these cases, the elastic body can be modeled by the plane elasticity problem. Two typical types of plane elasticity problems are the *plane strain problem* and the *plane stress problem*. The plain strain problem models the behavior of a cylindrical body under external forces parallel to its cross sections. The model is based on an assumption that there is no displacement along the direction of the axis of the cylindrical body. The plain stress problem models the behavior of a thin plate under external forces parallel to the plate. The model is based on an assumption that the stress is confined on the plane parallel to the plate. Figure 1 illustrates these two models.

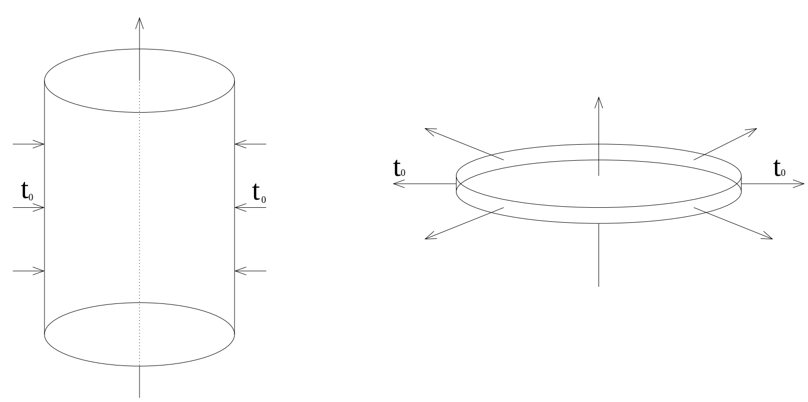

FIGURE 1. Typical plane strain model and plane stress model.

C. Variational principles

Assume that Ω is a bounded, open subset of $\mathbb{R}^2$ with a Lipschitz continuous boundary. Furthermore, we assume that both $\mathbb{C}$ and $\mathbb{A}$ are uniformly positive definite in Ω and bounded above. Consider Problem (2.6) under boundary conditions (2.7). There are three different ways to formulate it variationally: the primal variational principle, the dual variational principle and the mixed variational principle. First, we introduce the primal variational principle. Assume temporarily that the displacement boundary condition satisfies $\boldsymbol{u}_0 = \mathbf{0}$.

(Primal variational principle) Given $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega, \mathbb{R}^2)$ and $\boldsymbol{t}_0 \in (\boldsymbol{H}^{1/2}_{00}(\Gamma_T, \mathbb{R}^2))^*$, find $\boldsymbol{u} \in \boldsymbol{H}^1_{0,\Gamma_D}(\Omega, \mathbb{R}^2)$ such that $\boldsymbol{u}$ minimizes the potential energy functional

$$\int_\Omega \frac{1}{2}\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{\varepsilon}(\boldsymbol{v})\, d\boldsymbol{x} - (\boldsymbol{f}, \boldsymbol{v}) - < \boldsymbol{t}_0, \boldsymbol{v} >_{\Gamma_T},$$

over the space $\boldsymbol{H}^1_{0,\Gamma_D}(\Omega, \mathbb{R}^2)$. The corresponding weak problem can be written as: *Find* $\boldsymbol{u} \in H^1_{0,\Gamma_D}(\Omega, \mathbb{R}^2)$ *such that*

$$(\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}), \boldsymbol{\varepsilon}(\boldsymbol{v})) = (\boldsymbol{f}, \boldsymbol{v}) + < \boldsymbol{t}_0, \boldsymbol{v} >_{\Gamma_T} \qquad \text{for all } \boldsymbol{v} \in H^1_{0,\Gamma_D}(\Omega, \mathbb{R}^2). \tag{2.8}$$

Korn's Inequality [21, 25] states that for all $\boldsymbol{v} \in \boldsymbol{H}^1_{0,\Gamma_D}(\Omega, \mathbb{R}^2)$ (when $\Gamma_D \neq \emptyset$) or $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega, \mathbb{R}^2)/RM$ (when $\Gamma_D = \emptyset$), there exists a positive constant C independent of $\boldsymbol{v}$ such that

$$\|\boldsymbol{\varepsilon}(\boldsymbol{v})\|_{0,\Omega} \geq C\|\boldsymbol{v}\|_{1,\Omega}.$$

By the Lax-Milgram Lemma, we have the following theorem about the existence and the uniqueness of the weak solution.

Theorem II.3. *If $\Gamma_D \neq \emptyset$, then Problem (2.8) has a unique solution in $\boldsymbol{H}^1_{0,\Gamma_D}(\Omega, \mathbb{R}^2)$.*

If $\Gamma_D = \emptyset$*, and we have the following compatibility condition,*

$$(\boldsymbol{f}, \boldsymbol{v}) + < \boldsymbol{t}_0, \boldsymbol{v} >_{\partial\Omega} = 0 \quad \textit{for all } \boldsymbol{v} \in RM, \tag{2.9}$$

then Problem (2.8) has a solution and the solution is unique in $\boldsymbol{H}^1(\Omega, \mathbb{R}^2)/RM$*.*

In the case that $\boldsymbol{u}_0 \neq \boldsymbol{0}$, assume that $\boldsymbol{u}_0 \in \boldsymbol{H}^{1/2}(\Gamma_D, \mathbb{R}^2)$. There exists a $\tilde{\boldsymbol{u}}_0 \in \boldsymbol{H}^1(\Omega, \mathbb{S}^2)$ such that $\tilde{\boldsymbol{u}}_0|_{\Gamma_D} = \boldsymbol{u}_0$. Consider the following problem: *Find* $\boldsymbol{w} \in H^1_{0,\Gamma_D}(\Omega, \mathbb{R}^2)$ *such that*

$$(\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{w}), \boldsymbol{\varepsilon}(\boldsymbol{v})) = (\boldsymbol{f}, \boldsymbol{v}) + < \boldsymbol{t}_0, \boldsymbol{v} >_{\Gamma_T} - (\mathbb{C}\boldsymbol{\varepsilon}(\tilde{\boldsymbol{u}}_0), \boldsymbol{\varepsilon}(\boldsymbol{v})) \qquad \text{for all } \boldsymbol{v} \in H^1_{0,\Gamma_D}(\Omega, \mathbb{R}^2).$$

The right hand side of the above equation is a well-defined functional on $H^1_{0,\Gamma_D}(\Omega, \mathbb{R}^2)$. Hence the problem is well posed. Clearly, $\boldsymbol{w} + \tilde{\boldsymbol{u}}_0$ is the weak solution for Problem (2.6) with boundary conditions (2.7). Therefore, the non-homogeneous boundary problem can be reduced to a homogeneous boundary problem with a different right-hand side. Hence the theoretical analysis only needs to be done on Problem (2.8).

Notice that $\boldsymbol{w} + \tilde{\boldsymbol{u}}_0$ is the solution of the following problem: *Find* $\boldsymbol{u} \in H^1(\Omega, \mathbb{R}^2)$ *such that* $\boldsymbol{u}|_{\Gamma_D} = \boldsymbol{u}_0$ *and*

$$(\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}), \boldsymbol{\varepsilon}(\boldsymbol{v})) = (\boldsymbol{f}, \boldsymbol{v}) + < \boldsymbol{t}_0, \boldsymbol{v} >_{\Gamma_T} \qquad \text{for all } \boldsymbol{v} \in H^1_{0,\Gamma_D}(\Omega, \mathbb{R}^2). \tag{2.10}$$

Therefore in the implementation of finite element methods, we actually approximate Problem (2.10). The solution for Problem (2.10) exists and is unique in the same sense as stated in Theorem II.3.

Next, we consider the dual variational principle and the mixed variational principle for linear plane elasticity. Assume temporarily that $\boldsymbol{t}_0 = \boldsymbol{0}$. Similar to the analysis in Chapter 1 of [22], the primal formulation can be transformed to

(Dual variational principle) Given $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega, \mathbb{R}^2)$ and $\boldsymbol{u}_0 \in \boldsymbol{H}^{1/2}(\Gamma_D, \mathbb{R}^2)$, find $\boldsymbol{\sigma} \in \boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,, \Omega, \mathbb{S}_2)$ such that $-\mathbf{div}\,\boldsymbol{\sigma} = \boldsymbol{f}$ and $\boldsymbol{\sigma}$ minimizes the complimentary energy functional

$$\int_\Omega \frac{1}{2}\mathbb{A}\boldsymbol{\tau} : \boldsymbol{\tau}\, d\boldsymbol{x} - < \boldsymbol{\tau n}, \boldsymbol{u}_0 >_{\Gamma_D},$$

over the space $\boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,, \Omega, \mathbb{S}_2)$;

(Mixed variational principle) Given $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega, \mathbb{R}^2)$ and $\boldsymbol{u}_0 \in \boldsymbol{H}^{1/2}(\Gamma_D, \mathbb{R}^2)$, find $(\boldsymbol{\sigma}, \boldsymbol{u}) \in (\boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,, \Omega, \mathbb{S}_2), \boldsymbol{L}^2(\Omega, \mathbb{R}^2))$ which is the critical point of the following saddle-point problem

$$\inf_{\boldsymbol{\tau} \in \boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,, \Omega, \mathbb{S}_2)} \sup_{\boldsymbol{v} \in \boldsymbol{L}^2(\Omega, \mathbb{R}^2)} \int_\Omega (\frac{1}{2}\mathbb{A}\boldsymbol{\tau} : \boldsymbol{\tau} + \mathbf{div}\,\boldsymbol{\tau} \cdot \boldsymbol{v})\, d\boldsymbol{x} + (\boldsymbol{f}, \boldsymbol{v}) - < \boldsymbol{\tau n}, \boldsymbol{u}_0 >_{\Gamma_D}.$$

Notice that the term $< \boldsymbol{\tau n}, \boldsymbol{u}_0 >_{\Gamma_D}$ in both the dual variational principle and the mixed variational principle is defined in the sense of (2.3).

Among the three variational principles we stated, only the dual formulation does not involve the displacement field $\boldsymbol{u}$. Consider the problem of solving for $\boldsymbol{u} \in \mathbb{R}^2$ corresponding to a given symmetric matrix field $\boldsymbol{\varepsilon} = (\varepsilon_{ij})_{1 \le i,j \le 2}$. There are three independent equations but only two unknowns. The system is overdetermined. To make the problem solvable, $\boldsymbol{\varepsilon}$ has to satisfy the following compatibility condition:

$$\frac{\partial^2}{\partial y^2}\varepsilon_{11} + \frac{\partial^2}{\partial x^2}\varepsilon_{22} = 2\frac{\partial^2}{\partial x \partial y}\varepsilon_{12}. \tag{2.11}$$

The compatibility condition on $\boldsymbol{\varepsilon}$ implies that all symmetric matrix functions are not possible strain fields. This causes problem for the dual principle, since condition (2.11) has to be explicitly added to the formulation to make the solution meaningful. Therefore, it is impractical to use the dual variational principle in real applications.

We will focus on the mixed variational principle, which is commonly referred to as the Hellinger-Reissner Principle. Compared to the primal-based methods, mixed methods have some well-known advantages [2, 22]. For example, the dual variable $\boldsymbol{\sigma}$, which is usually the variable of primary interest, is computed directly as a fundamental unknown. Mixed methods also have some obvious disadvantages, such as the necessity of constructing stable pairs of finite element spaces and the fact that the resulting discrete system is indefinite.

From the mixed variational principle we can derive the mixed formulation for the linear elasticity problem, which is often referred to as the *stress-displacement formulation*. Assume that $\boldsymbol{t}_0 = \mathbf{0}$. The mixed problem is: *given* $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega, \mathbb{R}^2)$ *and* $\boldsymbol{u}_0 \in \boldsymbol{H}^{1/2}(\Gamma_D, \mathbb{R}^2)$, *find* $(\boldsymbol{\sigma}, \boldsymbol{u}) \in (\boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,, \Omega, \mathbb{S}_2), \boldsymbol{L}^2(\Omega, \mathbb{R}^2))$ *such that*

$$\begin{cases} (\mathbb{A}\boldsymbol{\sigma}, \boldsymbol{\tau}) + (\mathbf{div}\,\boldsymbol{\tau}, \boldsymbol{u}) = < \boldsymbol{\tau}\boldsymbol{n}, \boldsymbol{u}_0 >_{\Gamma_D} & \text{for all } \boldsymbol{\tau} \in \boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,, \Omega, \mathbb{S}_2), \\ (\mathbf{div}\,\boldsymbol{\sigma}, \boldsymbol{v}) = (-\boldsymbol{f}, \boldsymbol{v}) & \text{for all } \boldsymbol{v} \in \boldsymbol{L}^2(\Omega, \mathbb{R}^2). \end{cases} \tag{2.12}$$

Theorem II.4. *If* $\Gamma_D \neq \emptyset$, *then the mixed problem (2.12) has a unique solution. If* $\Gamma_D = \emptyset$ *and* $\boldsymbol{f}$ *satisfies the compatibility condition (2.9), then Problem (2.12) has a solution and the solution is unique in* $(\boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,, \Omega, \mathbb{S}_2), \boldsymbol{L}^2(\Omega, \mathbb{R}^2)/RM)$.

Proof. Let $\boldsymbol{u}$ be the solution for Problem (2.10) with $\boldsymbol{t}_0 = \mathbf{0}$ and define $\boldsymbol{\sigma} = \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u})$. By Green's formula, it is not hard to see that $\boldsymbol{\sigma} \in \boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,, \Omega, \mathbb{S}_2)$ and $\mathbf{div}\,\boldsymbol{\sigma} = -\boldsymbol{f}$. Therefore, by Theorem II.3, $Im(\mathbf{div}\,) = \boldsymbol{L}^2(\Omega, \mathbb{R}^2)$ when $\Gamma_D \neq \emptyset$ and $Im(\mathbf{div}\,) = \boldsymbol{L}^2(\Omega, \mathbb{R}^2)/RM$ when $\Gamma_D = \emptyset$. Hence $Im(\mathbf{div}\,)$ is closed in $\boldsymbol{L}^2(\Omega, \mathbb{R}^2)$. Since we have assumed that $\mathbb{A}$ is uniformly positive definite in Ω, so $(\mathbb{A}\cdot, \cdot)$ is coercive on $Ker(\mathbf{div}\,)$. By Theorem 1.1 in Chapter II of [22], Problem (2.12) has a solution and the solution is unique in $(\boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,, \Omega, \mathbb{S}_2), Im(\mathbf{div}\,))$. This completes the proof of the theorem. $\square$

Consider the case when $\boldsymbol{t}_0 \neq \mathbf{0}$. Assume that $\boldsymbol{t}_0 \in (\boldsymbol{H}_{00}^{1/2}(\Omega, \mathbb{R}^2))^*$. By Theorem II.3, there exists a $\tilde{\boldsymbol{\sigma}} \in \boldsymbol{H}(\mathbf{div}\,, \Omega, \mathbb{S}_2)$ such that $\tilde{\boldsymbol{\sigma}}\boldsymbol{n}|_{\Gamma_T} = \boldsymbol{t}_0$. Similar to the derivation of Problem (2.10), the non-homogeneous boundary problem can be reduced to a homogeneous boundary problem with a different right-hand side. Using linearity, it is easy to see that the mixed problem with $\boldsymbol{t}_0 \neq \mathbf{0}$ can be written as: *Find* $(\boldsymbol{\sigma}, \boldsymbol{u}) \in (\boldsymbol{H}(\mathbf{div}\,, \Omega, \mathbb{S}_2), \boldsymbol{L}^2(\Omega, \mathbb{R}^2))$ *such that* $\boldsymbol{\sigma n}|_{\Gamma_T} = \boldsymbol{t}_0$ *and*

$$\begin{cases} (\mathbb{A}\boldsymbol{\sigma}, \boldsymbol{\tau}) + (\mathbf{div}\,\boldsymbol{\tau}, \boldsymbol{u}) =< \boldsymbol{\tau n}, \boldsymbol{u}_0 >_{\Gamma_D} & \text{for all } \boldsymbol{\tau} \in \boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,, \Omega, \mathbb{S}_2), \\ (\mathbf{div}\,\boldsymbol{\sigma}, \boldsymbol{v}) = (-\boldsymbol{f}, \boldsymbol{v}) & \text{for all } \boldsymbol{v} \in \boldsymbol{L}^2(\Omega, \mathbb{R}^2). \end{cases} \tag{2.13}$$

Notice that in the mixed formulation, the traction boundary condition becomes the essential boundary condition and the displacement boundary condition becomes the natural boundary condition. Problem (2.13) is used in the implementation of mixed finite element methods, while theoretical analysis only needs to be done on Problem (2.12).

Clearly, the solution for Problem (2.13) exists and is unique in the same sense as stated in Theorem II.3 and Theorem II.4. Let $\boldsymbol{u}$ be the solution for Problem (2.10) and define $\boldsymbol{\sigma} = \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u})$. By Green's formula, it is not hard to see that $\boldsymbol{\sigma} \in \boldsymbol{H}(\mathbf{div}\,, \Omega, \mathbb{S}_2)$ and $\mathbf{div}\,\boldsymbol{\sigma} = -\boldsymbol{f}$. Furthermore, one can easily verify that $\boldsymbol{\sigma n}|_{\Gamma_T} = \boldsymbol{t}_0$ and

$$(\mathbb{A}\boldsymbol{\sigma}, \boldsymbol{\tau}) + (\mathbf{div}\,\boldsymbol{\tau}, \boldsymbol{u}) =< \boldsymbol{\tau n}, \boldsymbol{u}_0 >_{\Gamma_D} .$$

Therefore, $(\boldsymbol{\sigma}, \boldsymbol{u})$ is also the solution for Problem (2.13). Hence the mixed problem (2.13) is equivalent to the primal based problem (2.10) in the sense that they have the same solution.

Finally, we discuss some regularity results for elasticity problems. Assume that $\boldsymbol{f}$ is in $\boldsymbol{L}^2(\Omega, \mathbb{R}^2)$, $\boldsymbol{t}_0 = \mathbf{0}$ and $\boldsymbol{u}_0 = \mathbf{0}$. We say that the weak solution of Problem (2.13) has $\boldsymbol{H}^s$-regularity for some $0 < s \leq 1$ if the solution $(\boldsymbol{\sigma}, \boldsymbol{u})$ satisfies $\boldsymbol{\sigma} \in$

$\boldsymbol{H}(\mathbf{div}\,,\Omega,\mathbb{S}_2)\cap\boldsymbol{H}^s(\Omega,\mathbb{S}_2)$, $\boldsymbol{u}\in\boldsymbol{H}^{1+s}(\Omega,\mathbb{R}^2)$ and

$$\|\boldsymbol{\sigma}\|_{s,\Omega}+\|\boldsymbol{u}\|_{1+s,\Omega}\le C_R\|\boldsymbol{f}\|_{0,\Omega}, \tag{2.14}$$

where C_R is a positive constant depending only on Ω, Γ_T and $\mathbb{A}$.

We state some regularity results on a polygonal domain, which were proved by Grisvard [38, 39, 40]. Let Ω be a polygon with N corners. When we consider corners of Ω, we always include the points where Γ_T and Γ_D meet, even if they may not be actual corners of Ω. Let ω_j, $1\le j\le N$, be the measure of the interior angle at the j-th corner. Consider the Lamé system, that is, $\boldsymbol{\sigma}=2\mu\boldsymbol{\varepsilon}+\lambda tr(\boldsymbol{\varepsilon})\mathbf{I}$. We have

- for the pure displacement boundary problem, Problem (2.13) has $\boldsymbol{H}^s$-regularity for $s\le 1$ such that

$$s<\inf_{j=1,\dots,N}\{Re\,z;\ \sin^2(z\omega_j)=\frac{(\lambda+\mu)^2}{(\lambda+3\mu)^2}\,z^2\sin^2\omega_j,\ 0<Re\,z<1\};$$

- for the mixed boundary problem, Problem (2.13) has $\boldsymbol{H}^s$-regularity for $s\le 1$ such that

$$s<\inf_{j=1,\dots,N}\{Re\,z;\ \sin^2(z\omega_j)=\frac{(\lambda+2\mu)^2-(\lambda+\mu)^2z^2\sin^2\omega_j}{(\lambda+\mu)(\lambda+3\mu)},\ 0<Re\,z<1\};$$

- for the pure traction boundary problem, Problem (2.13) has $\boldsymbol{H}^s$-regularity for $\boldsymbol{\sigma}\in\boldsymbol{H}^s(\Omega,\mathbb{S}_2)$ for $s\le 1$ such that

$$s<\inf_{j=1,\dots,N}\{Re\,z;\ \sin^2(z\omega_j)=z^2\sin^2\omega_j,\ 0<Re\,z<1\}.$$

It turns out that:

- (for the pure displacement boundary problem or the pure traction boundary problem) Problem (2.13) has $\boldsymbol{H}^1$-regularity when Ω is a convex polygon. For a non-convex polygonal domain where no internal angle is equal to 2π, Problem

(2.13) has $\boldsymbol{H}^s$-regularity for $s \in (0, s_0)$, where $1/2 < s_0 < 1$ depends on the Lamé coefficients.

- (for the mixed boundary problem) Convexity of the domain no longer guarantees $\boldsymbol{H}^1$-regularity. For a non-convex polygonal domain where no internal angle is equal to 2π, Problem (2.13) has $\boldsymbol{H}^s$-regularity for $s \in (0, s_0)$, where s_0 depends on the Lamé coefficients and is not necessarily greater than 1/2.

Notice that the constant C_R in Inequality (2.14) depends on the Lamé coefficients λ and μ. When $\lambda \to \infty$, C_R may get large. A recent result by Bacuta and Bramble [11] states that if Ω is a convex polygon, then for the pure displacement boundary problem we have

$$\|\boldsymbol{u}\|_{2,\Omega} \leq c\|\boldsymbol{f}\|_{0,\Omega},$$

where c is a positive constant independent of the Lamé coefficients.

D. The Airy operator

The Airy stress function was first introduced by British astronomer, Sir George Biddell Airy (1801-1892). Recall that we have the equilibrium equation $-\mathbf{div}\,\boldsymbol{\sigma} = \boldsymbol{f}$. If the body force $\boldsymbol{f} = \mathbf{0}$, Airy noticed that the following type of symmetric matrix field,

$$\boldsymbol{\sigma} = \mathbf{airy}\, q = \begin{pmatrix} \frac{\partial^2}{\partial y^2} q & -\frac{\partial^2}{\partial x \partial y} q \\ -\frac{\partial^2}{\partial x \partial y} q & \frac{\partial^2}{\partial x^2} q \end{pmatrix},$$

satisfies the equilibrium equation for arbitrary $q \in C^3(\Omega)$. On the other hand, for a given stress field $\boldsymbol{\sigma}$ which satisfies $\mathbf{div}\,\boldsymbol{\sigma} = \mathbf{0}$, if there exists a q such that $\mathbf{airy}\, q = \boldsymbol{\sigma}$, then we say that q is the *Airy stress function* for $\boldsymbol{\sigma}$.

Define

$$\operatorname{curl} q = \begin{pmatrix} \frac{\partial}{\partial y} q \\ -\frac{\partial}{\partial x} q \end{pmatrix} \quad \text{and} \quad \mathbf{curl}\, \boldsymbol{v} = \begin{pmatrix} \frac{\partial}{\partial y} v_1 & -\frac{\partial}{\partial x} v_1 \\ \frac{\partial}{\partial y} v_2 & -\frac{\partial}{\partial x} v_2 \end{pmatrix}.$$

It is clear that

$$\mathbf{airy}\, q = \mathbf{curl} \operatorname{curl} q.$$

Define

$$\boldsymbol{u} \times \boldsymbol{v} = u_1 v_2 - u_2 v_1, \qquad \text{for } \boldsymbol{u} = (u_i)_{1 \le i \le 2},\ \boldsymbol{v} = (v_i)_{1 \le i \le 2} \in \mathbb{R}^2,$$

$$\boldsymbol{\tau} \times \boldsymbol{v} = \begin{pmatrix} \tau_{11} v_2 - \tau_{12} v_1 \\ \tau_{21} v_2 - \tau_{22} v_1 \end{pmatrix}, \qquad \text{for } \boldsymbol{\tau} = (\tau_{ij})_{1 \le i,j \le 2} \in \mathbb{R}^4,\ \boldsymbol{v} = (v_i)_{1 \le i \le 2} \in \mathbb{R}^2.$$

Assume that Ω is a polygon. For $\boldsymbol{\sigma} \in C^\infty(\Omega, \mathbb{S}_2)$ satisfying $\mathbf{div}\, \boldsymbol{\sigma} = \mathbf{0}$, it is fundamental to show that there exists $q \in C^\infty(\Omega)$ such that $\boldsymbol{\sigma} = \mathbf{airy}\, q$. Indeed, let $\boldsymbol{x}_0$ be a fixed point in Ω and for all $\boldsymbol{x} \in \Omega$, let $\gamma_{\boldsymbol{x}}$ be a smooth path from $\boldsymbol{x}_0$ to $\boldsymbol{x}$. Define

$$w_1(\boldsymbol{x}) = \int_{\gamma_{\boldsymbol{x}}} \begin{pmatrix} \sigma_{22} \\ -\sigma_{21} \end{pmatrix} \cdot \boldsymbol{t}\, ds,$$

$$w_2(\boldsymbol{x}) = \int_{\gamma_{\boldsymbol{x}}} \begin{pmatrix} -\sigma_{12} \\ \sigma_{11} \end{pmatrix} \cdot \boldsymbol{t}\, ds,$$

where $\boldsymbol{t}$ is the unit tangential vector on $\gamma_{\boldsymbol{x}}$ pointing toward $\boldsymbol{x}$. By Green's formula, one can easily show that w_1 and w_2 are independent of the choice of $\gamma_{\boldsymbol{x}}$. Define

$$q = \int_{\gamma_{\boldsymbol{x}}} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \cdot \boldsymbol{t}\, ds.$$

It is not hard to see that q is also independent of the choice of $\gamma_{\boldsymbol{x}}$ and $\mathbf{airy}\, q = \boldsymbol{\sigma}$. Notice that the above argument also works for $\boldsymbol{\sigma} \in \boldsymbol{H}(\mathbf{div}\,, \Omega, \mathbb{S}_2)$ satisfying $\mathbf{div}\, \boldsymbol{\sigma} = \mathbf{0}$ and the resulting q is in $H^2(\Omega)$. Combining the above analysis and Theorem II.3, we have the following lemma.

Lemma II.4. *The following exact sequence holds:*

$$0 \to \; P_1(\Omega) \; \xrightarrow{\subset} \; H^2(\Omega) \; \xrightarrow{\textbf{airy}} \; \boldsymbol{H}(\textbf{div}\,, \Omega, \mathbb{S}_2) \; \xrightarrow{\textbf{div}} \; \boldsymbol{L}^2(\Omega, \mathbb{R}^2) \; \to 0,$$

We want to derive a similar exact sequence for $\boldsymbol{H}_{0,\Gamma_T}(\textbf{div}\,, \Omega, \mathbb{S}_2)$ where $\Gamma_T \neq \emptyset$. First, we state the following lemmas.

Lemma II.5. *For $q \in H^2(\Omega)$ and $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega, \mathbb{R}^2)$, we have*

$$< (\textbf{airy}\, q)\boldsymbol{n}, \boldsymbol{v} >_{\partial\Omega} = < \operatorname{curl} q, \nabla \boldsymbol{v} \times \boldsymbol{n} >_{\partial\Omega},$$

where $\boldsymbol{n}$ is the unit outward normal vector.

Proof. By Green's formula, for $q \in C^\infty(\Omega)$ and $\boldsymbol{v} \in C^\infty(\Omega, \mathbb{R}^2)$,

$$\begin{aligned} < (\textbf{airy}\, q)\boldsymbol{n}, \boldsymbol{v} >_{\partial\Omega} &= (\textbf{div}\, \textbf{airy}\, q, \boldsymbol{v}) + (\textbf{airy}\, q, \nabla \boldsymbol{v}) \\ &= 0 + (\textbf{curl}\, \operatorname{curl} q, \nabla \boldsymbol{v}) \\ &= < \operatorname{curl} q, \nabla \boldsymbol{v} \times \boldsymbol{n} >_{\partial\Omega} . \end{aligned}$$

Using a density argument, we can show that the result is true for $q \in H^2(\Omega)$ and $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega, \mathbb{R}^2)$. □

Remark II.3. *Notice that for $\boldsymbol{u}$, $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega, \mathbb{R}^2)$, by the Green's formula and a density argument,*

$$< \boldsymbol{u}, \nabla \boldsymbol{v} \times \boldsymbol{n} >_{\partial\Omega} = (\textbf{curl}\, \boldsymbol{u}, \nabla \boldsymbol{v}) \leq \|\boldsymbol{u}\|_{1,\Omega} \|\boldsymbol{v}\|_{1,\Omega}.$$

By setting $\boldsymbol{u}$ to be harmonic in Ω, we have $\|\boldsymbol{u}\|_{1,\Omega} \leq c\|\boldsymbol{u}\|_{1/2,\partial\Omega}$ where c is a positive constant independent of $\boldsymbol{u}$. Therefore, $\nabla \boldsymbol{v} \times \boldsymbol{n}$ is in $H^{-1/2}(\partial\Omega)$. Furthermore, for

$\boldsymbol{v} \in C^\infty(\Omega, \mathbb{R}^2)$, the result of Lemma II.5 can be written as

$$< (\mathbf{airy}\, q)\boldsymbol{n}, \boldsymbol{v} >_{\partial\Omega} = \sum_{i=1}^{N} < \operatorname{curl} q, \nabla \boldsymbol{v} \times \boldsymbol{n}_i >_{\Gamma_i} \qquad \textit{for all } q \in H^2(\Omega),$$

where Γ_i, $i = 1, \ldots, N$, are the edges of the polygon Ω and $\boldsymbol{n}_i$ are the outward normal vectors on each Γ_i.

Lemma II.6. *Consider the line segment $l = (0, 1)$. For $q \in L^2(l)$ satisfying*

$$\int_l q \frac{dv}{dx}\, dx = 0 \quad \textit{for all } v \in C_0^\infty(l),$$

we can conclude that $q = constant$ on l.

Proof. Define $L_0^2(l) = \{v \in L^2(l) \mid \int_l v\, dx = 0\}$. First, we show that $\frac{d}{dx} C_0^\infty(l)$ is dense in $L_0^2(l)$. For $v \in L_0^2(l)$, let $w = \int_0^x v(s)\, ds$. It is clear that $w \in H_0^1(l)$ and there exists $\{w_n\}_{n=1}^\infty \subset C_0^\infty(l)$ such that $w_n \to w$ in $H_0^1(l)$. Therefore $\frac{dw_n}{dx} \in \frac{d}{dx} C_0^\infty(l)$ and $\frac{dw_n}{dx} \to v$ in $L_0^2(l)$. Hence $\frac{d}{dx} C_0^\infty(l)$ is dense in $L_0^2(l)$.

Therefore, q is orthogonal to $L_0^2(l)$ in the L^2 inner product. It is obvious that $L^2(l) = L_0^2(l) \oplus span\{1\}$. Hence $q = constant$ on l. $\square$

Define

$$\tilde{P}_1^{\Gamma_T}(\Omega) = P_1(\Omega) \cap H^2_{0,\Gamma_T}(\Omega).$$

Notice that $\tilde{P}_1^{\Gamma_T}(\Omega) = P_1(\Omega)$ when $\Gamma_T = \emptyset$ and $\tilde{P}_1^{\Gamma_T}(\Omega) = \emptyset$ when $\Gamma_T \neq \emptyset$. Now we can prove the following lemma:

Lemma II.7. *The following exact sequences are true:*

1. *If $\Gamma_T \neq \partial\Omega$ and Γ_T is connected,*

$$\tilde{P}_1^{\Gamma_T}(\Omega) \xrightarrow{\subset} H^2_{0,\Gamma_T}(\Omega) \xrightarrow{\mathbf{airy}} \boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}, \Omega, \mathbb{S}_2) \xrightarrow{\mathbf{div}} \boldsymbol{L}^2(\Omega, \mathbb{R}^2) \to 0,$$

2. If $\Gamma_T = \partial\Omega$

$$0 \to H_0^2(\Omega) \xrightarrow{\textbf{airy}} \boldsymbol{H}_0(\textbf{div}\,, \Omega, \mathbb{S}_2) \xrightarrow{\textbf{div}} \boldsymbol{L}^2(\Omega, \mathbb{R}^2) \to RM \to 0.$$

Proof. First, as stated in the proof of Theorem II.4, we have $Im(\textbf{div}\,) = \boldsymbol{L}^2(\Omega, \mathbb{R}^2)$ when $\Gamma_D \neq \emptyset$ and $Im(\textbf{div}\,) = \boldsymbol{L}^2(\Omega, \mathbb{R}^2)/RM$ when $\Gamma_D = \emptyset$.

By Lemma II.4, for $\boldsymbol{\tau} \in \boldsymbol{H}_{0,\Gamma_T}(\textbf{div}\,, \Omega, \mathbb{S}_2)$ satisfying $\textbf{div}\,\boldsymbol{\tau} = \boldsymbol{0}$, there exists $q \in H^2(\Omega)$, which is not unique, such that $\textbf{airy}\,q = \boldsymbol{\tau}$. We only need to show that q can be chosen in $H^2_{0,\Gamma_T}(\Omega)$. By Lemma II.5 and the definition of $\boldsymbol{H}_{0,\Gamma_T}(\textbf{div}\,, \Omega, \mathbb{S}_2)$, for all $\boldsymbol{v} \in \boldsymbol{H}^1_{0,\Gamma_D}(\Omega, \mathbb{R}^2)$,

$$< \operatorname{curl} q, \nabla \boldsymbol{v} \times \boldsymbol{n} >_{\partial\Omega} = 0.$$

Let l be an edge of the polygon Ω and assume that $l \in \Gamma_T$. For any $\boldsymbol{v} \in C_0^\infty(l, \mathbb{R}^2)$, we can extend $\boldsymbol{v}$ to $\partial\Omega$ by zero outside l and the resulting function, which is still denoted by $\boldsymbol{v}$, is in $C^\infty(\partial\Omega, \mathbb{R}^2)$. Let $\boldsymbol{t}$ be the unit tangential vector on l. Then

$$\begin{aligned} < \operatorname{curl} q, \frac{\partial}{\partial \boldsymbol{t}} \boldsymbol{v} >_l &= - < \operatorname{curl} q, \nabla \boldsymbol{v} \times \boldsymbol{n} >_l \\ &= - < \operatorname{curl} q, \nabla \boldsymbol{v} \times \boldsymbol{n} >_{\partial\Omega} \\ &= 0. \end{aligned}$$

By Lemma II.6, $\operatorname{curl} q = \mathit{constant}$ on l. Therefore, ∇q is piecewise constant on Γ_T. Since Γ_T is connected and $\nabla q \in \boldsymbol{H}^{1/2}(\Gamma_T)$, it is not hard to see that $\nabla q = \mathit{constant}$ on Γ_T. Notice that $Ker(\textbf{airy}\,) = P_1(\Omega)$. Without loss of generality, we can set $\nabla q = \boldsymbol{0}$ on Γ_T, which implies that $q = \mathit{constant}$ on Γ_T. Again, we can use the fact that $Ker(\textbf{airy}\,) = P_1(\Omega)$ to set $q = 0$ on Γ_T. Therefore, we have $q \in H^2_{0,\Gamma_T}(\Omega)$.

Finally, notice that $Ker(\textbf{airy}\,) = P_1(\Omega)$ and this implies the exactness condition on $H^2_{0,\Gamma_T}(\Omega)$. This completes the proof of the lemma. $\square$

Remark II.4. *The condition that* Γ_T *is connected is essential to the proof for Lemma*

II.7. One can easily construct counterexamples when Γ_T is not connected.

Lemma II.7 implies a decomposition of the space $\boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,, \Omega, \mathbb{S}_2)$:

$$\boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,, \Omega, \mathbb{S}_2) = \mathcal{H}_0 + \mathcal{H}_1,$$

where

$$\mathcal{H}_0 = Ker(\mathbf{div}\,) = \mathbf{airy}\,(H^2_{0,\Gamma_T}(\Omega)),$$

$$\mathcal{H}_1 = \{\boldsymbol{\tau} \in \boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,, \Omega, \mathbb{S}_2) \,|\, \boldsymbol{\tau} \perp \mathcal{H}_0 \text{ in the } \boldsymbol{H}(\mathbf{div}\,, \Omega, \mathbb{S}_2) \text{ inner product}\}.$$

Notice that $\mathcal{H}_0$ and $\mathcal{H}_1$ are orthogonal in both the $\boldsymbol{H}(\mathbf{div}\,, \Omega, \mathbb{S}_2)$ inner product and the $\boldsymbol{L}^2$ inner product.

Clearly, $\mathbf{div}$ is a bijection from $\mathcal{H}_1$ onto $Im(\mathbf{div}\,)$. Therefore, an inverse operator $\mathbf{div}^{-1}$ can be defined, which maps

$$\begin{cases} \boldsymbol{L}^2(\Omega) \text{ onto } \mathcal{H}_1, \text{ if } \Gamma_D \neq \emptyset; \\ \boldsymbol{L}^2(\Omega)/RM \text{ onto } \mathcal{H}_1, \text{ if } \Gamma_D = \emptyset. \end{cases}$$

Indeed, let $\boldsymbol{\sigma} \in \boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,, \Omega, \mathbb{S}_2)$ and $\boldsymbol{u} \in \boldsymbol{L}^2(\Omega, \mathbb{R}^2)$ satisfy

$$\begin{cases} (\boldsymbol{\sigma}, \boldsymbol{\tau}) + (\mathbf{div}\,\boldsymbol{\tau}, \boldsymbol{u}) = 0 & \text{for all } \boldsymbol{\tau} \in \boldsymbol{H}_{0.\Gamma_T}(\mathbf{div}\,, \Omega, \mathbb{S}_2), \\ (\mathbf{div}\,\boldsymbol{\sigma}, \boldsymbol{w}) = (\boldsymbol{v}, \boldsymbol{w}) & \text{for all } \boldsymbol{w} \in \boldsymbol{L}^2(\Omega, \mathbb{R}^2). \end{cases} \tag{2.15}$$

Then $\boldsymbol{\sigma} = \mathbf{div}^{-1}\boldsymbol{v} \in \mathcal{H}_1$. Furthermore, assuming $\boldsymbol{H}^s$-regularity for the above system, then by Inequality (2.14), we have $\mathbf{div}^{-1}\boldsymbol{v} \in \boldsymbol{H}^s(\Omega, \mathbb{S}_2)$ and

$$\|\mathbf{div}^{-1}\boldsymbol{v}\|_{s,\Omega} \leq C_R \|\boldsymbol{v}\|_{0,\Omega}.$$

We can conclude that, for all $\boldsymbol{\sigma} \in \boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,, \Omega, \mathbb{S}_2)$, there exists a unique decomposition

$$\boldsymbol{\sigma} = \mathbf{airy}\, q + \mathbf{div}^{-1}\boldsymbol{v},$$

where $q \in H^2_{0,\Gamma_T}(\Omega)$ and $\boldsymbol{v} = \mathbf{div}\,\boldsymbol{\sigma}$. We point out that, q can be considered as the solution of the following biharmonic problem: *Find* $q \in H^2_{0,\Gamma_T}(\Omega)$ *such that*

$$(\mathbf{airy}\, q, \mathbf{airy}\, p) = (\boldsymbol{\sigma}, \mathbf{airy}\, p) \quad \text{for all } p \in H^2_{0,\Gamma_T}(\Omega).$$

CHAPTER III

MIXED FINITE ELEMENT DISCRETIZATION

In this chapter, we discuss the mixed finite element discretization for the linear plane elasticity problem (2.12). The Arnold-Winther finite element spaces are introduced and proved to be stable under certain regularity assumptions. We also show that the divergence free part of the Arnold-Winther finite element space is connected to the Argyris finite element space by the Airy operator, which results in an orthogonal decomposition of the Arnold-Winther finite element space. This decomposition is, in fact, a discrete version of the decomposition introduced in Chapter II.

In this chapter, we assume that Ω is a polygon and Γ_D is connected. We always consider the points where Γ_D and Γ_T meet as corners of the polygon, even if the internal angers associated with these points may be π. Also, we assume that the compliance tensor $\mathbb{A}$ is uniformly positive definite in Ω and bounded above.

A. Mixed finite element method

As we have stated before, although Problem (2.13) is used in the implementation of mixed finite element methods, theoretical analysis only needs to be done on Problem (2.12), where the traction boundary condition satisfies $\boldsymbol{t}_0 = \mathbf{0}$. Let $(\mathbf{\Sigma}, \boldsymbol{V}) \subset (\boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,, \Omega, \mathbb{S}_2), \boldsymbol{L}^2(\Omega, \mathbb{R}^2))$ be a pair of finite dimensional subspaces. Problem (2.12) can be approximated by the following discrete problem: *Find* $\boldsymbol{\sigma} \in \mathbf{\Sigma}$ *and* $\boldsymbol{u} \in \boldsymbol{V}$ *such that*

$$\begin{cases} (\mathbb{A}\boldsymbol{\sigma}, \boldsymbol{\tau}) + (\mathbf{div}\,\boldsymbol{\tau}, \boldsymbol{u}) = < \boldsymbol{\tau n}, \boldsymbol{u}_0 >_{\Gamma_D} & \forall\, \boldsymbol{\tau} \in \mathbf{\Sigma}, \\ (\mathbf{div}\,\boldsymbol{\sigma}, \boldsymbol{v}) = (-\boldsymbol{f}, \boldsymbol{v}) & \forall\, \boldsymbol{v} \in \boldsymbol{V}. \end{cases} \tag{3.1}$$

As we shall see below, under certain conditions imposed on $(\boldsymbol{\Sigma}, \boldsymbol{V})$, the discrete problem (3.1) has a unique solution and the discrete solution is a good approximation to the weak solution of Problem (2.13). In the following, we will state these conditions.

Let $\boldsymbol{\Sigma}^*$ and $\boldsymbol{V}^*$ be the dual spaces of $\boldsymbol{\Sigma}$ and $\boldsymbol{V}$ respectively. Define linear operators $\mathcal{A} : \boldsymbol{\Sigma} \to \boldsymbol{\Sigma}^*$ and $\mathcal{B} : \boldsymbol{\Sigma} \to \boldsymbol{V}^*$ by

$$\begin{aligned} < \mathcal{A}\boldsymbol{\sigma}, \boldsymbol{\tau} > &= (\mathbb{A}\boldsymbol{\sigma}, \boldsymbol{\tau}) && \text{for all } \boldsymbol{\tau} \in \boldsymbol{\Sigma}, \\ < \mathcal{B}\boldsymbol{\sigma}, \boldsymbol{v} > &= (\mathbf{div}\, \boldsymbol{\sigma}, \boldsymbol{v}) && \text{for all } \boldsymbol{v} \in \boldsymbol{V}. \end{aligned} \tag{3.2}$$

Let $\mathcal{B}^t : \boldsymbol{V} \to \boldsymbol{\Sigma}^*$ be the adjoint of $\mathcal{B}$.

The discrete problem (3.1) can be written in the following operator form:

$$\begin{cases} \mathcal{A}\boldsymbol{\sigma} + \mathcal{B}^t\boldsymbol{u} = F & \text{in } \boldsymbol{\Sigma}^*, \\ \mathcal{B}\boldsymbol{\sigma} = G & \text{in } \boldsymbol{V}^*. \end{cases} \tag{3.3}$$

The functional G in (3.3) is defined by $(-\boldsymbol{f}, \cdot)$. Notice that for each $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega, \mathbb{R}^2)/Ker(\mathcal{B}^t)$,

$$G(\boldsymbol{w}) = (-\boldsymbol{f}, \boldsymbol{w}) = 0 \qquad \text{for all } \boldsymbol{w} \in Ker(\mathcal{B}^t).$$

Therefore, there exists a $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}$ such that

$$G(\boldsymbol{w}) = < \boldsymbol{\sigma}, \mathcal{B}^t\boldsymbol{w} > = < \mathcal{B}\boldsymbol{\sigma}, \boldsymbol{w} > \qquad \text{for all } \boldsymbol{w} \in \boldsymbol{V}.$$

This implies that G is in $Im(\mathcal{B})$ as long as $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega, \mathbb{R}^2)/Ker(\mathcal{B}^t)$.

Since we assumed that $\mathbb{A}$ is uniformly bounded above, there exists a positive number $\|\mathbb{A}\|$ such that

$$\begin{aligned} (\mathbb{A}\boldsymbol{\sigma}, \boldsymbol{\tau}) &\le \|\mathbb{A}\| \|\boldsymbol{\sigma}\|_{\boldsymbol{H}(\mathbf{div}, \Omega, \mathbb{S}_2)} \|\boldsymbol{\tau}\|_{\boldsymbol{H}(\mathbf{div}, \Omega, \mathbb{S}_2)} && \text{for all } \boldsymbol{\sigma}, \boldsymbol{\tau} \in \boldsymbol{\Sigma}, \\ (\mathbf{div}\, \boldsymbol{\sigma}, \boldsymbol{v}) &\le \|\boldsymbol{\sigma}\|_{\boldsymbol{H}(\mathbf{div}, \Omega, \mathbb{S}_2)} \|\boldsymbol{v}\|_{\boldsymbol{L}^2(\Omega, \mathbb{R}^2)} && \text{for all } \boldsymbol{\sigma} \in \boldsymbol{\Sigma},\ \boldsymbol{v} \in \boldsymbol{V}. \end{aligned}$$

In other words, both $\mathcal{A}$ and $\mathcal{B}$ are bounded operators.

We say that the pair of mixed finite element spaces $(\boldsymbol{\Sigma}, \boldsymbol{V})$ is stable if there exist positive constants C_A and C_B which only depend on Ω, Γ_T and $\mathbb{A}$ such that

$$(\mathbb{A}\boldsymbol{\sigma}, \boldsymbol{\sigma}) \geq C_A \|\boldsymbol{\sigma}\|^2_{\boldsymbol{H}(\mathbf{div}, \Omega, \mathbb{S}_2)} \qquad \text{for all } \boldsymbol{\sigma} \in Ker(\mathcal{B}), \tag{3.4}$$

$$\sup_{\boldsymbol{\tau} \in \boldsymbol{\Sigma}} \frac{(\mathbf{div}\, \boldsymbol{\tau}, \boldsymbol{v})}{\|\boldsymbol{\tau}\|_{\boldsymbol{H}(\mathbf{div}, \Omega, \mathbb{S}_2)}} \geq C_B \|\boldsymbol{v}\|_{\boldsymbol{L}^2(\Omega, \mathbb{R}^2)} \qquad \text{for all } \boldsymbol{v} \in \boldsymbol{V}/Ker(\mathcal{B}^t). \tag{3.5}$$

Inequality (3.5) is normally referred to as the discrete inf-sup condition. The importance of constructing a stable pair of finite element spaces is explained in the following theorems (see [22]).

Theorem III.1. *If* $(\boldsymbol{\Sigma}, \boldsymbol{V})$ *is stable and* $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega, \mathbb{R}^2)/Ker(\mathcal{B}^t)$*, then the discrete problem (3.1) has a unique solution in* $(\boldsymbol{\Sigma}, \boldsymbol{V}/Ker(\mathcal{B}^t))$*.*

If $(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{u}})$ is a solution of the weak problem (2.13) and $(\boldsymbol{\sigma}, \boldsymbol{u})$ is a solution of the discrete problem (3.1), we have the following error estimates [22]:

Theorem III.2. *If* $(\boldsymbol{\Sigma}, \boldsymbol{V})$ *is stable, then*

$$\begin{aligned} \|\hat{\boldsymbol{\sigma}} - \boldsymbol{\sigma}\|_{\boldsymbol{H}(\mathbf{div}, \Omega, \mathbb{S}_2)} \leq & \left(1 + \frac{\|\mathbb{A}\|}{C_A}\right)\left(1 + \frac{1}{C_B}\right) \inf_{\boldsymbol{\tau} \in \boldsymbol{\Sigma}} \|\hat{\boldsymbol{\sigma}} - \boldsymbol{\tau}\|_{\boldsymbol{H}(\mathbf{div}, \Omega, \mathbb{S}_2)} \\ & + \frac{1}{C_A} \inf_{\boldsymbol{v} \in \boldsymbol{V}} \|\hat{\boldsymbol{u}} - \boldsymbol{v}\|_{\boldsymbol{L}^2(\Omega, \mathbb{R}^2)}, \end{aligned}$$

and

$$\begin{aligned} \|\hat{\boldsymbol{u}} - \boldsymbol{u}\|_{\boldsymbol{L}^2(\Omega, \mathbb{R}^2)/Ker(\mathcal{B}^t)} \leq & \left(1 + \frac{1}{C_B}\right) \inf_{\boldsymbol{v} \in \boldsymbol{V}} \|\hat{\boldsymbol{u}} - \boldsymbol{v}\|_{\boldsymbol{L}^2(\Omega, \mathbb{R}^2)} \\ & + \frac{\|\mathbb{A}\|}{C_B} \|\hat{\boldsymbol{\sigma}} - \boldsymbol{\sigma}\|_{\boldsymbol{H}(\mathbf{div}, \Omega, \mathbb{S}_2)}. \end{aligned}$$

Furthermore, if $Ker(\mathcal{B}) \subset Ker(\mathbf{div})$*, then we have*

$$\|\hat{\boldsymbol{\sigma}} - \boldsymbol{\sigma}\|_{\boldsymbol{H}(\mathbf{div}, \Omega, \mathbb{S}_2)} \leq \left(1 + \frac{\|\mathbb{A}\|}{C_A}\right)\left(1 + \frac{1}{C_B}\right) \inf_{\boldsymbol{\tau} \in \boldsymbol{\Sigma}} \|\hat{\boldsymbol{\sigma}} - \boldsymbol{\tau}\|_{\boldsymbol{H}(\mathbf{div}, \Omega, \mathbb{S}_2)}.$$

For decades, extensive research has been done on developing stable pairs $(\boldsymbol{\Sigma}, \boldsymbol{V})$. According to the strong resemblance between system (3.1) and the mixed system for the Laplace equation, one may first think of using the lowest order Raviart-Thomas element RT_0 to construct the space $\mathbf{\Sigma}$, i.e.

$$\mathbf{\Sigma}^{RT_0}(\mathcal{T},\Omega,\Gamma_T) = \{\boldsymbol{\sigma} \in \boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,,\Omega,\mathbb{S}_2) \text{ such that } \boldsymbol{\sigma}|_T \in (RT_0)^2 \text{ for all } T \in \mathcal{T}\},$$

where

$$RT_0 = span\left\{\begin{pmatrix} a+cx \\ b+cy \end{pmatrix}\right\}.$$

However, notice that $\boldsymbol{\sigma} \in \mathbf{\Sigma}^{RT_0}(\mathcal{T},\Omega,\Gamma_T)$ has to be symmetric. Therefore, $\sigma_{12} = \sigma_{21}$ has to be a constant. This constraint results in the loss of approximating properties.

Indeed, it was very difficult to develop stable pairs of finite element spaces for linear elasticity problems because of the symmetry requirement on $\mathbf{\Sigma}$ (see Chapter VII of [22]). Stable finite elements on triangular meshes that have been developed include the Johnson-Mercier element [44], the Arnold-Douglas-Gupta element [4] and the Arnold-Winther element [8]. We will study, in detail, the Arnold-Winther element in Section B and briefly introduce other finite elements in Section D.

B. Arnold-Winther elements

Arnold and Winther proposed a new family of mixed finite elements for elasticity in [8]. First, we introduce the lowest order Arnold-Winther element.

Let $\mathcal{T}$ be a quasi-uniform triangulation of Ω with characteristic mesh size h which aligns with the corners of Ω. On each triangle $T \in \mathcal{T}$, define $P_i(T)$ to be the space consisting of polynomials of degree less than or equal to i. Let $P_i(T,\mathbb{S}_2) = (P_i(T))^3$ and $P_i(T,\mathbb{R}^2) = (P_i(T))^2$ be the spaces of polynomial tensors and polynomial vectors respectively. Define

$$\begin{aligned}\boldsymbol{\Sigma}_T &= \{\boldsymbol{\tau} \in P_3(T, \mathbb{S}_2) \text{ such that } \mathbf{div}\,\boldsymbol{\tau} \in P_1(T, \mathbb{R}^2)\},\\ \boldsymbol{V}_T &= P_1(T, \mathbb{R}^2).\end{aligned}$$

There are 24 degrees of freedom (dof) for $\boldsymbol{\Sigma}_T$ and 6 degrees of freedom for $\boldsymbol{V}_T$. The degrees of freedom for $\boldsymbol{\Sigma}_T$ are:

- the nodal values of the three components of $\boldsymbol{\tau}$ at each vertex of T (9 dofs);
- the moments of degree 0 and 1 of the two normal components of $\boldsymbol{\tau}$ on each edge of T (12 dofs);
- the moments of degree 0 of the three components of $\boldsymbol{\tau}$ on T (3 dofs).

The degrees of freedom of $\boldsymbol{V}_T$ are given as the zeroth and first order moments on T. Figure 2 illustrates the degrees of freedom for $\boldsymbol{\Sigma}_T$ and $\boldsymbol{V}_T$.

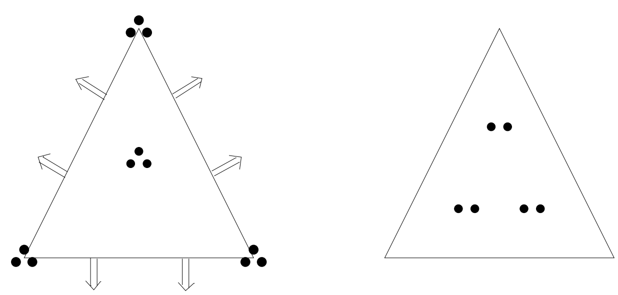

FIGURE 2. Finite elements $\boldsymbol{\Sigma}_T$ and $\boldsymbol{V}_T$.

It has been shown that the degrees of freedom defined above are unisolvent for $\boldsymbol{\Sigma}_T$ and $\boldsymbol{V}_T$ (see [8]). Furthermore, we clearly have $\mathbf{div}\,\boldsymbol{\Sigma}_T = \boldsymbol{V}_T$. Indeed, for $\boldsymbol{v} = (v_1, v_2)^t$ in $\boldsymbol{V}_T$, let $\boldsymbol{\sigma}$ be defined by $\frac{\partial}{\partial x}\sigma_{11} = v_1$, $\frac{\partial}{\partial y}\sigma_{22} = v_2$ and $\sigma_{12} = \sigma_{21} = 0$. Then $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_T$ and $\mathbf{div}\,\boldsymbol{\sigma} = \boldsymbol{v}$.

The Arnold-Winther element is affine under the matrix Piola transformation [8]. Let $\hat{T}$ be a reference triangle and J be the Jacobian matrix of the affine mapping from $\hat{T}$ to T. Define Π_T to be the nodal value interpolation from $C^2(T, \mathbb{S}_2)$ to $\mathbf{\Sigma}_T$ associated with the degrees of freedom for the Arnold-Winther element. For $\boldsymbol{\sigma} \in C^2(T, \mathbb{S}_2)$, define $\hat{\boldsymbol{\sigma}} \in C^2(\hat{T}, \mathbb{S}_2)$ by the matrix Piola transformation

$$\hat{\boldsymbol{\sigma}}(\hat{\boldsymbol{x}}) = J^{-1}\boldsymbol{\sigma}(\boldsymbol{x})J^{-t}. \tag{3.6}$$

Then

$$\Pi_T\boldsymbol{\sigma}(\boldsymbol{x}) = J\Pi_{\hat{T}}\hat{\boldsymbol{\sigma}}(\hat{\boldsymbol{x}})J^t. \tag{3.7}$$

Another important property is that ([8]), for $\boldsymbol{\sigma} \in C^2(T, \mathbb{S}_2)$,

$$\mathbf{div}\,\Pi_T\boldsymbol{\sigma} = \mathbf{P}_{\boldsymbol{V}_T}\mathbf{div}\,\boldsymbol{\sigma}, \tag{3.8}$$

where $\mathbf{P}_{\boldsymbol{V}_T}$ is the $\boldsymbol{L}^2$ projection onto $\boldsymbol{V}_T$.

The finite element spaces on the mesh $\mathcal{T}$ and domain Ω are defined as follows:

$$\begin{aligned} \mathbf{\Sigma}(\mathcal{T}, \Omega, \Gamma_T) = \{&\boldsymbol{\tau} \text{ defined on } \Omega \text{ satisfying } \boldsymbol{\tau}|_T \in \mathbf{\Sigma}_T \text{ for each } T \in \mathcal{T}, \\ &\boldsymbol{\tau} \text{ is continuous on the degrees of freedom on each vertex} \\ &\text{and each edge of } \mathcal{T} \text{ and } \boldsymbol{\tau}\boldsymbol{n}|_{\Gamma_T} = \mathbf{0}.\} \\ \boldsymbol{V}(\mathcal{T}, \Omega) = \{&\boldsymbol{v} \in \boldsymbol{L}_2(\Omega, \mathbb{R}^2) \text{ such that } \boldsymbol{v}|_T \in \boldsymbol{V}_T \text{ for each } T \in \mathcal{T}\}. \end{aligned}$$

The boundary condition $\boldsymbol{\tau}\boldsymbol{n}|_{\Gamma_T} = \mathbf{0}$ implies two linear relations among three components of $\boldsymbol{\tau}$ on boundary nodes. Hence on each corner of the polygon Ω where two traction boundary edges meet, we have $\boldsymbol{\tau} = \mathbf{0}$. This fact was noticed by Arnold and Winther in [8]. For simplicity, when $\Gamma_T = \emptyset$, we denote $\mathbf{\Sigma}(\mathcal{T}, \Omega, \Gamma_T)$ by $\mathbf{\Sigma}(\mathcal{T}, \Omega)$.

Notice that for $\boldsymbol{\tau} \in \mathbf{\Sigma}(\mathcal{T}, \Omega, \Gamma_T)$, $\boldsymbol{\tau}\boldsymbol{n}$ is continuous on the shared edge of two triangular elements. Therefore, by Lemma II.3,

$$\mathbf{\Sigma}(\mathcal{T},\Omega,\Gamma_T) \subset \boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,,\Omega,\mathbb{S}_2).$$

Another immediate observation is that

$$\mathbf{div}\,\mathbf{\Sigma}(\mathcal{T},\Omega,\Gamma_T) \subset \boldsymbol{V}(\mathcal{T},\Omega).$$

Remark III.1. *In the definition of* $\mathbf{\Sigma}(\mathcal{T},\Omega,\Gamma_T)$*, we require the continuity on each vertex. This does not seem to be natural for defining a subspace of* $\boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,,\Omega,\mathbb{S}_2)$*, in which nodal values are meaningless. A seemingly natural way to define the finite element space is*

$$\tilde{\mathbf{\Sigma}}(\mathcal{T},\Omega,\Gamma_T) = \{\boldsymbol{\tau} \in \boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,,\Omega,\mathbb{S}_2) \text{ such that } \boldsymbol{\tau}|_T \in \mathbf{\Sigma}_T \text{ for each } T \in \mathcal{T}\}.$$

The only continuity requirement for $\boldsymbol{\tau} \in \tilde{\mathbf{\Sigma}}(\mathcal{T},\Omega,\Gamma_T)$ *is that* $\boldsymbol{\tau n}$ *has to be continuous across each internal edge of* $\mathcal{T}$*. Notice that* $\mathbf{\Sigma}(\mathcal{T},\Omega,\Gamma_T)$ *and* $\tilde{\mathbf{\Sigma}}(\mathcal{T},\Omega,\Gamma_T)$ *may not be equal. For example, if one consider the clusters of triangles as shown in Figure 3, a function* $\boldsymbol{\tau} \in \tilde{\mathbf{\Sigma}}(\mathcal{T},\Omega,\Gamma_T)$ *does not necessarily need to have continuous nodal value on vertex v, although it has to be continuous on vertex w. Therefore, it is difficult to determine a basis for* $\tilde{\mathbf{\Sigma}}(\mathcal{T},\Omega,\Gamma_T)$*. This is the reason why we would like to explicitly define continuous nodal values in* $\mathbf{\Sigma}(\mathcal{T},\Omega,\Gamma_T)$*.*

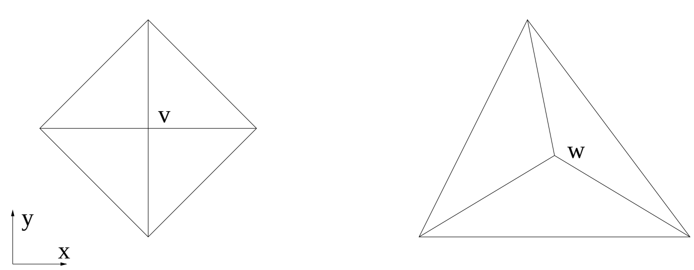

FIGURE 3. Clusters of triangles.

We need to show that under certain conditions, $(\boldsymbol{\Sigma}(\mathcal{T},\Omega,\Gamma_T),\boldsymbol{V}(\mathcal{T},\Omega))$ is stable. In [8], the stability of $(\boldsymbol{\Sigma}(\mathcal{T},\Omega),\boldsymbol{V}(\mathcal{T},\Omega))$ for the pure displacement boundary problem is proved. In the rest of this section, we generalize their proof to $(\boldsymbol{\Sigma}(\mathcal{T},\Omega,\Gamma_T),\boldsymbol{V}(\mathcal{T},\Omega))$ under the assumption of $\boldsymbol{H}^s$-regularity for $s > 1/2$.

To do this, we first construct an operator $\Pi_{\boldsymbol{\Sigma}} : \boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,,\Omega,\mathbb{S}_2)\cap\boldsymbol{H}^s(\Omega,\mathbb{S}_2) \to \boldsymbol{\Sigma}(\mathcal{T},\Omega,\Gamma_T)$ such that the following commutative diagram holds:

$$\begin{array}{ccc} \boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,,\Omega,\mathbb{S}_2)\cap\boldsymbol{H}^s(\Omega,\mathbb{S}_2) & \xrightarrow{\mathbf{div}} & \boldsymbol{L}^2(\Omega,\mathbb{R}^2) \\ \downarrow\Pi_{\boldsymbol{\Sigma}} & & \downarrow\mathbf{P}_{\boldsymbol{V}} \\ \boldsymbol{\Sigma}(\mathcal{T},\Omega,\Gamma_T) & \xrightarrow{\mathbf{div}} & \boldsymbol{V}(\mathcal{T},\Omega) \end{array} \tag{3.9}$$

where $\mathbf{P}_{\boldsymbol{V}} : \boldsymbol{L}^2(\Omega,\mathbb{R}^2) \to \boldsymbol{V}(\mathcal{T},\Omega)$ is the L^2 orthogonal projection and there exists a positive constant C_Π independent of the mesh size h such that for all $\boldsymbol{\tau} \in \boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,,\Omega,\mathbb{S}_2)\cap\boldsymbol{H}^s(\Omega,\mathbb{S}_2)$, $s > 1/2$,

$$\|\Pi_{\boldsymbol{\Sigma}}\boldsymbol{\tau}\|_{0,\Omega} \le C_\Pi\|\boldsymbol{\tau}\|_{s,\Omega}. \tag{3.10}$$

The moments of degree 0 and 1 of the two normal components on each edge are well defined for functions in $\boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,,\Omega,\mathbb{S}_2)\cap\boldsymbol{H}^s(\Omega,\mathbb{S}_2)$, where $s > 1/2$. Therefore, in the construction of $\Pi_{\boldsymbol{\Sigma}}$, we have no problem in dealing with the degrees of freedom on each edge. However, because of degrees of freedom on each vertex, the natural interpolation associated with the degrees of freedom is not bounded with respect to the norm $\|\cdot\|_{s,\Omega}$ for $s \le 1$. We need to consider a Clément type interpolation.

One important feature of the $\Pi_{\boldsymbol{\Sigma}}$ that we construct is that it preserves the essential boundary condition $\boldsymbol{\tau}\boldsymbol{n}|_{\Gamma_T} = \mathbf{0}$. Notice that τ_{ij} and $(\Pi_{\boldsymbol{\Sigma}}\boldsymbol{\tau})_{ij}$, $1 \le i,j \le 2$, need not be zero on Γ_T. Hence, we want $\Pi_{\boldsymbol{\Sigma}}$ to preserve the boundary condition in a natural way. For this purpose, we resort to the interpolation operator defined by Scott and Zhang in [50].

Consider the triangles in $\mathcal{T}$ as closed subsets of $\overline{\Omega}$ which contain their boundary. For each triangle $T \in \mathcal{T}$, define $S_T = \bigcup\{T_i | T_i \cap T \neq \emptyset, \quad T_i \in \mathcal{T}\}$. Let R_h be the interpolation operator from $H^s(\Omega)$, $s > 1/2$, onto the space of C^0-quadratics with respect to $\mathcal{T}$, as defined by Scott and Zhang [50]. We only need to pay attention that when we choose the integration simplex (see [50] or Appendix A for details) for boundary points where Γ_D and Γ_T meet, the simplex should be chosen to be a subset of Γ_T. The degrees of freedom for the C^0-quadratic element are nodal values on each vertex and the center point of each edge of $\mathcal{T}$. We call these points "nodes".

R_h is a linear interpolation satisfying (see [50] or Appendix A for the proof):

1. $R_h p = p$ for all C^0-quadratics p defined on mesh $\mathcal{T}$;

2. let $l \subset \Gamma_T$ be a boundary edge and $v \in H^s(\Omega)$ satisfy $v|_l = 0$, then $R_h v|_l = 0$;

3. (stability) for $1/2 < s \leq 1$, there exists a positive constant c independent of h such that

$$\|R_h v\|_{0,T} \leq c(\|v\|_{0,S_T} + h^s |v|_{s,S_T}) \quad \text{for all } v \in H^s(\Omega); \tag{3.11}$$

4. (approximability) for $1/2 < s \leq 3$ and $0 \leq t \leq min(s, 1)$, there exists a positive constant c independent of h such that

$$\|v - R_h v\|_{t,T} \leq ch^{s-t} |v|_{s,S_T} \quad \text{for all } v \in H^s(\Omega). \tag{3.12}$$

Define $\mathbf{R}_h$ which maps $\boldsymbol{H}^s(\Omega, \mathbb{S}_2)$, $s > 1/2$, to the space of symmetric tensors of C^0-quadratics with respect to the mesh $\mathcal{T}$ by

- on each corner $\boldsymbol{x}$ of the polygon Ω such that both of the boundary lines adjacent to $\boldsymbol{x}$ are in Γ_T,

$$\mathbf{R}_h(\boldsymbol{\tau})(\boldsymbol{x}) = \mathbf{0};$$

- on all the other nodes $\boldsymbol{x}$,

$$\mathbf{R}_h(\boldsymbol{\tau})(\boldsymbol{x}) = \begin{pmatrix} (R_h\tau_{11})(\boldsymbol{x}) & (R_h\tau_{12})(\boldsymbol{x}) \\ (R_h\tau_{21})(\boldsymbol{x}) & (R_h\tau_{22})(\boldsymbol{x}) \end{pmatrix}.$$

Lemma III.1. $\mathbf{R}_h$ *is an interpolation which satisfies*

1. *for* $\boldsymbol{\tau} \in \boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,, \Omega, \mathbb{S}_2) \cap \boldsymbol{H}^s(\Omega, \mathbb{S}_2)$, $s > 1/2$, *we have* $(\mathbf{R}_h\boldsymbol{\tau})\boldsymbol{n}|_{\Gamma_T} = \mathbf{0}$*;*
2. *(stability) for* $1/2 < s \leq 1$, *there exists a positive constant* c *independent of* h *such that*

$$\|\mathbf{R}_h\boldsymbol{\tau}\|_{0,T} \leq c(\|\boldsymbol{\tau}\|_{0,S_T} + h^s|\boldsymbol{\tau}|_{s,S_T}) \quad \textit{for all } \boldsymbol{\tau} \in \boldsymbol{H}^s(\Omega, \mathbb{S}_2); \tag{3.13}$$

3. *(approximability) for* $1/2 < s \leq 3$ *and* $0 \leq t \leq min(s, 1)$, *there exists a positive constant* c *independent of* h *such that*

$$\|\mathbf{R}_h\boldsymbol{\tau} - \boldsymbol{\tau}\|_{t,T} \leq ch^{s-t}|\boldsymbol{\tau}|_{s,S_T} \quad \textit{for all } \boldsymbol{\tau} \in \boldsymbol{H}^s(\Omega, \mathbb{S}_2) \cap \boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,, \Omega, \mathbb{S}_2). \tag{3.14}$$

Proof. First, we show that $\mathbf{R}_h$ preserves the essential boundary condition on Γ_T. Indeed, we only need to show that the boundary condition is preserved on all nodes on $\overline{\Gamma_T}$. We divide those nodes into three categories:

(I) $\boldsymbol{x}$ is inside a boundary line, or in other words, $\boldsymbol{x}$ is not a corner of the polygon Ω;

(II) $\boldsymbol{x}$ is a corner of the polygon Ω where Γ_D and Γ_T meets;

(III) $\boldsymbol{x}$ is a corner of the polygon Ω and the two boundary lines connected to $\boldsymbol{x}$ are both in Γ_T.

Let $\boldsymbol{\tau} \in \boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,, \Omega, \mathbb{S}_2) \cap \boldsymbol{H}^s(\Omega, \mathbb{S}_2)$. For $\boldsymbol{x}$ of type (I) and (II), $\boldsymbol{x}$ is either inside a boundary edge $l \subset \Gamma_T$ or is one end of a boundary edge $l \subset \Gamma_T$. Let $\boldsymbol{n} = (n_1, n_2)^t$

be the outward normal vector on l. Then $\boldsymbol{\tau n}|_l = \mathbf{0}$. Since R_h is linear, clearly

$$(\mathbf{R}_h\boldsymbol{\tau})\boldsymbol{n}|_{\boldsymbol{x}} = \begin{pmatrix} R_h\tau_{11} & R_h\tau_{12} \\ R_h\tau_{21} & R_h\tau_{22} \end{pmatrix} \begin{pmatrix} n_1 \\ n_2 \end{pmatrix}\Bigg|_{\boldsymbol{x}} = \begin{pmatrix} R_h(\tau_{11}n_1 + \tau_{12}n_2)|_{\boldsymbol{x}} \\ R_h(\tau_{21}n_1 + \tau_{22}n_2)|_{\boldsymbol{x}} \end{pmatrix} = \mathbf{0}.$$

For $\boldsymbol{x}$ of type (III), $\mathbf{R}_h\boldsymbol{\tau}(\boldsymbol{x})$ is forced to be $\mathbf{0}$ by definition. Therefore, the boundary condition is preserved.

The stability result (3.13) follows immediately from the definition of $\mathbf{R}_h$ and Inequality (3.11). We need to prove the approximability result (3.14). By the Bramble-Hilbert Lemma, there exists a $\boldsymbol{\rho} \in P_2(S_T, \mathbb{S}_2)$ (see Appendix A for details) such that

$$\|\boldsymbol{\tau} - \boldsymbol{\rho}\|_{s',S_T} \le ch^{s-s'}|\boldsymbol{\tau}|_{s,S_T}, \qquad 0 \le s' \le s \le 3,$$

where c is a positive constant independent of h and S_T. Hence by the triangle inequality, inverse inequality and the stability result (3.13), for $0 \le t \le min(s,1)$,

$$\begin{aligned} &\|\mathbf{R}_h\boldsymbol{\tau} - \boldsymbol{\tau}\|_{t,T} \\ \le\ & \|\boldsymbol{\tau} - \boldsymbol{\rho}\|_{t,T} + \|\mathbf{R}_h(\boldsymbol{\tau} - \boldsymbol{\rho})\|_{t,T} + \|\boldsymbol{\rho} - \mathbf{R}_h\boldsymbol{\rho}\|_{t,T} \\ \le\ & \|\boldsymbol{\tau} - \boldsymbol{\rho}\|_{t,T} + ch^{-t}(\|\boldsymbol{\tau} - \boldsymbol{\rho}\|_{0,S_T} + h^{min(s,1)}|\boldsymbol{\tau} - \boldsymbol{\rho}|_{min(s,1),S_T}) + \|\boldsymbol{\rho} - \mathbf{R}_h\boldsymbol{\rho}\|_{t,T} \\ \le\ & ch^{s-t}|\boldsymbol{\tau}|_{s,S_T} + \|\boldsymbol{\rho} - \mathbf{R}_h\boldsymbol{\rho}\|_{t,T}. \end{aligned}$$

By the definition of $\mathbf{R}_h$, $\boldsymbol{\rho} - \mathbf{R}_h\boldsymbol{\rho}$ has none-zero nodal values only at the corners of polygon Ω which connect two boundary edges in Γ_T. Denote V_c to be the set of such corner nodes. Then $\|\boldsymbol{\rho} - \mathbf{R}_h\boldsymbol{\rho}\|^2_{t,T} \le ch^{2(1-t)}\sum_{\boldsymbol{x}\in V_c\cap T}|\boldsymbol{\rho}(\boldsymbol{x})|^2$. Now we evaluate $|\boldsymbol{\rho}(\boldsymbol{x})|$ for each $\boldsymbol{x} \in V_c$. It is easy to see that $\boldsymbol{x}$ is the intersection of two edges γ_1, γ_2 of the mesh $\mathcal{T}_h$ and $\gamma_1, \gamma_2 \subset \Gamma_T \cap \partial S_T$. Denote $\boldsymbol{n}_1$, $\boldsymbol{n}_2$ to be the outward normal vectors on γ_1, γ_2 respectively. Figure 4 is an example of T and $\boldsymbol{x} \in V_c$. The polygon inside the thick line contour is S_T.

Notice that $\boldsymbol{n}_1 \ne \boldsymbol{n}_2$. This guarantees that $|\boldsymbol{\rho}(\boldsymbol{x})| \le c(|(\boldsymbol{\rho n}_1)(\boldsymbol{x})| + |(\boldsymbol{\rho n}_2)(\boldsymbol{x})|)$,

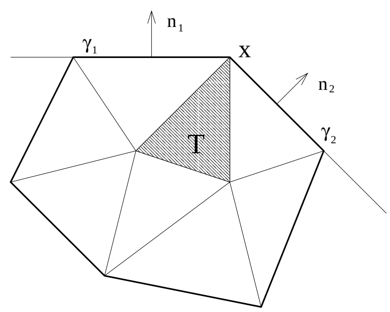

FIGURE 4. An example of T and $\boldsymbol{x} \in V_c$.

where c is independent of h, even if $\boldsymbol{n}_1$ or $\boldsymbol{n}_2$ is parallel to the x-axis or y-axis. By the boundary condition of $\boldsymbol{\tau}$ and the trace theorem,

$$\begin{aligned} h|\boldsymbol{\rho}(\boldsymbol{x})|^2 &\leq c\sum_{i=1}^{2} \|\boldsymbol{\rho}\boldsymbol{n}_i\|_{0,\gamma_i}^2 = c\sum_{i=1}^{2} \|(\boldsymbol{\tau}-\boldsymbol{\rho})\boldsymbol{n}_i\|_{0,\gamma_i}^2 \\ &\leq ch(h^{-2}\|(\boldsymbol{\tau}-\boldsymbol{\rho})\|_{0,S_T}^2 + h^{2min(s,1)-2}|(\boldsymbol{\tau}-\boldsymbol{\rho})|_{min(s,1),S_T}^2) \\ &\leq ch^{2s-1}|\boldsymbol{\tau}|_{s,S_T}^2. \end{aligned}$$

Hence $|\boldsymbol{\rho}(\boldsymbol{x})| \leq ch^{s-1}|\boldsymbol{\tau}|_{s,S_T}$ and consequently $\|\boldsymbol{\rho} - \mathbf{R}_h\boldsymbol{\rho}\|_{t,T} \leq ch^{s-t}|\boldsymbol{\tau}|_{s,S_T}$. This completes the proof of approximability for $\mathbf{R}_h$. □

Define $\Pi_{\boldsymbol{\Sigma}}^0 : \boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,, \Omega, \mathbb{S}_2) \cap \boldsymbol{H}^s(\Omega, \mathbb{S}_2) \to \boldsymbol{\Sigma}(\mathcal{T}, \Omega, \Gamma_T)$ by setting

(α) $(\Pi_{\boldsymbol{\Sigma}}^0\boldsymbol{\tau})(\boldsymbol{x}) = \mathbf{0}$ for all vertices $\boldsymbol{x}$ in $\mathcal{T}$;

(β) $\int_e (\Pi_{\boldsymbol{\Sigma}}^0\boldsymbol{\tau})\boldsymbol{n}\,ds = \int_e \boldsymbol{\tau}\boldsymbol{n}\,ds$ and $\int_e (\Pi_{\boldsymbol{\Sigma}}^0\boldsymbol{\tau})\boldsymbol{n}\,s\,ds = \int_e \boldsymbol{\tau}\boldsymbol{n}\,s\,ds$ for all edges e in $\mathcal{T}$;

(γ) $\int_T \Pi_{\boldsymbol{\Sigma}}^0\boldsymbol{\tau}\,d\boldsymbol{x} = \int_T \boldsymbol{\tau}\,d\boldsymbol{x}$ for all triangles T in $\mathcal{T}$.

Clearly, $\Pi_{\boldsymbol{\Sigma}}^0$ preserves the homogeneous essential boundary condition.

Lemma III.2. *For all* $\boldsymbol{\tau} \in \boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,, \Omega, \mathbb{S}_2) \cap \boldsymbol{H}^s(\Omega, \mathbb{S}_2)$, $1/2 < s \leq 1$, *we have*

$$\|\Pi_{\boldsymbol{\Sigma}}^0\boldsymbol{\tau}\|_{0,T} \leq c(\|\boldsymbol{\tau}\|_{0,T} + h^s\|\boldsymbol{\tau}\|_{s,T}) \quad \textit{for all } T \in \mathcal{T}, \tag{3.15}$$

where c is a positive constant independent of h and T.

Proof. Recall that the Arnold-Winther element is affine under the matrix Piola transformation (3.6). Let $\hat{T}$ be a reference triangle and J be the Jacobian matrix of the affine mapping from $\hat{T}$ to T. For $\boldsymbol{\tau} \in \boldsymbol{H}(\mathbf{div}\,, T, \mathbb{S}_2) \cap \boldsymbol{H}^s(T, \mathbb{S}_2)$, define $\hat{\boldsymbol{\tau}} \in \boldsymbol{H}(\mathbf{div}\,, \hat{T}, \mathbb{S}_2) \cap \boldsymbol{H}^s(\hat{T}, \mathbb{S}_2)$ by the matrix Piola transformation (3.6). Let Π^0_T be the restriction of $\Pi^0_{\boldsymbol{\Sigma}}$ to T. In [8], the authors have shown that

$$\Pi^0_T \boldsymbol{\tau}(\boldsymbol{x}) = J \Pi^0_{\hat{T}} \hat{\boldsymbol{\tau}}(\hat{\boldsymbol{x}}) J^t.$$

Notice that

$$\|\Pi^0_{\hat{T}} \hat{\boldsymbol{\tau}}\|_{0,\hat{T}} \leq c \|\hat{\boldsymbol{\tau}}\|_{s,\hat{T}} \quad \text{for all } \hat{\boldsymbol{\tau}} \in \boldsymbol{H}(\mathbf{div}\,, \hat{T}, \mathbb{S}_2) \cap \boldsymbol{H}^s(\hat{T}, \mathbb{S}_2),$$

where c is a constant depending only on $\hat{T}$. Inequality (3.15) follows from a standard scaling argument. □

Define

$$\Pi_{\boldsymbol{\Sigma}} = \Pi^0_{\boldsymbol{\Sigma}}(\mathbf{I} - \mathbf{R}_h) + \mathbf{R}_h = \Pi^0_{\boldsymbol{\Sigma}} + (\mathbf{I} - \Pi^0_{\boldsymbol{\Sigma}})\mathbf{R}_h.$$

Notice that for $\boldsymbol{\tau} \in \boldsymbol{H}_{0,\Gamma}(\mathbf{div}\,, \Omega, \mathbb{S}_2) \cap \boldsymbol{H}^s(\Omega, \mathbb{S}_2)$, $s > 1/2$, $(\mathbf{I} - \Pi^0_{\boldsymbol{\Sigma}})\mathbf{R}_h \boldsymbol{\tau}$ has nonzero degrees of freedom only on nodal values on each vertex. $\Pi_{\boldsymbol{\Sigma}}$ clearly preserves the boundary condition of $\boldsymbol{H}_{0,\Gamma}(\mathbf{div}\,, \Omega, \mathbb{S}_2)$ and we have the following lemma.

Lemma III.3. *For $s > 1/2$, $\Pi_{\boldsymbol{\Sigma}} : \boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,, \Omega, \mathbb{S}_2) \cap \boldsymbol{H}^s(\Omega, \mathbb{S}_2) \to \boldsymbol{\Sigma}(\mathcal{T}, \Omega, \Gamma_T)$ satisfies (3.9) and (3.10). Furthermore, we have the following approximation property*

$$\|\boldsymbol{\tau} - \Pi_{\boldsymbol{\Sigma}} \boldsymbol{\tau}\|_{0,\Omega} \leq c h^s \|\boldsymbol{\tau}\|_{s,\Omega} \quad \text{for } 1/2 < s \leq 3, \tag{3.16}$$

where c is a positive constant independent of h.

Proof. By Green's formula, for $\boldsymbol{\tau} \in \boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,, \Omega, \mathbb{S}_2) \cap \boldsymbol{H}^s(\Omega, \mathbb{S}_2)$, $\boldsymbol{v} \in \boldsymbol{V}_T$ and

$T \in \mathcal{T}$,

$$\begin{aligned}
&\int_T \mathbf{div}\,(\boldsymbol{\tau} - \Pi_{\boldsymbol{\Sigma}}\boldsymbol{\tau}) \cdot \boldsymbol{v}\, d\boldsymbol{x} \\
=& -\int_T (\boldsymbol{\tau} - \Pi_{\boldsymbol{\Sigma}}\boldsymbol{\tau}) : \nabla \boldsymbol{v}\, d\boldsymbol{x} + \int_{\partial T} (\boldsymbol{\tau} - \Pi_{\boldsymbol{\Sigma}}\boldsymbol{\tau})\boldsymbol{n} \cdot \boldsymbol{v}\, ds \\
=& -\int_T [(I - \Pi^0_{\boldsymbol{\Sigma}})(I - \mathbf{R}_h)\boldsymbol{\tau}] : \nabla \boldsymbol{v}\, d\boldsymbol{x} + \int_{\partial T} [(I - \Pi^0_{\boldsymbol{\Sigma}})(I - \mathbf{R}_h)\boldsymbol{\tau}]\boldsymbol{n} \cdot \boldsymbol{v}\, ds \\
=& 0.
\end{aligned}$$

The last equality followed from (β) and (γ) in the definition of $\Pi^0_{\boldsymbol{\Sigma}}$. Therefore (3.9) is true.

For $\boldsymbol{\tau} \in \boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,, \Omega, \mathbb{S}_2) \cap \boldsymbol{H}^s(\Omega, \mathbb{S}_2)$, $1/2 < s \le 3$, we have

$$\begin{aligned}
\|\boldsymbol{\tau} - \Pi_{\boldsymbol{\Sigma}}\boldsymbol{\tau}\|_{0,\Omega} &\le \|\boldsymbol{\tau} - \mathbf{R}_h\boldsymbol{\tau}\|_{0,\Omega} + \|\Pi^0_{\boldsymbol{\Sigma}}(\boldsymbol{\tau} - \mathbf{R}_h\boldsymbol{\tau})\|_{0,\Omega} \\
&\le c \sum_{T \in \mathcal{T}} (\|\boldsymbol{\tau} - \mathbf{R}_h\boldsymbol{\tau}\|_{0,T} + h^{min(s,1)}\|\boldsymbol{\tau} - \mathbf{R}_h\boldsymbol{\tau}\|_{min(s,1),T}) \\
&\le c \sum_{T \in \mathcal{T}} h^s |\boldsymbol{\tau}|_{s,S_T} \le c h^s |\boldsymbol{\tau}|_{s,\Omega}.
\end{aligned}$$

Notice that the last step of the above inequality is true because of the finite overlapping property of S_T when $\mathcal{T}$ is quasi-uniform. This proves (3.16) and consequently, (3.10) is true. $\square$

Remark III.2. *Another important consequence of the operator $\Pi_{\boldsymbol{\Sigma}}$ is that, by (3.2) we immediately have $Ker(\mathcal{B}^t) = \emptyset$ when $\Gamma_D \neq \emptyset$ and $Ker(\mathcal{B}^t) = RM$ when $\Gamma_D = \emptyset$. Therefore, the compatibility condition $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega, \mathbb{R}^2)/Ker(\mathcal{B}^t)$ in Theorem III.1 is consistent with the compatibility condition (2.9) for the weak problem.*

Now we can prove the following theorem:

Theorem III.3. *If the weak problem (2.12) has $\boldsymbol{H}^s$-regularity for $s > 1/2$, then the pair of finite element spaces $(\boldsymbol{\Sigma}(\mathcal{T}, \Omega, \Gamma_T), \boldsymbol{V}(\mathcal{T}, \Omega))$ is stable with C_A depending only on $\mathbb{A}$ and $C_B = 1/\sqrt{C_\Pi^2 C_R^2 + 1}$, where C_R is the constant in Inequality (2.14).*

Proof. Since $\mathbf{div}\,\boldsymbol{\Sigma}(\mathcal{T},\Omega,\Gamma_T) \subset \boldsymbol{V}(\mathcal{T},\Omega)$ and $\mathbb{A}$ is uniformly positive definite in Ω, clearly Inequality (3.4) is true with C_A depending only on $\mathbb{A}$.

Next we prove that Inequality (3.5) is true. Let $\boldsymbol{v} \in \boldsymbol{V}/Ker(\mathcal{B}^t)$. As stated in Remark III.2, we have $Ker(\mathcal{B}^t) = \emptyset$ when $\Gamma_D \neq \emptyset$ and $Ker(\mathcal{B}^t) = RM$ when $\Gamma_D = \emptyset$. Therefore, $\boldsymbol{v}$ satisfies the compatibility condition (2.9). By the regularity assumption, there exists a $\boldsymbol{\sigma} \in \boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,,\Omega,\mathbb{S}_2) \cap \boldsymbol{H}^s(\Omega,\mathbb{S}_2)$ such that $\mathbf{div}\,\boldsymbol{\sigma} = \boldsymbol{v}$ and $\|\boldsymbol{\sigma}\|_{s,\Omega} \le C_R\|\boldsymbol{v}\|_{0,\Omega}$, where $C_R > 0$ only depends on Ω, Γ_T and $\mathbb{A}$. By Lemma III.3,

$$\begin{aligned}
\|\Pi_{\boldsymbol{\Sigma}}\boldsymbol{\sigma}\|^2_{\boldsymbol{H}(\mathbf{div}\,,\Omega,\mathbb{S}_2)} &= \|\Pi_{\boldsymbol{\Sigma}}\boldsymbol{\sigma}\|^2_{0,\Omega} + \|\mathbf{P}_{\boldsymbol{V}}\mathbf{div}\,\boldsymbol{\sigma}\|^2_{0,\Omega}\\
&\le C_\Pi^2\|\boldsymbol{\sigma}\|^2_{s,\Omega} + \|\boldsymbol{v}\|^2_{0,\Omega}\\
&\le (C_\Pi^2 C_R^2 + 1)\|\boldsymbol{v}\|^2_{0,\Omega}.
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
\|\boldsymbol{v}\|_{0,\Omega} = \frac{(\boldsymbol{v},\boldsymbol{v})}{\|\boldsymbol{v}\|_{0,\Omega}} &= \frac{(\mathbf{P}_{\boldsymbol{V}}\mathbf{div}\,\boldsymbol{\sigma},\boldsymbol{v})}{\|\boldsymbol{v}\|_{0,\Omega}}\\
&\le \sqrt{C_\Pi^2 C_R^2 + 1}\,\frac{(\mathbf{div}\,\Pi_{\boldsymbol{\Sigma}}\boldsymbol{\sigma},\boldsymbol{v})}{\|\Pi_{\boldsymbol{\Sigma}}\boldsymbol{\sigma}\|_{\boldsymbol{H}(\mathbf{div}\,,\Omega,\mathbb{S}_2)}}\\
&\le \sqrt{C_\Pi^2 C_R^2 + 1}\sup_{\boldsymbol{\tau}\in\boldsymbol{\Sigma}(\mathcal{T},\Omega,\Gamma_T)}\frac{(\mathbf{div}\,\boldsymbol{\tau},\boldsymbol{v})}{\|\boldsymbol{\tau}\|_{\boldsymbol{H}(\mathbf{div}\,,\Omega)}}.
\end{aligned}$$

This completes the proof of the theorem. □

Remark III.3. *In [8], the authors proved the stability of* $(\boldsymbol{\Sigma}(\mathcal{T},\Omega),\boldsymbol{V}(\mathcal{T},\Omega))$. *The proof of Theorem III.3 followed their idea. The result in [8] can be viewed as a simplified case of Theorem III.3 in two aspects. First,* $\Pi_{\boldsymbol{\Sigma}}$ *does not need to preserve essential boundary conditions and hence a normal Clément type operator [26] can be used instead of a Scott-Zhang type operator. Second, for each* $\boldsymbol{v} \in \boldsymbol{L}^2(\Omega,\mathbb{R}^2)$, *there exists a* $\boldsymbol{\sigma} \in \boldsymbol{H}^1(\Omega,\mathbb{S}_2)$ *such that* $\mathbf{div}\,\boldsymbol{\sigma} = \boldsymbol{v}$ *and* $\|\boldsymbol{\sigma}\|_{1,\Omega} \le c\|\boldsymbol{v}\|_{0,\Omega}$, *where* c *is a positive constant independent of* $\boldsymbol{v}$. *Indeed, this* $\boldsymbol{\sigma}$ *can be easily obtained by solving an elasticity problem on a convex polygon containing* Ω *with* $\mathbf{div}\,\boldsymbol{\sigma} = \boldsymbol{v}$ *inside* Ω *and*

$\mathbf{div}\,\boldsymbol{\sigma} = \mathbf{0}$ *outside* Ω. *Therefore, no regularity assumption is required in this case. Notice that this has only been done when there is no essential boundary condition on* $\boldsymbol{\sigma}$.

Let $(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{u}})$ be a solution of the weak problem (2.13) and $(\boldsymbol{\sigma}, \boldsymbol{u})$ be a solution of the discrete problem (3.1). One can derive an error estimate by Theorem III.2. However, similar to Theorem 5.1 in [8], we have the following more precise error estimates:

Theorem III.4. *If the weak problem (2.13) has* $\boldsymbol{H}^s$*-regularity for* $1/2 < s \leq 3$, *then*

$$
\begin{aligned}
\|\hat{\boldsymbol{\sigma}} - \boldsymbol{\sigma}\|_{0,\Omega} &\leq ch^s \|\hat{\boldsymbol{\sigma}}\|_{s,\Omega}, \\
\|\mathbf{div}\,\hat{\boldsymbol{\sigma}} - \mathbf{div}\,\boldsymbol{\sigma}\|_{0,\Omega} &\leq ch^{\max(0,s-1)} \|\mathbf{div}\,\hat{\boldsymbol{\sigma}}\|_{\max(0,s-1),\Omega}, \\
\|\hat{\boldsymbol{u}} - \boldsymbol{u}\|_{0,\Omega} &\leq ch^{\min(s,2)} \|\hat{\boldsymbol{u}}\|_{\min(s,2)+1,\Omega},
\end{aligned}
$$

where c *is a general constant independent of* h.

$\square$

We have introduced the lowest order element of the family of Arnold-Winther elements. In general, the k-th order ($k \geq 1$) Arnold-Winther element is defined by

$$
\begin{aligned}
&\boldsymbol{\Sigma}_T^k = \{\text{symmetric tensors } \boldsymbol{\tau} \in P_{k+2}(T, \mathbb{S}_2) \text{ such that } \mathbf{div}\,\boldsymbol{\tau} \in P_k(T, \mathbb{R}^2)\}, \\
&\boldsymbol{V}_T^k = P_k(T, \mathbb{R}^2),
\end{aligned}
$$

with $dim(\boldsymbol{\Sigma}_T^k) = (3k^2 + 17k + 28)/2$ and $dim(\boldsymbol{V}_T^k) = (k+1)(k+2)$. The degrees of freedom for the k-th order $\boldsymbol{\Sigma}_T^k$ are:

- the nodal values of the three components of $\boldsymbol{\tau}$ at each vertex of T (9 dofs);
- the moments of degree $0, 1, \ldots, k$ of the two normal components of $\boldsymbol{\tau}$ on each edge of T ($6k + 6$ dofs);

- the moments $\int_T \boldsymbol{\tau} : \boldsymbol{\rho}\, d\boldsymbol{x}$ for all

$$\boldsymbol{\rho} \in \boldsymbol{\varepsilon}(P_k(T, \mathbb{R}^2)) + \mathbf{airy}\,(\lambda_1^2\lambda_2^2\lambda_3^2 P_{k-2}(T)),$$

  where $(\lambda_1, \lambda_2, \lambda_3)$ are the barycentric coordinate functions of T ($(3k^2+5k-2)/2$ dofs).

The degrees of freedom for $\boldsymbol{V}_T^k$ are the moments of degree $0, \ldots, k$ on T.

We would also like to mention a nonconforming version of the Arnold-Winther element recently introduced in [9]. Consider the lowest order Arnold-Winther element. Notice that the vertex degrees of freedom is unnatural for a subspace of $\boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,, \Omega, \Gamma_T)$, but unavoidable for defining conforming finite elements. However, one can drop the vertex degrees of freedom and define a nonconforming element. Define

$$\boldsymbol{\Sigma}_T^{NC} = \{\boldsymbol{\tau} \in P_2(T, \mathbb{S}_2) \text{ such that } \boldsymbol{n}^t\boldsymbol{\tau}\boldsymbol{n} \in P_1(e) \text{ for each edge } e \text{ of } T\}.$$

There are 15 degrees of freedom for $\boldsymbol{\Sigma}_T^{NC}$ which are exactly the degrees of freedom for $\boldsymbol{\Sigma}_T$ except for the nodal values. The space V_T is defined to be $P_1(T, \mathbb{R}^2)$. It has been proved in [9] that the nonconforming Arnold-Winther element is stable.

C. Airy operator on the discrete level

In this section, we show that on the discrete level, there exist exact sequences similar to the exact sequences in Lemma II.4 and Lemma II.7.

Denote Q_T to be the lowest order Argyris element [25] on T. It is a quintic element and the degrees of freedom are

- the nodal value on each vertex (3 dofs), the first derivatives at each vertex (6 dofs) and the second derivatives at each vertex (9 dofs);

- the moments of degree 0 of $\frac{\partial}{\partial \boldsymbol{n}} q$ on the edges of T (3 dofs).

Figure 5 illustrates the degrees of freedom for the Argyris element.

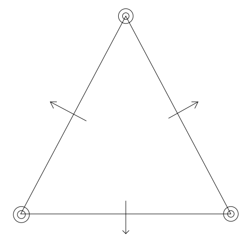

FIGURE 5. The Argyris finite element.

Notice that $\mathbf{airy}\, Q_T \subset \boldsymbol{\Sigma}_T$, $Ker(\mathbf{airy}\,) = P_1(T)$ and $\mathbf{div}\, \boldsymbol{\Sigma}_T = \boldsymbol{V}_T$. We immediately observe the following exact sequence by counting dimensions:

$$0 \to \; P_1(T) \xrightarrow{\subset} \mathrm{Q}_T \xrightarrow{\mathbf{airy}} \boldsymbol{\Sigma}_T \xrightarrow{\mathbf{div}} \boldsymbol{V}_T \; \to 0.$$

Define

$$\begin{aligned} \mathrm{Q}(\mathcal{T}, \Omega, \Gamma_T) = \{ & q \text{ defined on } \Omega \text{ satisfying } q|_T \in \mathrm{Q}_T \text{ for each } T \in \mathcal{T}, \\ & q \text{ is continuous on the degrees of freedom on each vertex} \\ & \text{and each edge of } \mathcal{T} \text{ and } q|_{\Gamma_T} = 0, \; \nabla q|_{\Gamma_T} = \mathbf{0} \}. \end{aligned}$$

It is not difficult to see that $\mathrm{Q}(\mathcal{T}, \Omega, \Gamma_T) \subset H^2_{0,\Gamma_T}(\Omega)$ and

$$\mathbf{airy}\, \mathrm{Q}(\mathcal{T}, \Omega, \Gamma_T) \subset \boldsymbol{\Sigma}(\mathcal{T}, \Omega, \Gamma_T).$$

Define

$$\tilde{P}_1^{\Gamma_T}(\Omega) = P_1(\Omega) \cap \mathrm{Q}(\mathcal{T}, \Omega, \Gamma_T).$$

Notice that $\tilde{P}_1^{\Gamma_T}(\Omega) = P_1(\Omega)$ when $\Gamma_T = \emptyset$ and $\tilde{P}_1^{\Gamma_T}(\Omega) = \emptyset$ when $\Gamma_T \neq \emptyset$.

Lemma III.4. *Assume that Problem (2.12) has $\boldsymbol{H}^s$-regularity, $s > 1/2$. Then the following exact sequences are true:*

1. *If $\Gamma_T \neq \partial\Omega$ and Γ_T is connected,*

$$\tilde{P}_1^{\Gamma_T}(\Omega) \xrightarrow{\subset} \mathrm{Q}(\mathcal{T},\Omega,\Gamma_T) \xrightarrow{\textbf{airy}} \boldsymbol{\Sigma}(\mathcal{T},\Omega,\Gamma_T) \xrightarrow{\textbf{div}} \boldsymbol{V}(\mathcal{T},\Omega) \to 0,$$

2. *If $\Gamma_T = \partial\Omega$*

$$0 \to \mathrm{Q}(\mathcal{T},\Omega,\Gamma_T) \xrightarrow{\textbf{airy}} \boldsymbol{\Sigma}(\mathcal{T},\Omega,\Gamma_T) \xrightarrow{\textbf{div}} \boldsymbol{V}(\mathcal{T},\Omega) \to RM \to 0.$$

Proof. First, by Theorem III.1 and Theorem III.3, $\textbf{div}$ maps $\boldsymbol{\Sigma}(\mathcal{T},\Omega,\Gamma_T)$ onto $\boldsymbol{V}(\mathcal{T},\Omega)$ when $\Gamma_T \neq \partial\Omega$ and $\textbf{div}$ maps $\boldsymbol{\Sigma}(\mathcal{T},\Omega,\partial\Omega)$ onto $\boldsymbol{V}(\mathcal{T},\Omega)/RM$.

Next, consider the divergence free part of $\boldsymbol{\Sigma}(\mathcal{T},\Omega,\Gamma_T)$. Clearly, $\textbf{airy}\,\mathrm{Q}(\mathcal{T},\Omega,\Gamma_T)$ is divergence free. Conversely, let $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}(\mathcal{T},\Omega,\Gamma_T)$ satisfies $\textbf{div}\,\boldsymbol{\sigma} = \mathbf{0}$. Similar to Section D in Chapter II, we can construct q such that $\textbf{airy}\,q = \boldsymbol{\tau}$ and it is elementary to show that q is a piece-wise quintic polynomial with continuous second order derivatives at the vertices of $\mathcal{T}$. Furthermore, following the proof of Lemma II.7, q can be chosen to satisfy the boundary conditions $q|_{\Gamma_T} = 0$ and $\nabla q|_{\Gamma_T} = \mathbf{0}$. Therefore, q is in $\mathrm{Q}(\mathcal{T},\Omega,\Gamma_T)$.

Finally, notice that $Ker(\textbf{airy}\,) = P_1(\Omega)$ and this implies the exactness condition on $\mathrm{Q}(\mathcal{T},\Omega,\Gamma_T)$. This completes the proof of this lemma. □

Lemma III.4 implies a decomposition of the discrete space $\boldsymbol{\Sigma}(\mathcal{T},\Omega,\Gamma_T)$:

$$\boldsymbol{\Sigma}(\mathcal{T},\Omega,\Gamma_T) = \mathcal{H}_0(\mathcal{T},\Omega,\Gamma_T) + \mathcal{H}_1(\mathcal{T},\Omega,\Gamma_T),$$

where

$$\mathcal{H}_0(\mathcal{T},\Omega,\Gamma_T) = Ker(\mathcal{B}) = \mathbf{airy}\,(\mathrm{Q}(\mathcal{T},\Omega,\Gamma_T)),$$

$$\mathcal{H}_1(\mathcal{T},\Omega,\Gamma_T) = \{\boldsymbol{\tau} \in \boldsymbol{\Sigma}(\mathcal{T},\Omega,\Gamma_T)\,|\,\boldsymbol{\tau} \perp \mathcal{H}_0(\mathcal{T},\Omega,\Gamma_T)$$

$$\text{in the } \boldsymbol{H}(\mathbf{div}\,,\Omega,\mathbb{S}_2) \text{ inner product}\}.$$

$\mathcal{H}_0(\mathcal{T},\Omega,\Gamma_T)$ and $\mathcal{H}_1(\mathcal{T},\Omega,\Gamma_T)$ are orthogonal in both the $\boldsymbol{H}(\mathbf{div}\,,\Omega,\mathbb{S}_2)$ inner product and the $\boldsymbol{L}^2$ inner product.

Clearly, $\mathbf{div}$ is a bijection from $\mathcal{H}_1(\mathcal{T},\Omega,\Gamma_T)$ onto $\boldsymbol{V}(\mathcal{T},\Omega)$ when $\Gamma_D \neq \emptyset$ or onto $\boldsymbol{V}(\mathcal{T},\Omega)/RM$ when $\Gamma_D = \emptyset$. An inverse operator $\mathbf{div}_{\mathcal{T}}^{-1}$ can be defined which maps

$$\begin{cases} \boldsymbol{V}(\mathcal{T},\Omega) \text{ onto } \mathcal{H}_1(\mathcal{T},\Omega,\Gamma_T), \text{ if } \Gamma_D \neq \emptyset; \\ \boldsymbol{V}(\mathcal{T},\Omega)/RM \text{ onto } \mathcal{H}_1(\mathcal{T},\Omega,\Gamma_T), \text{ if } \Gamma_D = \emptyset. \end{cases}$$

Let $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}(\mathcal{T},\Omega,\Gamma_T)$ and $\boldsymbol{u} \in \boldsymbol{V}(\mathcal{T},\Omega)$ satisfy

$$\begin{cases} (\boldsymbol{\sigma},\boldsymbol{\tau}) + (\mathbf{div}\,\boldsymbol{\tau},\boldsymbol{u}) = 0 \quad \text{for all } \boldsymbol{\tau} \in \boldsymbol{\Sigma}(\mathcal{T},\Omega,\Gamma_T), \\ (\mathbf{div}\,\boldsymbol{\sigma},\boldsymbol{w}) = (\boldsymbol{v},\boldsymbol{w}) \quad \text{for all } \boldsymbol{w} \in \boldsymbol{V}(\mathcal{T},\Omega). \end{cases} \tag{3.17}$$

Then $\mathbf{div}_{\mathcal{T}}^{-1}\boldsymbol{v}$ is defined to be $\boldsymbol{\sigma} \in \mathcal{H}_1(\mathcal{T},\Omega,\Gamma_T)$.

Therefore, for all $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}(\mathcal{T},\Omega,\Gamma_T)$, there exists a unique decomposition

$$\boldsymbol{\sigma} = \mathbf{airy}\,q + \mathbf{div}_{\mathcal{T}}^{-1}\boldsymbol{v},$$

where $q \in \mathrm{Q}(\mathcal{T},\Omega,\Gamma_T)/\tilde{P}_1^{\Gamma_T}(\Omega)$ and $\boldsymbol{v} = \mathbf{div}\,\boldsymbol{\sigma}$.

By Theorem III.4, we have the following lemma.

Lemma III.5. *Assume that Problem (2.12) has $\boldsymbol{H}^s$-regularity for $s > 1/2$. Then, for $\boldsymbol{v} \in \boldsymbol{V}(\mathcal{T},\Omega)$ when $\Gamma_D \neq \emptyset$ or $\boldsymbol{v} \in \boldsymbol{V}(\mathcal{T},\Omega)/RM$ when $\Gamma_D = \emptyset$,*

$$\|\mathbf{div}^{-1}\boldsymbol{v} - \mathbf{div}_{\mathcal{T}}^{-1}\boldsymbol{v}\|_{0,\Omega} \leq ch^s\|\boldsymbol{v}\|_{0,\Omega},$$

where c is a positive constant independent of h.

Finally, we would like to mention additional properties of $\mathcal{H}_0(\mathcal{T},\Omega,\Gamma_T)$, the divergence free part of $\boldsymbol{\Sigma}(\mathcal{T},\Omega,\Gamma_T)$. Let T be a triangle in $\mathcal{T}$ and $\boldsymbol{\tau} \in \boldsymbol{\Sigma}_T$. Let $\boldsymbol{x}_i$, $i = 1, 2, 3$ be the vertices of T and denote e_i to be the edge opposite to the vertex $\boldsymbol{x}_i$. Since $\operatorname{\mathbf{div}} \boldsymbol{\tau} \in \boldsymbol{V}_T = P_1(T, \mathbb{R}^2)$, we know that $\operatorname{\mathbf{div}} \boldsymbol{\tau} = \mathbf{0}$ if an only if

$$\int_T (\operatorname{\mathbf{div}} \boldsymbol{\tau}) \cdot \boldsymbol{v}\, d\boldsymbol{x} = 0 \quad \text{for all } \boldsymbol{v} \in P_1(T, \mathbb{R}^2).$$

By Green's formula and the fact that $\boldsymbol{\tau}$ is symmetric, we have

$$\int_T (\operatorname{\mathbf{div}} \boldsymbol{\tau}) \cdot \boldsymbol{v}\, d\boldsymbol{x} = \sum_{i=1}^{3} \int_{e_i} (\boldsymbol{\tau}\boldsymbol{n}) \cdot \boldsymbol{v}\, ds - \int_T \boldsymbol{\tau} : \boldsymbol{\varepsilon}(\boldsymbol{v})\, d\boldsymbol{x}.$$

Then, $\operatorname{\mathbf{div}} \boldsymbol{\tau} = \mathbf{0}$ if the right hand side of the above equation is zero for all $\boldsymbol{v} \in P_1(T, \mathbb{R}^2)$. Clearly, this is true if $\boldsymbol{\tau}$ has nonzero degrees of freedom only on nodal values on each vertex of T. Let $\boldsymbol{\tau}_i$, $i = 1, 2, 3$ be given constant symmetric tensors and define $\boldsymbol{\tau} \in \boldsymbol{\Sigma}_T$ by:

$$\begin{aligned}
\boldsymbol{\tau}(\boldsymbol{x}_i) &= \boldsymbol{\tau}_i && \text{for } i = 1, 2, 3,\\
\int_{e_i} \boldsymbol{\tau}\boldsymbol{n}\, ds = \int_{e_i} \boldsymbol{\tau}\boldsymbol{n} s\, ds &= \mathbf{0} && \text{for } i = 1, 2, 3,\\
\int_T \boldsymbol{\tau}\, d\boldsymbol{x} &= \mathbf{0}.
\end{aligned}$$

Then $\operatorname{\mathbf{div}} \boldsymbol{\tau} = \mathbf{0}$, which implies that there exists a $q \in Q_T$ such that $\boldsymbol{\tau} = \operatorname{\mathbf{airy}} q$. Define $q \in \mathrm{Q}_T$ by:

$$\begin{aligned}
\operatorname{\mathbf{airy}} q(\boldsymbol{x}_i) &= \boldsymbol{\tau}_i && \text{for } i = 1, 2, 3,\\
q(\boldsymbol{x}_i) = 0, \quad \nabla q(\boldsymbol{x}_i) &= \mathbf{0} && \text{for } i = 1, 2, 3,\\
\int_{e_i} \frac{\partial}{\partial \boldsymbol{n}} q\, ds &= 0 && \text{for } i = 1, 2, 3.
\end{aligned}$$

Note that setting $\operatorname{\mathbf{airy}} q(\boldsymbol{x}_i) = \boldsymbol{\tau}_i$ reduces to defining the second order derivative at the vertices. Hence q is defined by setting the values on each of the degrees of freedom.

Lemma III.6. *For $\boldsymbol{\tau}$ and q defined above, we have $\boldsymbol{\tau} = \mathbf{airy}\, q$.*

Proof. We will show that $\boldsymbol{\tau}$ and $\mathbf{airy}\, q$ are identical on all degrees of freedom. Notice that we already have $\mathbf{airy}\, q(\boldsymbol{x}_i) = \boldsymbol{\tau}(\boldsymbol{x}_i)$ for $i = 1, 2, 3$.

Let $\boldsymbol{n}$ and $\boldsymbol{s}$ denote the outward normal vector and the unit tangential vector on ∂T respectively. A simple calculation shows that

$$\begin{aligned} \frac{\partial^2}{\partial \boldsymbol{s}^2} q &= \boldsymbol{n}^t (\mathbf{airy}\, q) \boldsymbol{n}, \\ \frac{\partial^2}{\partial \boldsymbol{n} \partial \boldsymbol{s}} q &= -\boldsymbol{n}^t (\mathbf{airy}\, q) \boldsymbol{s}. \end{aligned}$$

By integration by parts, we have

$$\begin{aligned} \int_{e_i} \boldsymbol{n}^t (\mathbf{airy}\, q) \boldsymbol{n}\, ds &= \int_{e_i} \frac{\partial^2}{\partial \boldsymbol{s}^2} q\, ds = 0, \\ \int_{e_i} \boldsymbol{n}^t (\mathbf{airy}\, q) \boldsymbol{s}\, ds &= \int_{e_i} \frac{\partial^2}{\partial \boldsymbol{n} \partial \boldsymbol{s}} q\, ds = 0, \\ \int_{e_i} \boldsymbol{n}^t (\mathbf{airy}\, q) \boldsymbol{n}\, s\, ds &= \int_{e_i} \frac{\partial^2}{\partial \boldsymbol{s}^2} q\, s\, ds = -\int_{e_i} \frac{\partial}{\partial \boldsymbol{s}} q\, ds = 0, \\ \int_{e_i} \boldsymbol{n}^t (\mathbf{airy}\, q) \boldsymbol{s}\, s\, ds &= \int_{e_i} \frac{\partial^2}{\partial \boldsymbol{n} \partial \boldsymbol{s}} q\, s\, ds = -\int_{e_i} \frac{\partial}{\partial \boldsymbol{n}} q\, ds = 0. \end{aligned}$$

Notice that $\boldsymbol{n}$ and $\boldsymbol{s}$ are constants on each e_i. Therefore, a linear combination gives

$$\begin{aligned} \int_{e_i} (\mathbf{airy}\, q)\, \boldsymbol{n}\, ds &= \mathbf{0}, \\ \int_{e_i} (\mathbf{airy}\, q)\, \boldsymbol{n} s\, ds &= \mathbf{0}. \end{aligned}$$

Finally, let $\boldsymbol{n} = (n_1, n_2)^t$. Since $\int_{e_i} \frac{\partial}{\partial \boldsymbol{s}} q\, ds = 0$ and $\int_{e_i} \frac{\partial}{\partial \boldsymbol{n}} q\, ds = 0$ implies $\int_{e_i} \frac{\partial}{\partial x} q\, ds = \int_{e_i} \frac{\partial}{\partial y} q\, ds = 0$, by Green's formula,

$$\int_T \mathbf{airy}\, q\, d\boldsymbol{x} = \sum_{i=1}^{3} \int_{e_i} \begin{pmatrix} n_2 \frac{\partial}{\partial y} q & -n_1 \frac{\partial}{\partial y} q \\ -n_2 \frac{\partial}{\partial x} q & n_1 \frac{\partial}{\partial x} q \end{pmatrix} ds = \mathbf{0}.$$

Combining all the above, we have $\boldsymbol{\tau} = \mathbf{airy}\, q$. $\square$

D. Other mixed elements

In this section, we introduce several other mixed finite elements for linear plane elasticity. For simplicity, we only consider the pure displacement boundary problem, that is, $\Gamma_T = \emptyset$.

As we mentioned in Section A, it is not easy to construct stable pairs of finite element spaces for the linear plane elasticity problem. Brezzi and Fortin discussed this difficulty in detail in Chapter VII of [22]. They also point out two ways to solve this problem: one can relax the symmetry requirement for the stress by using Lagrangian functionals, or one can resort to composite elements.

For more on using the Lagrange multiplier to relax the symmetry constraint on the stress, we refer to [22, 3, 5]. Notice that in their formulations, the problem ends up with three variables: the stress, the displacement, and the Lagrange multiplier.

We turn to the second way mentioned above, the use of composite elements. The idea was introduced by Johnson and Mercier in [44]. Later Arnold, Douglas and Gupta developed another composite element in [4]. Next, we briefly introduce these two composite elements.

Divide each triangle $T \in \mathcal{T}$ into three sub-triangles T_1,T_2 and T_3 by connecting each vertex with the barycenter of T. The Johnson-Mercier elements on T are defined by

$$\boldsymbol{\Sigma}_T^{JM} = \{\boldsymbol{\tau} \in \boldsymbol{H}(\mathbf{div}\,, T, \mathbb{S}_2) \text{ such that } \boldsymbol{\tau}|_{T_i} \in P_1(T_i, \mathbb{S}_2), \text{ for } i = 1, 2, 3\},$$

$$\boldsymbol{V}_T^{JM} = \{\boldsymbol{v} \in P_1(T, \mathbb{R}^2)\}.$$

The degrees of freedom for $\boldsymbol{\Sigma}_T^{JM}$ are:

- the moments of degree 0 and 1 of the two normal components of $\boldsymbol{\tau}$ on each edge of T (12 dofs);

- the moments of degree 0 of the three components of $\boldsymbol{\tau}$ on T (3 dofs).

The degrees of freedom of $\boldsymbol{V}_T^{JM}$ are the zeroth and the first order moments on T. Figure 6 shows the degrees of freedom for the Johnson-Mercier element.

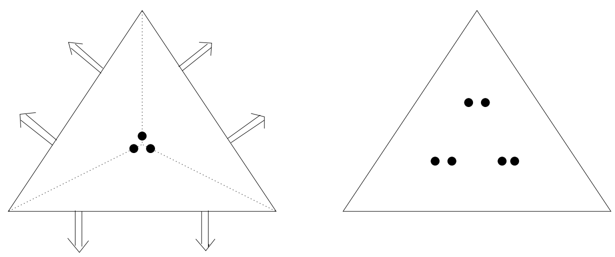

FIGURE 6. The Johnson-Mercier finite element.

It has been proved that the Johnson-Mercier element is stable [44]. Furthermore, the Johnson-Mercier element is related to the Hsieh-Clough-Tocher element [25] by the **airy** operator as the Arnold-Winther element is related to the Argyris element. However, notice that $\mathbf{div}\,\boldsymbol{\Sigma}_T^{JM} \not\subseteq \boldsymbol{V}_T^{JM}$, so we do not have the exact sequences as in Lemma III.4 or a commutative diagram as (3.9).

Next, we define the Arnold-Douglas-Gupta element. Define

$$\begin{aligned}\Xi_T^k = \{\boldsymbol{\tau} \in \boldsymbol{H}(\mathbf{div}\,, T, \mathbb{S}_2) \text{ such that } \boldsymbol{\tau}|_{T_i} \in P_k(T_i, \mathbb{S}_2),\ \text{for } i = 1, 2, 3 \\ \text{and } \mathbf{div}\,\boldsymbol{\tau} \in P_{k-1}(T, \mathbb{R}^2)\}.\end{aligned}$$

Let $\boldsymbol{\tau}^1$, $\boldsymbol{\tau}^2$ and $\boldsymbol{\tau}^3$ be three linearly independent elements in $\Xi_T^k/P_k(T, \mathbb{S}_2)$. The choice for $\boldsymbol{\tau}^i$ is arbitary and not unique. Then, the k-th order ($k \geq 2$) Arnold-Douglas-Gupta elements on T are defined by

$$\boldsymbol{\Sigma}_T^{ADG_k} = span\{P_k(T,\mathbb{S}_2), \boldsymbol{\tau}^1, \boldsymbol{\tau}^2, \boldsymbol{\tau}^3\},$$
$$\boldsymbol{V}_T^{ADG_k} = \{\boldsymbol{v} \in P_{k-1}(T,\mathbb{R}^2)\}.$$

The degrees of freedom for $\boldsymbol{\Sigma}_T^{ADG_k}$ are:

- the moments of degree $0, 1, \ldots, k$ of the two normal components of $\boldsymbol{\tau}$ on each edge of T ($6k + 6$ dofs);

- the moments $\int_T \boldsymbol{\tau} : \boldsymbol{\rho}\, d\boldsymbol{x}$ for all

$$\boldsymbol{\rho} \in \boldsymbol{\varepsilon}(P_{k-1}(T,\mathbb{R}^2)) + \mathbf{airy}\,(\lambda_1^2\lambda_2^2\lambda_3^2 P_{k-4}(T)),$$

  where $(\lambda_1, \lambda_2, \lambda_3)$ are the barycentric coordinate functions of T ($\frac{3}{2}k(k-1)$ dofs).

The degrees of freedom for $\boldsymbol{V}_T^k$ are the moments of degree $0, \ldots, k-1$ on T. Since $span\{\boldsymbol{\tau}^1, \boldsymbol{\tau}^2, \boldsymbol{\tau}^3\}$ is not uniquely determined, neither is the space $\boldsymbol{\Sigma}_T^{ADG_k}$. The authors of paper [4] showed that $\boldsymbol{\tau}^i$, $i = 1, 2, 3$ can be selected so that $\boldsymbol{\Sigma}_T^{ADG_k}$ is invariant under affine transformations of T onto itself.

Figure 7 shows the degrees of freedom for the lowest order ($k = 2$) Arnold-Douglas-Gupta elements, which has 21 degrees of freedom for the stress on each triangle:

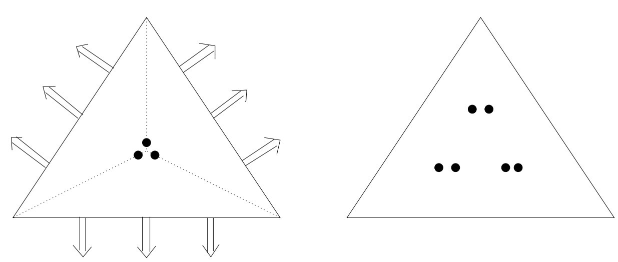

FIGURE 7. The Arnold-Douglas-Gupta finite element with $k = 2$.

Notice that $\mathbf{div}\,\boldsymbol{\Sigma}_T^{ADG_k} \subset \boldsymbol{V}_T^{ADG_k}$. In particular, a commutative diagram similar to (3.9) is true for the Arnold-Douglas-Gupta elements, where $s = 1$ and $\Pi_{\boldsymbol{\Sigma}}$ is the natural interpolation associated with the degrees of freedom. We refer to [4] for more details and results.

There are several other methods to discretize the mixed formulation for linear plane elasticity. One method which is frequently used employs the stabilizing technique [34]. A stabilizing term is added to the original formulation which can generalize the choice of finite element spaces for the stress and the displacement. Another method is to use least-square methods for the stress-displacement formulation [23, 24], where least-square functionals based either on the $\boldsymbol{L}^2$ norm or on the $\boldsymbol{H}^{-1}$ norm are defined.

CHAPTER IV

ITERATIVE SOLVERS FOR SADDLE POINT PROBLEMS

In this chapter, we discuss the iterative solvers for the linear systems which result from the mixed finite element discretization. First, we introduce the MINRES algorithm for solving symmetric indefinite linear problems and point out that the convergence rate of the MINRES method depends on the eigenvalue distribution of the linear system, which makes preconditioning important. Next, we discuss in detail the idea of preconditioning the mixed system by using norm equivalence. Using this idea, the problem of preconditioning the mixed system (3.1) can be reduced to the problem of preconditioning a $\boldsymbol{H}(\mathbf{div})$ problem in the symmetric matrix space, which will be the starting point of Chapter V and Chapter VI. Finally, several other iterative solvers are briefly introduced.

A. The preconditioned MINRES method

Let H be a n-dimensional real inner product space with inner product $(\cdot,\cdot)$ and corresponding norm $\|\cdot\|$. Let $M: H \to H$ be a linear operator which is symmetric with respect to $(\cdot,\cdot)$. M is not necessarily positive definite. Our purpose is to find a $u \in H$ such that

$$Mu = f, \tag{4.1}$$

for a given $f \in Im(M) \subset H$.

One way to solve Problem (4.1) is to use the minimum residual (MINRES) method [48]. Consider the Krylov subspace for $r \in H$:

$$K_m(M,r) = span\{r, Mr, \ldots, M^{m-1}r\}.$$

Given $u^0 \in H$, set $r^0 = f - Mu^0$. For $m = 1, 2, \ldots$, the mth step of the MINRES method calculates u^m which satisfies

$$\|f - Mu^m\| = \min_{v \in u^0 + K_m(M, r^0)} \|f - Mv\|.$$

Then u^m is an approximation to u. The minimization process using the Lanczos procedure leads to a 3-recurrence algorithm for the MINRES method:

Algorithm IV.1. *(MINRES) Given* $u^0 \in H/Ker(M)$*, set* $p^0 = r^0 = f - Mu^0$*. For* $i = 1, \ldots$ *until convergence, do*

1. $\alpha_i = (r^i, Mp^i)/(Mp^i, Mp^i)$;
2. $u^{i+1} = u^i + \alpha_i p^i$, $r^{i+1} = r^i - \alpha_i Mp^i$;
3. $\beta_i = (r^{i+1}, Mr^{i+1})/(r^i, Mr^i)$;
4. $p^{i+1} = r^{i+1} + \beta_i p^i$.

Notice that $Im(M) = H/Ker(M)$. The condition $u^0 \in H/Ker(M)$ guarantees that $u^i \in H/Ker(M)$ for each i.

Since M is a symmetric operator, all its eigenvalues are real. Assume that all the nonzero eigenvalues $\{\lambda_i\}_{i=1}^m$, $m \leq n$, are in the set

$$E = [a, b] \cup [c, d], \qquad \text{where } a < b < 0 < c < d.$$

Using the eigenvector expansion, we have

$$\|r^i\| \leq \delta_i(E)\|r^0\|,$$

where

$$\delta_i(E) = \inf_{p \in P_i} \sup_{t \in E} |p(t)|$$

and $P_i = \{$all polynomials p with degrees no more than i and satisfying $p(0) = 1\}$.

Let

$$\kappa(M) = \frac{max\{|a|, d\}}{min\{|b|, c\}}. \tag{4.2}$$

By using Chebyshev polynomials, one can derive an upper bound for $\delta_i(E)$:

$$\delta_i(E) \leq \inf_{\tilde{p} \in P_{[i/2]}} \sup_{t \in E} |\tilde{p}(t^2)| \leq 2 \left(\frac{\kappa(M) - 1}{\kappa(M) + 1} \right)^{[i/2]}.$$

Further discussion about the MINRES method and its convergence rate can be found in [33, 36].

Let $\{x_i\}_{i=1}^m$ be an orthonormal set of eigenfunctions of M corresponding to the non-zero eigenvalues $\{\lambda_i\}_{i=1}^m$ and spanning $Im(M)$. Let $M^\dagger$ be the pseudo-inverse of M defined by

$$M^\dagger u = \sum_{i=1}^m \lambda_i^{-1} (u, x_i) x_i. \tag{4.3}$$

It is clear that

$$\kappa(M) = \|M\| \|M^\dagger\| \leq \|M\|_* \|M^\dagger\|_*,$$

where the operator norms $\|\cdot\|$ and $\|\cdot\|_*$ are defined by

$$\|M\| = \sup_{\substack{v \in H \\ v \neq 0}} \frac{\|Mv\|}{\|v\|}, \qquad \|M\|_* = \sup_{\substack{v \in H \\ v \neq 0}} \frac{\|Mv\|_*}{\|v\|_*},$$

where $\|\cdot\|_*$ is any norm on H.

The error estimate indicates that MINRES converges relatively slow when $\kappa(M)$ is large. A solution to this problem is the preconditioned MINRES method. Let $S : H \to H$ be a linear operator which is symmetric with respect to $(\cdot, \cdot)$ and positive definite. Instead of solving system (4.1), one can solve

$$SMu = Sf.$$

Notice that $(S^{-1}\cdot, \cdot)$ is also an inner product on H and SM is symmetric with respect to $(S^{-1}\cdot, \cdot)$. We have the following preconditioned MINRES algorithm:

Algorithm IV.2. *(Preconditioned MINRES) Given* $u^0 \in H$*, set* $r^0 = f - Mu^0$*,* $p^0 = z^0 = Sr^0$*. For* $i = 1, \ldots$ *until convergence, do*

1. $\alpha_i = (z^i, Mp^i)/(Mp^i, SMp^i)$;
2. $u^{i+1} = u^i + \alpha_i p^i$, $r^{i+1} = r^i - \alpha_i M p^i$;
3. $z^{i+1} = z^i - \alpha_i SMp^i$;
4. $\beta_i = (z^{i+1}, Mz^{i+1})/(z^i, Mz^i)$;
5. $p^{i+1} = z^{i+1} + \beta_i p^i$.

It is clear that we want to choose S so that:

- the action of S on a vector is not too expensive;
- $\kappa(SM) \ll \kappa(M)$.

Finally, we mention the matrix form of the above algorithm. Let $\{\phi_i\}_{i=1}^n$ be a computational basis for H. All $u \in H$ can be uniquely written as $u = \sum_{i=1}^n \overline{u}_i \phi_i$, where $\overline{u}_i$, $1 \leq i \leq n$, are real numbers. Denote $\overline{u} = (\overline{u}_i)_{1 \leq i \leq n} \in \mathbb{R}^n$. Define the vector form of the right hand side $f \in H$ by

$$(\underline{f})_i = (f, \phi_i) \qquad \text{for } 1 \leq i \leq n. \tag{4.4}$$

Let $[\cdot, \cdot]$ be the Euclidean inner product on $\mathbb{R}^n$. Notice that $\underline{f} \neq \overline{f}$ and

$$(u, v) = [\overline{u}, \underline{v}] = [\underline{u}, \overline{v}].$$

Define the matrix forms of M and S by

$$\begin{aligned} (\underline{\underline{M}})_{ij} &= (M\,\phi_i, \phi_j) \qquad && \text{for } 1 \leq i, j \leq n, \\ ((\overline{\overline{S}})^{-1})_{ij} &= (S^{-1}\,\phi_i, \phi_j) \qquad && \text{for } 1 \leq i, j \leq n. \end{aligned}$$

Remark IV.1. *Let* $\{\psi_i\}_{i=1}^n$ *be a basis for* H *which is bi-orthonormal to* $\{\phi_i\}_{i=1}^n$, *that is,* $(\phi_i, \psi_j) = \delta_{ij}$, *where* δ_{ij} *is the Kronecker delta. Clearly,* $f = \sum_{i=1}^n \underline{f}_i \psi_i$. *Notice that* $\{(\psi_i, \cdot)\}_{i=1}^n$ *forms a basis for* H^*, *in other words,* $\underline{f}$ *is a vector representation of* $(f, \cdot) \in H^*$. *Let operators* $\mathcal{M} : H \to H^*$ *and* $\mathcal{S} : H^* \to H$ *satisfy*

$$< \mathcal{M}u, v > = (Mu, v) \qquad \text{for all } u, v \in H,$$

$$< \mathcal{S}^{-1}u, v > = (S^{-1}u, v) \qquad \text{for all } u, v \in H.$$

It is not hard to see that $\underline{\underline{M}}$ *is the matrix representation of* $\mathcal{M}$ *and* $\overline{\overline{S}}$ *is the matrix representation of* $\mathcal{S}$.

Normally, in real applications, instead of computing $\overline{\overline{S}}$, *one only computes the action of* $\overline{\overline{S}}$ *on a vector* $\underline{f}$ *by* $\overline{\overline{S}}\underline{f} = \overline{(S\,f)}$, *which is equivalent to the action of* S *on* f *or the action of* $\mathcal{S}$ *on* $(f, \cdot) \in H^*$.

Clearly,

$$\underline{\underline{M}}\,\overline{u} = \underline{(Mu)} \qquad \text{for all } u \in H,$$

$$\overline{\overline{S}}\,\underline{\underline{M}}\,\overline{u} = \overline{(S\,M\,u)} \qquad \text{for all } u \in H.$$

Problem (4.1) can be rewritten as

$$\overline{\overline{S}}\,\underline{\underline{M}}\,\overline{u} = \overline{\overline{S}}\,\underline{f}.$$

The matrix form of preconditioned MINRES algorithm is as follows:

Algorithm IV.3. *(Preconditioned MINRES in matrix form) Given* $\overline{u}^0 \in \mathbb{R}^n$, *set* $\underline{r}^0 = \underline{f} - \underline{\underline{M}}\,\overline{u}^0$, $\overline{p}^0 = \overline{z}^0 = \overline{\overline{S}}\,\underline{r}^0$. *For* $i = 1, \ldots$ *until convergence, do*

1. $\alpha_i = [\overline{z}^i, \underline{\underline{M}}\,\overline{p}^i]/[\underline{\underline{M}}\,\overline{p}^i, \overline{\overline{S}}\,\underline{\underline{M}}\,\overline{p}^i]$;
2. $\overline{u}^{i+1} = \overline{u}^i + \alpha_i \overline{p}^i$, $\underline{r}^{i+1} = \underline{r}^i - \alpha_i \underline{\underline{M}}\,\overline{p}^i$;
3. $\overline{z}^{i+1} = \overline{z}^i - \alpha_i \overline{\overline{S}}\,\underline{\underline{M}}\,\overline{p}^i$;
4. $\beta_i = [\overline{z}^{i+1}, \underline{\underline{M}}\,\overline{z}^{i+1}]/[\overline{z}^i, \underline{\underline{M}}\,\overline{z}^i]$;

5. $\overline{p}^{i+1} = \overline{z}^{i+1} + \beta_i \overline{p}^i$.

It is easy to see that Algorithm IV.3 is equivalent to Algorithm IV.2. Therefore, they have the same convergence rate estimate, which depends on $\kappa(SM)$.

Remark IV.2. *Another implementation of MINRES method is based on using the Lanczos procedure to tridiagonalize the matrix (see [48, 36] for details). It is mathematically equivalent to the MINRES algorithm stated here, although some authors believe that it is more stable.*

B. Preconditioning the saddle-point problem using norm equivalence

Assume that H_1 and H_2 are two Hilbert spaces with norms $\|\cdot\|_{H_1}$ and $\|\cdot\|_{H_2}$. Denote $\mathcal{L}(H_1, H_2)$ to be the set of all linear operators which map H_1 to H_2. For $\mathcal{O} \in \mathcal{L}(H_1, H_2)$, define operator norm

$$\|\mathcal{O}\| = \sup_{\substack{x \in H_1 \\ x \neq 0}} \frac{\|\mathcal{O}x\|_{H_2}}{\|x\|_{H_1}}.$$

To be explicit, we consider system (3.3). Assume that we have $\boldsymbol{H}^s$-regularity, where $s > 1/2$. For simplicity, denote $\boldsymbol{\Sigma}(\mathcal{T}, \Omega, \Gamma_T)$ by $\boldsymbol{\Sigma}$ and $\boldsymbol{V}(\mathcal{T}, \Omega)$ by $\boldsymbol{V}$. Let $\|\cdot\|_{\boldsymbol{\Sigma}}$ be the $\boldsymbol{H}(\mathbf{div}, \Omega, \mathbb{S}_2)$ norm on $\boldsymbol{\Sigma}$ and $\|\cdot\|_{\boldsymbol{\Sigma}^*}$ be its dual norm. Denote $\|\cdot\|_{\boldsymbol{V}}$ to be the $\boldsymbol{L}^2(\Omega, \mathbb{R}^2)$ norm on $\boldsymbol{V}$ and $\|\cdot\|_{\boldsymbol{V}^*}$ to be its dual norm. System (3.3) can be written as

$$\mathcal{M}\begin{pmatrix} \boldsymbol{\sigma} \\ \boldsymbol{u} \end{pmatrix} = \begin{pmatrix} \mathcal{A} & \mathcal{B}^t \\ \mathcal{B} & 0 \end{pmatrix}\begin{pmatrix} \boldsymbol{\sigma} \\ \boldsymbol{u} \end{pmatrix} = \begin{pmatrix} F \\ G \end{pmatrix}, \tag{4.5}$$

where $F \in \boldsymbol{\Sigma}^*$, $G \in Im(\mathcal{B}) \subset \boldsymbol{V}^*$ and $\mathcal{M} \in \mathcal{L}(\boldsymbol{\Sigma} \times \boldsymbol{V}, \boldsymbol{\Sigma}^* \times \boldsymbol{V}^*)$.

Lemma IV.1. *Let $F \in \boldsymbol{\Sigma}^*$ and $G \in Im(\mathcal{B})$. Let $(\boldsymbol{\sigma}, \boldsymbol{u})$ be a solution of the equation*

(4.5), then

$$c_0(\|F\|_{\mathbf{\Sigma}^*} + \|G\|_{\boldsymbol{V}^*}) \le \|\boldsymbol{\sigma}\|_{\mathbf{\Sigma}} + \|\boldsymbol{u}\|_{\boldsymbol{V}/Ker(\mathcal{B}^t)} \le c_1(\|F\|_{\mathbf{\Sigma}^*} + \|G\|_{\boldsymbol{V}^*}),$$

where c_0 and c_1 are positive and independent of the mesh size h.

Proof. By the stability of the finite elements spaces $(\mathbf{\Sigma}, \boldsymbol{V})$ and Proposition 1.3 in [22],

$$\|\boldsymbol{\sigma}\|_{\mathbf{\Sigma}} + \|\boldsymbol{u}\|_{\boldsymbol{V}/Ker(\mathcal{B}^t)} \le c_1(\|F\|_{\mathbf{\Sigma}^*} + \|G\|_{\boldsymbol{V}^*}),$$

where c_1 is independent of h. By the Schwartz inequality, we have

$$\begin{aligned}
\|F\|_{\mathbf{\Sigma}^*} + \|G\|_{\boldsymbol{V}^*} &= \sup_{\boldsymbol{\tau}\in\mathbf{\Sigma}} \frac{F(\boldsymbol{\tau})}{\|\boldsymbol{\tau}\|_{\mathbf{\Sigma}}} + \sup_{\boldsymbol{v}\in\boldsymbol{V}} \frac{G(\boldsymbol{v})}{\|\boldsymbol{v}\|_{\boldsymbol{V}}} \\
&= \sup_{\boldsymbol{\tau}\in\mathbf{\Sigma}} \frac{(\mathbb{A}\boldsymbol{\sigma}, \boldsymbol{\tau}) + (\mathbf{div}\,\boldsymbol{\tau}, \boldsymbol{u})}{\|\boldsymbol{\tau}\|_{\mathbf{\Sigma}}} + \sup_{\boldsymbol{v}\in\boldsymbol{V}} \frac{(\mathbf{div}\,\boldsymbol{\sigma}, \boldsymbol{v})}{\|\boldsymbol{v}\|_{\boldsymbol{V}}} \\
&\le c(\|\boldsymbol{\sigma}\|_{\mathbf{\Sigma}} + \|\boldsymbol{u}\|_{\boldsymbol{V}}),
\end{aligned}$$

where c is independent of h. Notice that the above inequality is true for $\boldsymbol{u}+\boldsymbol{w}$, where $\boldsymbol{w}$ is any element in $Ker(\mathcal{B}^t)$. Therefore

$$\|F\|_{\mathbf{\Sigma}^*} + \|G\|_{\boldsymbol{V}^*} \le c(\|\boldsymbol{\sigma}\|_{\mathbf{\Sigma}} + \|\boldsymbol{u}\|_{\boldsymbol{V}/Ker(\mathcal{B}^t)}).$$

$\square$

Define the $\boldsymbol{L}^2$ inner product over $\mathbf{\Sigma} \times \boldsymbol{V}$ by

$$\left(\begin{pmatrix} \boldsymbol{\sigma} \\ \boldsymbol{u} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\tau} \\ \boldsymbol{v} \end{pmatrix} \right) = (\boldsymbol{\sigma}, \boldsymbol{\tau}) + (\boldsymbol{u}, \boldsymbol{v}) \quad \text{for all } \boldsymbol{\sigma}, \boldsymbol{\tau} \in \mathbf{\Sigma} \text{ and } \boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{V}.$$

Let $\iota_1 : \mathbf{\Sigma}^* \to \mathbf{\Sigma}$ and $\iota_2 : \boldsymbol{V}^* \to \boldsymbol{V}$ be defined by

$$\begin{aligned}
(\iota_1(F), \boldsymbol{\tau}) &= F(\boldsymbol{\tau}) \qquad && \text{for all } \boldsymbol{\tau} \in \mathbf{\Sigma}, \\
(\iota_2(G), \boldsymbol{v}) &= G(\boldsymbol{v}) \qquad && \text{for all } \boldsymbol{v} \in \boldsymbol{V}.
\end{aligned} \tag{4.6}$$

Denote $\iota \in \mathcal{L}(\mathbf{\Sigma}^* \times \boldsymbol{V}^*, \mathbf{\Sigma} \times \boldsymbol{V})$ to be

$$\iota = \begin{pmatrix} \iota_1 & 0 \\ 0 & \iota_2 \end{pmatrix}.$$

Since $\mathbf{\Sigma}^* \times \boldsymbol{V}^*$ is a finite dimensional space, ι is invertible. Define an operator $M \in \mathcal{L}(\mathbf{\Sigma} \times \boldsymbol{V}, \mathbf{\Sigma} \times \boldsymbol{V})$ by $M = \iota \circ \mathcal{M}$. By (3.2), M can be written as

$$M = \begin{pmatrix} \iota_1 \circ \mathcal{A} & \iota_1 \circ \mathcal{B}^t \\ \iota_2 \circ \mathcal{B} & 0 \end{pmatrix} = \begin{pmatrix} A & B^t \\ B & 0 \end{pmatrix},$$

where $A : \mathbf{\Sigma} \to \mathbf{\Sigma}$ and $B : \mathbf{\Sigma} \to \boldsymbol{V}$ satisfy

$$\begin{aligned} (A\boldsymbol{\sigma}, \boldsymbol{\tau}) &= (\mathbb{A}\boldsymbol{\sigma}, \boldsymbol{\tau}) && \text{for all } \boldsymbol{\sigma}, \boldsymbol{\tau} \in \mathbf{\Sigma}, \\ (B\boldsymbol{\sigma}, \boldsymbol{v}) &= (\mathbf{div}\, \boldsymbol{\sigma}, \boldsymbol{v}) && \text{for all } \boldsymbol{\sigma} \in \mathbf{\Sigma} \text{ and } \boldsymbol{v} \in \boldsymbol{V}. \end{aligned}$$

System (4.5) is equivalent to

$$M \begin{pmatrix} \boldsymbol{\sigma} \\ \boldsymbol{u} \end{pmatrix} = \begin{pmatrix} \iota_1(F) \\ \iota_2(G) \end{pmatrix}. \tag{4.7}$$

Clearly, M is symmetric with respect to the $\boldsymbol{L}^2$ inner product over $\mathbf{\Sigma} \times \boldsymbol{V}$. Also, notice that the computing of the matrix form $\underline{\underline{M}}$ is straight forward since each one of $(A\cdot, \cdot)$, $(B\cdot, \cdot)$ and $(B^t\cdot, \cdot)$ is computable. By (4.4) and (4.6), the vector form of the right hand side $\underline{\iota_1(F)}$ and $\underline{\iota_2(G)}$ can be computed by $F(\cdot)$ and $G(\cdot)$. Therefore, we can use the MINRES algorithm defined in Section A to solve Problem (4.7).

The convergence rate of MINRES is determined by $\kappa(M)$. One can show that $\kappa(M)$ depends on h. Let μ_i, $i = 1, \ldots, m$, be the eigenvalues of A in ascending order and let σ_i, $i = 1, \ldots, k$, be the non-zero singular values of B in ascending order. It is clear that μ_i, $i = 1, \ldots, m$, are of order $O(1)$, σ_1 is of order $O(1)$ and σ_k is of order $O(h^{-1})$. A detailed analysis of $\kappa(M)$ is given in [49], which states that the non-zero eigenvalues of M are in $[a,b]\cap[c,d]$ where $a<b<0<c<d$ and

$$a \ge \frac{1}{2}\left(\mu_1 - \sqrt{\mu_1^2+4\sigma_k^2}\right), \qquad b \le \frac{1}{2}\left(\mu_m - \sqrt{\mu_m^2+4\sigma_1^2}\right),$$
$$c \ge \mu_1, \qquad d \le \frac{1}{2}\left(\mu_m - \sqrt{\mu_m^2+4\sigma_k^2}\right).$$

Combining all the above gives that $\kappa(M)$ is at least of order $O(h^{-1})$. To get an efficient algorithm, we need to find a preconditioner $S:\ \boldsymbol{\Sigma}\times\boldsymbol{V}\to\boldsymbol{\Sigma}\times\boldsymbol{V}$, which is symmetric with respect to the $\boldsymbol{L}^2$ inner product and positive definite, such that $\kappa(SM)$ is bounded by a positive number independent of the mesh size h.

Let $M^\dagger$ be the pseudo-inverse of M defined by (4.3). Then $\mathcal{M}^\dagger = M^\dagger\circ\iota$ is a pseudo-inverse of $\mathcal{M}$. By Lemma IV.1, we have $\|\mathcal{M}\|\le c_1$ and $\|\mathcal{M}^\dagger\|\le 1/c_0$. Define a linear operator $\mathcal{S}:\ \boldsymbol{\Sigma}^*\times\boldsymbol{V}^*\to\boldsymbol{\Sigma}\times\boldsymbol{V}$ by $\mathcal{S}=S\circ\iota$. Since SM is symmetric under $(S^{-1}\cdot,\cdot)$,

$$\begin{aligned}\kappa(SM) &\le \|SM\|\|M^\dagger S^{-1}\|\\ &= \|\mathcal{S}\mathcal{M}\|\|\mathcal{M}^\dagger\mathcal{S}^{-1}\|\\ &\le \|\mathcal{S}\|\|\mathcal{S}^{-1}\|\|\mathcal{M}\|\|\mathcal{M}^\dagger\|\\ &\le \frac{c_1}{c_0}\|\mathcal{S}\|\|\mathcal{S}^{-1}\|.\end{aligned} \tag{4.8}$$

Therefore, as long as both $\|\mathcal{S}\|$ and $\|\mathcal{S}^{-1}\|$ are bounded uniformly in h, $\kappa(SM)$ will be independent of h.

Consider those S in the form

$$S = \begin{pmatrix} S_1 & 0\\ 0 & S_2\end{pmatrix}, \tag{4.9}$$

where $S_1:\ \boldsymbol{\Sigma}\to\boldsymbol{\Sigma}$ and $S_2:\ \boldsymbol{V}\to\boldsymbol{V}$ are linear operators. Then

$$\mathcal{S} = \begin{pmatrix} S_1\circ\iota_1 & 0\\ 0 & S_2\circ\iota_2\end{pmatrix} = \begin{pmatrix}\mathcal{S}_1 & 0\\ 0 & \mathcal{S}_2\end{pmatrix}.$$

We want to find out conditions on S_1 and S_2 such that both $\mathcal{S}_1$, $\mathcal{S}_2$ and their inverses are bounded uniformly in h.

Define a bilinear form on $\boldsymbol{\Sigma}$ as follows:

$$\boldsymbol{\Lambda}(\boldsymbol{\sigma}, \boldsymbol{\tau}) = (\boldsymbol{\sigma}, \boldsymbol{\tau}) + (\mathbf{div}\, \boldsymbol{\sigma}, \mathbf{div}\, \boldsymbol{\tau}), \tag{4.10}$$

which is equal to to the $\boldsymbol{H}(\mathbf{div}\,, \Omega, \mathbb{S}_2)$ inner product. Use the same notation $\boldsymbol{\Lambda}$ to denote an operator induced by the bilinear form $\boldsymbol{\Lambda}$:

$$(\boldsymbol{\Lambda}\boldsymbol{\sigma}, \boldsymbol{\tau}) = \boldsymbol{\Lambda}(\boldsymbol{\sigma}, \boldsymbol{\tau}) \quad \text{for all } \boldsymbol{\sigma}, \boldsymbol{\tau} \in \boldsymbol{\Sigma}.$$

Here, we use the same notation for the bilinear form and the induced operator and it should cause no ambiguity. Consider the following problems: *for $F \in \boldsymbol{\Sigma}^*$, find $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}$ such that*

$$\boldsymbol{\Lambda}(\boldsymbol{\sigma}, \boldsymbol{\tau}) = F(\boldsymbol{\tau}) \quad \text{for all } \boldsymbol{\tau} \in \boldsymbol{\Sigma}, \tag{4.11}$$

and *for $G \in \boldsymbol{V}^*$, find $\boldsymbol{u} \in \boldsymbol{V}$ such that*

$$(\boldsymbol{u}, \boldsymbol{v}) = G(\boldsymbol{v}) \quad \text{for all } \boldsymbol{v} \in \boldsymbol{V}. \tag{4.12}$$

By the same analysis as before, we can conclude that the problem of finding a preconditioner S can be reduced to the problem of finding preconditioners for the $\boldsymbol{H}(\mathbf{div})$ problem (4.11) and Problem (4.12).

Lemma IV.2. *Let $S_1 : \boldsymbol{\Sigma} \to \boldsymbol{\Sigma}$ be a linear operator which is symmetric with respect to the $\boldsymbol{L}^2$ inner product and positive definite. Let $S_2 : \boldsymbol{V} \to \boldsymbol{V}$ be a linear operator which is symmetric with respect to the $\boldsymbol{L}^2$ inner product and positive definite. Assume* *that S_1 and S_2 satisfy*

$$\mu_0 \mathbf{\Lambda}(\boldsymbol{\sigma}, \boldsymbol{\sigma}) \le \mathbf{\Lambda}(S_1 \mathbf{\Lambda} \boldsymbol{\sigma}, \boldsymbol{\sigma}) \le \mu_1 \mathbf{\Lambda}(\boldsymbol{\sigma}, \boldsymbol{\sigma}) \qquad \textit{for all } \boldsymbol{\sigma} \in \mathbf{\Sigma}, \tag{4.13}$$

$$\mu_2(\boldsymbol{u}, \boldsymbol{u}) \le (S_2 \boldsymbol{u}, \boldsymbol{u}) \le \mu_3(\boldsymbol{u}, \boldsymbol{u}) \qquad \textit{for all } \boldsymbol{u} \in \boldsymbol{V}. \tag{4.14}$$

Then $\mathcal{S}_1 = S_1 \circ \iota_1$ and $\mathcal{S}_2 = S_2 \circ \iota_2$ satisfy

$$\|\mathcal{S}_1\| \le \mu_1, \qquad \|\mathcal{S}_1^{-1}\| \le \frac{1}{\mu_0},$$
$$\|\mathcal{S}_2\| \le \mu_3, \qquad \|\mathcal{S}_2^{-1}\| \le \frac{1}{\mu_2}.$$

A natural choice for S_2 is the identity operator. Then $\mathcal{S}_2 = \iota_2$ is the exact solver of Problem (4.12). The matrix representation $\underline{\underline{S}}_2$ for $\mathcal{S}_2$ is just the inverse of the mass matrix. Since the space $\boldsymbol{V}$ consists of discontinuous linears on the triangles, $\underline{\underline{S}}_2$ reduces to the inversion of a 3×3 block diagonal matrix.

All that is left is to construct a preconditioner S_1, which satisfies (4.13) with μ_0, μ_1 independent of h. Two possible constructions, the overlapping Schwarz preconditioner and the multigrid preconditioner, will be discussed and analyzed in Chapter V and Chapter VI respectively.

C. Other iterative solvers

In this section, we discuss several other iterative solvers for saddle-point problems. Notice that

$$M = \begin{pmatrix} I & 0 \\ BA^{-1} & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & -BA^{-1}B^t \end{pmatrix} \begin{pmatrix} I & A^{-1}B^t \\ 0 & I \end{pmatrix}. \tag{4.15}$$

Define $C = BA^{-1}B^t$. Clearly C is symmetric and semi-positive definite.

Several iterative solvers have been developed for system (4.7). We have already introduced the method of using preconditioned MINRES algorithm with a preconditioner constructed by norm equivalence in Section B. In this section, we briefly introduce some other methods or preconditioners.

We consider two basic type of methods:

Category (I): A linear iterative method with the $(i+1)$st step defined by

$$\hat{M}\left[\begin{pmatrix}\boldsymbol{\sigma}^{i+1}\\ \boldsymbol{u}^{i+1}\end{pmatrix}-\begin{pmatrix}\boldsymbol{\sigma}^{i}\\ \boldsymbol{u}^{i}\end{pmatrix}\right]=\begin{pmatrix}\boldsymbol{f}\\ \boldsymbol{g}\end{pmatrix}-M\begin{pmatrix}\boldsymbol{\sigma}^{i}\\ \boldsymbol{u}^{i}\end{pmatrix},$$

where $\hat{M}\in\mathcal{L}(\boldsymbol{\Sigma}\times\boldsymbol{V},\boldsymbol{\Sigma}\times\boldsymbol{V})$ satisfies

$$\rho(I-\hat{M}^{-1}M)\leq\delta<1. \tag{4.16}$$

Here $\rho(\cdot)$ denotes the spectral radius of the given operator;

Category (II): Preconditioned MINRES method with a symmetric positive definite preconditioner S such that $\kappa(SM)$ is bounded independent of the mesh size h.

Notice that $\hat{M}$ in Category (I) can not be symmetric positive definite, since otherwise $\hat{M}^{-1}M$ would be a symmetric indefinite operator under the inner product $(\hat{M}\cdot,\cdot)$ and hence $\rho(I-\hat{M}^{-1}M)>1$. Therefore, a linear iterative method in Category (I) can not induce a preconditioner in Category (II), and vice versa. This is different from the case of symmetric positive definite systems, where linear iterative methods are equivalent to preconditioners [53].

One possible way to construct a preconditioner S is given in the previous section. Here is another way to construct S [46]:

Lemma IV.3. *Assume that $\hat{A}:\boldsymbol{\Sigma}\to\boldsymbol{\Sigma}$ and $\hat{C}:\boldsymbol{V}\to\boldsymbol{V}$ are symmetric with respect to the $\boldsymbol{L}^2$ inner product and positive definite. Furthermore, assume that $Ker(C)$ and $\boldsymbol{V}/Ker(C)$ are invariant subspaces under operator $\hat{C}$. Assume there exist positive* *numbers μ_0, μ_1, μ_2 and μ_3 such that*

$$\mu_0(A\boldsymbol{\tau},\boldsymbol{\tau}) \le (\hat{A}\boldsymbol{\tau},\boldsymbol{\tau}) \le \mu_1(A\boldsymbol{\tau},\boldsymbol{\tau}) \qquad \textit{for all } \boldsymbol{\tau} \in \boldsymbol{\Sigma}, \tag{4.17}$$

$$\mu_2(C\boldsymbol{v},\boldsymbol{v}) \le (\hat{C}\boldsymbol{v},\boldsymbol{v}) \le \mu_3(C\boldsymbol{v},\boldsymbol{v}) \qquad \textit{for all } \boldsymbol{v} \in \boldsymbol{V}/Ker(C). \tag{4.18}$$

Define $S = \begin{pmatrix} \hat{A}^{-1} & 0 \\ 0 & \hat{C}^{-1} \end{pmatrix}$. *Then*

$$\kappa(SM) \le \frac{(\sqrt{5}+1)\max\{\mu_1,\mu_3\}}{2\min\{\mu_0,\mu_2\}}.$$

Proof. Define $S_0 = \begin{pmatrix} A^{-1} & 0 \\ 0 & C^{\dagger} \end{pmatrix}$, where $C^{\dagger}$ is a pseudo-inverse of C, and denote $T = S_0 M$. A simple calculation shows that

$$T(T-I)(T^2-T-I) = 0$$

and hence $\kappa(S_0M) = \frac{\sqrt{5}+1}{2}$ (see [46]). Since S_0 is symmetric with respect to the $\boldsymbol{L}^2$ inner product and semi-positive definite, it induces an operator norm on $\mathcal{L}(\boldsymbol{\Sigma} \times \boldsymbol{V}/Ker(C), \boldsymbol{\Sigma} \times \boldsymbol{V}/Ker(C))$ by

$$\|\mathcal{O}\|_{S_0^{-1}} = \sup_{\substack{x \in \boldsymbol{\Sigma}\times\boldsymbol{V}/Ker(C) \\ x \neq 0}} \frac{(S_0^{-1}\mathcal{O}x, \mathcal{O}x)^{1/2}}{(S_0^{-1}x, x)^{1/2}}.$$

It is not hard to see that $(\boldsymbol{\Sigma}\times\boldsymbol{V})/Ker(M) = Im(M) = \boldsymbol{\Sigma}\times\boldsymbol{V}/Ker(C)$. Therefore, in the estimate of $\kappa(SM)$, we can simply restrict all operators to $\boldsymbol{\Sigma}\times\boldsymbol{V}/Ker(C)$. Notice that $\rho(\mathcal{O}) = \|\mathcal{O}\|_{S_0^{-1}}$ if $\mathcal{O}$ is symmetric under the inner product $(S_0^{-1}\cdot,\cdot)$. We have

$$\begin{aligned}
\kappa(SM) &\le \|SM\|_{S_0^{-1}} \|M^{-1}S^{-1}\|_{S_0^{-1}} \\
&\le \|SS_0^{-1}\|_{S_0^{-1}} \|S_0M\|_{S_0^{-1}} \|M^{-1}S_0^{-1}\|_{S_0^{-1}} \|S_0S^{-1}\|_{S_0^{-1}} \\
&\le \frac{(\sqrt{5}+1)\max\{\mu_1,\mu_3\}}{2\min\{\mu_0,\mu_2\}}.
\end{aligned}$$

□

Next, we introduce different ways to construct $\hat{M}$ in Category (I). Intuitively, one wants to choose $\hat{M}$ so that $\hat{M}^{-1}$ is an inexpensive approximation of M. By Equation (4.15), we would like to first consider

$$M_1 = \begin{pmatrix} I & 0 \\ BA^{-1} & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & -\hat{C} \end{pmatrix} = \begin{pmatrix} A & 0 \\ B & -\hat{C} \end{pmatrix}$$

and a symmetric version

$$M_2 = \begin{pmatrix} I & 0 \\ BA^{-1} & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & -\hat{C} \end{pmatrix} \begin{pmatrix} I & A^{-1}B^t \\ 0 & I \end{pmatrix} = \begin{pmatrix} A & B^t \\ B & BA^{-1}B^t - \hat{C} \end{pmatrix}.$$

Lemma IV.4. *Assume that* $\rho(I - \hat{C}^{-1}C) \leq \delta < 1$. *Then*

$$\rho(I - M_1^{-1}M) \leq \delta,$$
$$\rho(I - M_2^{-1}M) \leq \delta.$$

Proof. The results follow clearly from

$$I - M_1^{-1}M = \begin{pmatrix} 0 & -A^{-1}B^t \\ 0 & I - \hat{C}^{-1}C \end{pmatrix},$$
$$I - M_2^{-1}M = \begin{pmatrix} I & -A^{-1}B^t \\ 0 & I \end{pmatrix} \begin{pmatrix} 0 & o \\ 0 & I - \hat{C}^{-1}C \end{pmatrix} \begin{pmatrix} I & A^{-1}B^t \\ 0 & I \end{pmatrix}.$$

□

Indeed, the linear iterative method using M_1 is the preconditioned Uzawa method [31, 17], in which the inner iteration requires the exact inverse of A. This, in some applications, can be expensive. Therefore, an inexact Uzawa method has been proposed:

$$\tilde{M}_1 = \begin{pmatrix} I & 0 \\ B\hat{A}^{-1} & I \end{pmatrix} \begin{pmatrix} \hat{A} & 0 \\ 0 & -\hat{C} \end{pmatrix} = \begin{pmatrix} \hat{A} & 0 \\ B & -\hat{C} \end{pmatrix}$$

where $\hat{A}^{-1}$ is an approximation of A and $\hat{C}^{-1}$ is an approximation of C. Also, a symmetric version of the inexact method can be defined by

$$\tilde{M}_2 = \begin{pmatrix} I & 0 \\ B\hat{A}^{-1} & I \end{pmatrix} \begin{pmatrix} \hat{A} & 0 \\ 0 & -\hat{C} \end{pmatrix} \begin{pmatrix} I & \hat{A}^{-1}B^t \\ 0 & I \end{pmatrix} = \begin{pmatrix} \hat{A} & B^t \\ B & B\hat{A}^{-1}B^t - \hat{C} \end{pmatrix}.$$

The linear iterative method with $\tilde{M}_2$ was introduced by Bank, Welfert and Yserentant in [12].

The convergence rate analysis for $\tilde{M}_1$ and $\tilde{M}_2$ is non-trivial (see [31, 17] for the analysis of $\tilde{M}_1$ and [12] for the analysis of $\tilde{M}_2$). Also, in a recent paper [55], the author analyzed both $\tilde{M}_1$ and $\tilde{M}_2$ and their convergence rates under a unified framework, which is based on the fact that the error reduction matrices in both methods can be transformed to a product of symmetric matrices and block diagonal matrices. We skip the convergence rate analysis here since it is not our purpose.

Remark IV.3. *Notice that the linear iterative method stated as Category (I) is equivalent to the Richardson method applied to the following system*

$$\hat{M}^{-1} M \begin{pmatrix} \boldsymbol{\sigma} \\ \boldsymbol{u} \end{pmatrix} = \hat{M}^{-1} \begin{pmatrix} \boldsymbol{f} \\ \boldsymbol{g} \end{pmatrix}. \tag{4.19}$$

Therefore, $\hat{M}$ is also called a preconditioner in the literature.

Finally, we mention the idea of positive definite reformulation introduced by

Bramble and Pasciak in [15]. Consider system (4.19). Set

$$\hat{M} = \begin{pmatrix} \hat{A} & 0 \\ B & -I \end{pmatrix},$$

where $\hat{A} : \mathbf{\Sigma} \to \mathbf{\Sigma}$ is symmetric positive definite and satisfies (4.17) with $\mu_1 < 1$. Set

$$T = \begin{pmatrix} A - \hat{A} & 0 \\ 0 & I \end{pmatrix}.$$

Clearly, T is symmetric positive definite. It has been shown ([15]) that $\hat{M}^{-1}M$ is symmetric positive definite under the inner product $(T\cdot, \cdot)$. Hence the iterative methods for symmetric positive definite problem can be applied for system (4.19) under the inner product $(T\cdot, \cdot)$ and preconditioners can be developed.

CHAPTER V

THE OVERLAPPING SCHWARZ PRECONDITIONER

In this chapter, we develop a two-level overlapping Schwarz preconditioner for the $\boldsymbol{H}(\mathbf{div})$ problem (4.11) and analyze the condition number of the preconditioned system. We start with introducing the framework of an overlapping Schwarz precoditioner in Section A. An abstract condition number estimate is given here. In Section B, we build the subspace decompositione and several important operators for the $\boldsymbol{H}(\mathbf{div})$ problem. Finally, in Section C, we prove that the assumptions for the abstract condition number estimate hold under the settings of Section B.

A. Framework of an overlapping Schwarz preconditioner

Denote $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\mathcal{T}, \Omega, \Gamma_T)$. Recall that we have defined the bilinear form $\boldsymbol{\Lambda}$ on $\boldsymbol{\Sigma}$ in (4.10) and used the same notation $\boldsymbol{\Lambda}$ for the operator induced by the bilinear form $\boldsymbol{\Lambda}$. For $i = 0, \ldots, K$, let $\boldsymbol{\Sigma}_i \subset \boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}\,, \Omega, \mathbb{S}_2)$ and assume that there exist linear operators $\mathbf{I}_i : \boldsymbol{\Sigma}_i \to \boldsymbol{\Sigma}$ such that

$$\boldsymbol{\Sigma} = \sum_{i=0}^{K} \mathbf{I}_i \boldsymbol{\Sigma}_i.$$

Both the $\boldsymbol{L}^2$ inner product $(\cdot, \cdot)$ and the $\boldsymbol{H}(\mathbf{div}\,, \Omega, \mathbb{S}_2)$ inner product $\boldsymbol{\Lambda}(\cdot, \cdot)$ are well defined on $\boldsymbol{\Sigma}_i$.

In the remainder of this chapter, we use $\lesssim$ to denote "less than or equal to" up to a positive constant independent of the mesh size h and the number of subspaces K.

For each i, define the operator $\boldsymbol{\Lambda}_i : \boldsymbol{\Sigma}_i \to \boldsymbol{\Sigma}_i$ by

$$(\boldsymbol{\Lambda}_i \boldsymbol{\sigma}, \boldsymbol{\tau}) = \boldsymbol{\Lambda}(\boldsymbol{\sigma}, \boldsymbol{\tau}) \quad \text{for all } \boldsymbol{\sigma}, \boldsymbol{\tau} \in \boldsymbol{\Sigma}_i.$$

Let $\mathbf{I}_i^t : \mathbf{\Sigma} \to \mathbf{\Sigma}_i$ be the $\boldsymbol{L}^2$-adjoint of $\mathbf{I}_i$ and $\mathbf{P}_i : \mathbf{\Sigma} \to \mathbf{\Sigma}_i$ be the $\mathbf{\Lambda}$-adjoint of $\mathbf{I}_i$. Clearly, we have

$$\mathbf{\Lambda}_i \mathbf{P}_i = \mathbf{I}_i^t \mathbf{\Lambda}.$$

This gives the subspace problem: *given* $\boldsymbol{f} \in \mathbf{\Sigma}$, *find* $\boldsymbol{\sigma}_i = \mathbf{P}_i \boldsymbol{\sigma} \in \mathbf{\Sigma}_i$, *where* $\boldsymbol{\sigma}$ *satisfies* $\mathbf{\Lambda}\boldsymbol{\sigma} = \boldsymbol{f}$, *such that*

$$(\mathbf{\Lambda}_i \boldsymbol{\sigma}_i, \boldsymbol{\tau}) = (\mathbf{I}_i^t \mathbf{\Lambda}\boldsymbol{\sigma}, \boldsymbol{\tau}) = (\mathbf{I}_i^t \boldsymbol{f}, \boldsymbol{\tau}) \quad \text{for all } \boldsymbol{\tau} \in \mathbf{\Sigma}_i. \tag{5.1}$$

Normally, each $\mathbf{\Sigma}_i$ is a low dimensional space compared to $\mathbf{\Sigma}$, which makes Problem (5.1) easier to solve than Problem (4.11). This gives rise to the overlapping Schwarz method [53, 29], which uses cheaper subspace solvers to build an iterative method or, equivalently, a preconditioner.

Let $\mathbf{R}_i : \mathbf{\Sigma}_i \to \mathbf{\Sigma}_i$, $i = 0, \ldots, K$, be linear operators which are symmetric with respect to the $\boldsymbol{L}^2$ inner product and positive definite. Assume that there exist positive constants r_0 and r_1 such that

$$r_0(\mathbf{\Lambda}_i^{-1}\boldsymbol{\sigma}, \boldsymbol{\sigma}) \le (\mathbf{R}_i \boldsymbol{\sigma}, \boldsymbol{\sigma}) \le r_1(\mathbf{\Lambda}_i^{-1}\boldsymbol{\sigma}, \boldsymbol{\sigma}) \qquad \text{for } i = 0, \ldots, K \text{ and } \boldsymbol{\sigma} \in \mathbf{\Sigma}_i. \tag{5.2}$$

For $i = 0, \ldots, k$, define the operator $\mathbf{T}_i : \mathbf{\Sigma} \to \mathbf{\Sigma}$ by

$$\mathbf{T}_i = \mathbf{I}_i \mathbf{R}_i \mathbf{\Lambda}_i \mathbf{P}_i = \mathbf{I}_i \mathbf{R}_i \mathbf{I}_i^t \mathbf{\Lambda}.$$

Clearly $\mathbf{T}_i$ is symmetric with respect to $\mathbf{\Lambda}(\cdot, \cdot)$. The additive and multiplicative Schwarz preconditioners (denoted by $\mathbf{B}_a$ and $\mathbf{B}_m$ respectively) are defined by:

$$\begin{aligned}
\mathbf{B}_a \mathbf{\Lambda} &= \sum_{i=0}^{k} \mathbf{T}_i; \\
\mathbf{B}_m \mathbf{\Lambda} &= \mathbf{I} - (\mathbf{I} - \mathbf{T}_k)(\mathbf{I} - \mathbf{T}_{k-1}) \cdots (\mathbf{I} - \mathbf{T}_0)^2 \cdots (\mathbf{I} - \mathbf{T}_{k-1})(\mathbf{I} - \mathbf{T}_k) \\
&= \mathbf{I} - \mathbf{E}^* \mathbf{E}.
\end{aligned}$$

Note that the computation of the action of $\mathbf{B}_a$ or $\mathbf{B}_m$ on a function $\boldsymbol{f} \in \mathbf{\Sigma}$ only involves the application of $\mathbf{\Lambda}$, the approximate solution of subspace problems using $\mathbf{R}_i$ and the application of the interpolation operator $\mathbf{I}_i$.

It is clear that $\mathbf{B}_a = \sum_{i=1}^{K} \mathbf{I}_i \mathbf{R}_i \mathbf{I}_i^t$. We have the following result [54]:

Lemma V.1. $\mathbf{B}_a^{-1} : \mathbf{\Sigma} \to \mathbf{\Sigma}$ *exists and*

$$(\mathbf{B}_a^{-1}\boldsymbol{\sigma}, \boldsymbol{\sigma}) = \min_{\substack{\boldsymbol{\sigma}_i \in \mathbf{\Sigma}_i \\ \boldsymbol{\sigma} = \sum_{i=0}^{K} \mathbf{I}_i \boldsymbol{\sigma}_i}} \sum_{i=0}^{K} (\mathbf{R}_i^{-1}\boldsymbol{\sigma}_i, \boldsymbol{\sigma}_i). \tag{5.3}$$

Next, we give abstract condition number estimates for both $\mathbf{B}_a\mathbf{\Lambda}$ and $\mathbf{B}_m\mathbf{\Lambda}$. The proof of the following result is standard [53, 52]:

Theorem V.1. *Assume that (5.2) holds and:*

(S.1) For all $\boldsymbol{\sigma} \in \mathbf{\Sigma}$, *there exists a decomposition* $\boldsymbol{\sigma} = \sum_{i=0}^{K} \mathbf{I}_i \boldsymbol{\sigma}_i$ *and a positive constant* C_0 *such that*

$$\sum_{i=0}^{K} \mathbf{\Lambda}(\boldsymbol{\sigma}_i, \boldsymbol{\sigma}_i) \le C_0 \mathbf{\Lambda}(\boldsymbol{\sigma}, \boldsymbol{\sigma});$$

(S.2) For all $\boldsymbol{\sigma}_i, \boldsymbol{\tau}_i \in \mathbf{\Sigma}_i$, $0 \le i \le K$, *there exists a positive constant* C_1 *such that*

$$\sum_{i,j=0}^{K} \mathbf{\Lambda}(\mathbf{I}_i\boldsymbol{\sigma}_i, \mathbf{I}_j\boldsymbol{\tau}_j) \le C_1 \left(\sum_{i=0}^{K} \mathbf{\Lambda}(\boldsymbol{\sigma}_i, \boldsymbol{\sigma}_i) \right)^{1/2} \left(\sum_{j=0}^{K} \mathbf{\Lambda}(\boldsymbol{\tau}_j, \boldsymbol{\tau}_j) \right)^{1/2}.$$

Then

$$\frac{r_0}{C_1} \mathbf{\Lambda}(\boldsymbol{\sigma}, \boldsymbol{\sigma}) \le (\mathbf{B}_a^{-1}\boldsymbol{\sigma}, \boldsymbol{\sigma}) \le r_1 C_0 \mathbf{\Lambda}(\boldsymbol{\sigma}, \boldsymbol{\sigma}) \qquad \textit{for all } \boldsymbol{\sigma} \in \mathbf{\Sigma}. \tag{5.4}$$

If, in addition to (S.1) and (S.2), we assume:

(S.3) There exists a constant C_2 *with* $0 \le C_2 \le 2$ *and*

$$\mathbf{\Lambda}(\mathbf{T}_i\boldsymbol{\sigma}, \boldsymbol{\sigma}) \le C_2 \mathbf{\Lambda}(\boldsymbol{\sigma}, \boldsymbol{\sigma}) \quad \textit{for all } \boldsymbol{\sigma} \in \mathbf{\Sigma} \textit{ and } 0 \le i \le K.$$

Then

$$\frac{2 - C_2}{C_0 r_1^2 (1 + C_1)^2} \mathbf{\Lambda}(\boldsymbol{\sigma}, \boldsymbol{\sigma}) \leq \mathbf{\Lambda}(\mathbf{B}_m \mathbf{\Lambda} \boldsymbol{\sigma}, \boldsymbol{\sigma}) \leq \mathbf{\Lambda}(\boldsymbol{\sigma}, \boldsymbol{\sigma}) \qquad \textit{for all } \boldsymbol{\sigma} \in \mathbf{\Sigma}. \tag{5.5}$$

Notice that (5.4) is equivalent to

$$\frac{1}{r_1 C_0} \mathbf{\Lambda}(\boldsymbol{\sigma}, \boldsymbol{\sigma}) \leq \mathbf{\Lambda}(\mathbf{B}_a \mathbf{\Lambda} \boldsymbol{\sigma}, \boldsymbol{\sigma}) \leq \frac{C_1}{r_0} \mathbf{\Lambda}(\boldsymbol{\sigma}, \boldsymbol{\sigma}) \qquad \text{for all } \boldsymbol{\sigma} \in \mathbf{\Sigma}.$$

B. An overlapping Schwarz preconditioner for the $\boldsymbol{H}(\mathbf{div})$ problem

In this section, we build the subspaces and operators needed for defining an overlapping Schwarz preconditioner for Problem (4.11) following the framework given in Section A. For simplicity, we only consider the pure traction boundary problem on a polygonal domain Ω. The analysis can easily be generalized to problems with general boundary conditions which have $\boldsymbol{H}^s$-regularity for $s > 1/2$.

Let $\mathcal{T}_H$ be a quasi-uniform mesh on Ω with characteristic mesh size H and $\mathcal{T}_h$ be a quasi-uniform refinement of $\mathcal{T}_H$ with characteristic mesh size h. Let $\tilde{\Omega}_i$, $i = 1, \ldots, K$ be a non-overlapping decomposition of Ω whose boundaries align with the coarse mesh $\mathcal{T}_H$. Extend $\tilde{\Omega}_i$ by one or more layers of fine elements to get Ω_i. Then we have an overlapping cover $\{\Omega_i\}_{i=1}^K$ of Ω whose boundaries align with the fine mesh $\mathcal{T}_h$. Figure 8 illustrates how the sub-domains are defined inside Ω and near the boundary of Ω. The bold line contour draws the boundary of $\tilde{\Omega}_i$ and the outermost dashed line contour draws the boundary of Ω_i. We have illustrated the case of one cell overlap although we may overlap many more cells in practice.

Denote the characteristic distance between $\partial\tilde{\Omega}_i \backslash \partial\Omega$ and $\partial\Omega_i \backslash \partial\Omega$ as δ. Furthermore, assume that there exists a positive integer N_c such that each $x \in \Omega$ is included

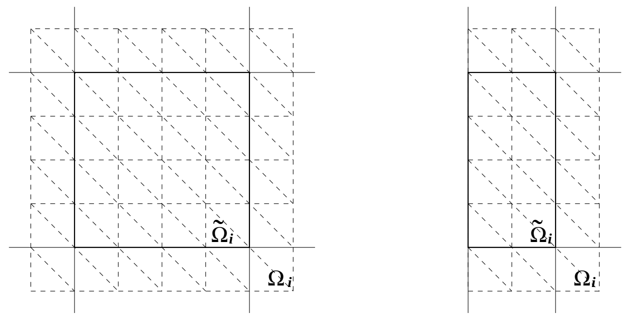

FIGURE 8. Sub-domains $\tilde{\Omega}_i$ and Ω_i.

in at most N_c sub-domains in $\{\Omega_i\}$. Define

$$\mathrm{Q}_0 = \mathrm{Q}(\mathcal{T}_H, \Omega, \partial\Omega), \qquad \mathbf{\Sigma}_0 = \mathbf{\Sigma}(\mathcal{T}_H, \Omega, \partial\Omega), \qquad \boldsymbol{V}_0 = \boldsymbol{V}(\mathcal{T}_H, \Omega),$$
$$\mathrm{Q} = \mathrm{Q}(\mathcal{T}_h, \Omega, \partial\Omega), \qquad \mathbf{\Sigma} = \mathbf{\Sigma}(\mathcal{T}_h, \Omega, \partial\Omega), \qquad \boldsymbol{V} = \boldsymbol{V}(\mathcal{T}_h, \Omega).$$

For $i = 1, \ldots, K$, define $\mathbf{\Sigma}_i$, $\boldsymbol{V}_i$ and Q_i to be the subspaces of $\mathbf{\Sigma}$, $\boldsymbol{V}$ and Q respectively, which vanish outside Ω_i. Recalling how we defined the boundary conditions for $\mathrm{Q}(\mathcal{T}, \Omega, \partial\Omega)$ and $\mathbf{\Sigma}(\mathcal{T}, \Omega, \partial\Omega)$, it is clear that

$$\mathrm{Q}_i \subsetneq \mathrm{Q}(\mathcal{T}_h, \Omega_i, \partial\Omega_i), \qquad \mathbf{\Sigma}_i \subsetneq \mathbf{\Sigma}(\mathcal{T}_h, \Omega_i, \partial\Omega_i) \quad \text{for all } i = 1, \ldots, K.$$

Hence the space $\mathbf{\Sigma}_i$ does not correspond to a natural stress tensor approximation subspace with pure traction boundary condition.

Denote $\Psi(\mathcal{T})$ to be the set of all vertices in mesh $\mathcal{T}$. We know that $\mathrm{Q}_0 \not\subseteq \mathrm{Q}$ and $\mathbf{\Sigma}_0 \not\subseteq \mathbf{\Sigma}$ since, for example, a function $\boldsymbol{\sigma} \in \mathbf{\Sigma}_0$ is not necessarily continuous at the points in $\Psi(\mathcal{T}_h)\backslash\Psi(\mathcal{T}_H)$ and a function $q \in \mathrm{Q}_0$ does not necessarily have continuous second order derivatives at the points in $\Psi(\mathcal{T}_h)\backslash\Psi(\mathcal{T}_H)$. Hence we need to define interpolation operators. The easiest way to do this is by using the nodal value interpolation on each $T \in \mathcal{T}_h$ and then taking average on the discontinuous degrees of freedom at vertices. For any point $v \in \Psi(\mathcal{T}_h)$, let $\Theta(v)$ be the set of all triangles in $\mathcal{T}_H$ which contain the vertex v and denote $|\Theta(v)|$ to be the number of triangles in $\Theta(v)$. For $q \in \mathrm{Q}_0$ and $\boldsymbol{\tau} \in \boldsymbol{\Sigma}_0$, define $\tilde{q}$ and $\tilde{\boldsymbol{\tau}}$ as follows. On each element $T \in \mathcal{T}_h$, let $\tilde{q}|_T \in \mathrm{Q}_T$ and $\tilde{\boldsymbol{\tau}}|_T \in \boldsymbol{\Sigma}_T$ satisfy

$$\begin{aligned} \mathbf{airy}\,\tilde{q}(v)|_T &= \left(\frac{1}{|\Theta(v)|}\sum_{T_v\in\Theta(v)} \mathbf{airy}\,q(v)|_{T_v}\right) - \mathbf{airy}\,q(v)|_T, \\ \tilde{\boldsymbol{\tau}}(v)|_T &= \left(\frac{1}{|\Theta(v)|}\sum_{T_v\in\Theta(v)} \boldsymbol{\tau}(v)|_{T_v}\right) - \boldsymbol{\tau}(v)|_T, \end{aligned} \tag{5.6}$$

on each vertex v of T and $\tilde{q}$, $\tilde{\boldsymbol{\tau}}$ vanish at all the other degrees of freedom. Define

$$\begin{aligned} \mathcal{I}_0 q &= q + \tilde{q} \qquad && \text{for all } q \in \mathrm{Q}_0, \\ \mathbf{I}_0 \boldsymbol{\tau} &= \boldsymbol{\tau} + \tilde{\boldsymbol{\tau}} \qquad && \text{for all } \boldsymbol{\tau} \in \boldsymbol{\Sigma}_0. \end{aligned} \tag{5.7}$$

It is not hard to see that $\mathcal{I}_0$ preserves the boundary condition $q = 0$, $\nabla q = \mathbf{0}$ on $\partial\Omega$ and $\mathbf{I}_0$ preserves the boundary condition $\boldsymbol{\tau n} = \mathbf{0}$ on $\partial\Omega$. Therefore, $\mathcal{I}_0$ maps Q_0 to Q and $\mathbf{I}_0$ maps $\boldsymbol{\Sigma}_0$ to $\boldsymbol{\Sigma}$. Consequently, we have $\tilde{q} \in H_0^2(\Omega)$ and $\tilde{\boldsymbol{\tau}} \in \boldsymbol{H}_0(\mathbf{div}\,, \Omega, \mathbb{S}_2)$.

Lemma V.2. *For all $q \in \mathrm{Q}_0$ and $i = 0, 1, 2$, we have*

$$|q - \mathcal{I}_0 q|_{i,\Omega} \lesssim h^{2-i}|q|_{2,\Omega}. \tag{5.8}$$

Proof. For $q \in \mathrm{Q}_0$, let $\tilde{q}$ be defined as in (5.6). The Argyris element is almost affine but not affine [25]. However, by using the technique in the proof of Theorem 6.1.1 in [25] and a scaling argument, we have

$$|\tilde{q}|_{i,T} \lesssim h^{2-i}|\tilde{q}|_{2,T} \qquad \text{for } i = 0, 1, 2.$$

Let v_i, $i = 1, 2, 3$, be vertices of T. Since $\mathbf{airy}\,q$ is a symmetric matrix of cubic polynomials, we have

$$|\tilde{q}|_{2,T}^2 \lesssim \|\mathbf{airy}\,\tilde{q}\|_{0,T}^2 \lesssim h^2 \sum_{i=1}^{3} |(\mathbf{airy}\,\tilde{q})(v_i)|^2 \lesssim |\mathbf{airy}\,q|_{0,T}^2 \lesssim |q|_{2,T}^2.$$

Combining all the above, we have

$$|q - \mathcal{I}_0 q|_{i,\Omega} = |\tilde{q}|_{i,\Omega} \lesssim h^{2-i} |\tilde{q}|_{i,\Omega} \lesssim h^{2-i} |q|_{2,\Omega} \qquad \text{for } i = 0, 1, 2.$$

$\square$

Remark V.1. *The spaces* Q_i, $\boldsymbol{V}_i$ *and the operator* $\mathcal{I}_0$ *are defined only for the purpose of theoretic analysis. They are not used in the implementation of the preconditioner.*

The following lemma shows the relations between the spaces defined above.

Lemma V.3. *The following commutative diagram of exact sequences holds:*

$$\begin{array}{ccccccccc} 0 & \rightarrow & \mathrm{Q}_0 & \xrightarrow{\mathbf{airy}} & \boldsymbol{\Sigma}_0 & \xrightarrow{\mathbf{div}} & \boldsymbol{V}_0 & \longrightarrow RM \rightarrow 0 \\ & & \downarrow \mathcal{I}_0 & & \downarrow \mathbf{I}_0 & & \downarrow id & \\ 0 & \rightarrow & \mathrm{Q} & \xrightarrow{\mathbf{airy}} & \boldsymbol{\Sigma} & \xrightarrow{\mathbf{div}} & \boldsymbol{V} & \longrightarrow RM \rightarrow 0 \end{array} \tag{5.9}$$

For each $i = 1, \ldots, K$, *we have the exact sequence*

$$0 \longrightarrow \mathrm{Q}_i \xrightarrow{\mathbf{airy}} \boldsymbol{\Sigma}_i \xrightarrow{\mathbf{div}} \boldsymbol{V}_i. \tag{5.10}$$

Proof. By Lemma III.4, in order to prove (5.9), it is sufficient to prove the commutativity property. For all $q \in \mathrm{Q}_0$ and $\boldsymbol{\tau} = \mathbf{airy}\,q$, we have $\tilde{\boldsymbol{\tau}}(v)|_T = \mathbf{airy}\,\tilde{q}(v)|_T$ at each vertex v of each $T \in \mathcal{T}_h$, where $\tilde{\boldsymbol{\tau}}$ and $\tilde{q}$ are defined by (5.6). By Lemma III.6, we can conclude that $\tilde{\boldsymbol{\tau}} = \mathbf{airy}\,\tilde{q}$, which implies that $\mathbf{airy}\,\mathcal{I}_0 = \mathbf{I}_0\mathbf{airy}$. Furthermore, for each $\boldsymbol{\tau} \in \boldsymbol{\Sigma}_0$ we have $\mathbf{div}\,\tilde{\boldsymbol{\tau}} = \mathbf{0}$, which implies that $\mathbf{div}\,\mathbf{I}_0\boldsymbol{\tau} = \mathbf{div}\,\boldsymbol{\tau}$. This completes the proof of (5.9).

By the definitions of Q_i and $\boldsymbol{\Sigma}_i$, for $i = 1, \ldots, K$, we can see that for each $q \in \mathrm{Q}_i$,

airy q vanishes on the vertices on $\partial\Omega_i$ and for each $\boldsymbol{\tau} \in \boldsymbol{\Sigma}_i$, $\boldsymbol{\tau}$ vanishes on the vertices on $\partial\Omega_i$. Hence by Lemma III.4 and Lemma III.6, (5.10) is clear. $\square$

The next lemma follows from the commutative diagram (5.9).

Lemma V.4. *For all $\boldsymbol{\tau} \in \boldsymbol{\Sigma}_0$, there exists a positive constant ω independent of h and H such that*

$$\boldsymbol{\Lambda}(\mathbf{I}_0\boldsymbol{\tau}, \mathbf{I}_0\boldsymbol{\tau}) \leq \omega\boldsymbol{\Lambda}(\boldsymbol{\tau}, \boldsymbol{\tau}). \tag{5.11}$$

Consequently, for all $\boldsymbol{\tau} \in \boldsymbol{\sigma}$,

$$\boldsymbol{\Lambda}(\mathbf{P}_0\boldsymbol{\tau}, \mathbf{P}_0\boldsymbol{\tau}) \leq \omega\boldsymbol{\Lambda}(\boldsymbol{\tau}, \boldsymbol{\tau}).$$

Proof. Since

$$\boldsymbol{\Lambda}(\mathbf{I}_0\boldsymbol{\tau}, \mathbf{I}_0\boldsymbol{\tau}) = \|\mathbf{I}_0\boldsymbol{\tau}\|^2_{0,\Omega} + \|\mathbf{div}\,\mathbf{I}_0\boldsymbol{\tau}\|^2_{0,\Omega} = \|\mathbf{I}_0\boldsymbol{\tau}\|^2_{0,\Omega} + \|\mathbf{div}\,\boldsymbol{\tau}\|^2_{0,\Omega},$$

we only need to show that $\|\mathbf{I}_0\boldsymbol{\tau}\|^2_{0,\Omega} \leq \omega\|\boldsymbol{\tau}\|^2_{0,\Omega}$. This follows from a standard scaling argument, the definition of $\mathbf{I}_0$ and the quasi-uniformity of the mesh. $\square$

Notice that $\boldsymbol{\Sigma}_i \subset \boldsymbol{\Sigma}$ for $i = 1, \ldots, K$. Therefore, $\mathbf{I}_i$, $i = 1, \ldots, K$ can be defined to be natural embeddings.

There are different ways to define $\mathbf{R}_i$, $i = 0, \ldots, K$, which satisfy (5.2). The simplest way is to define $\mathbf{R}_i = \boldsymbol{\Lambda}_i^{-1}$. Then $r_0 = r_1 = 1$. Notice that in this case, for all $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}$,

$$\boldsymbol{\Lambda}(\mathbf{T}_0\boldsymbol{\sigma}, \boldsymbol{\sigma}) = \boldsymbol{\Lambda}(\mathbf{P}_0\boldsymbol{\sigma}, \mathbf{P}_0\boldsymbol{\sigma}) \leq \omega\boldsymbol{\Lambda}(\boldsymbol{\tau}, \boldsymbol{\tau}).$$

Since ω is not necessarily less than 2, Assumption (S.3) may not hold. This can be dealt with by a simple modification to $\mathbf{R}_0$. Define

$$\mathbf{R}_i = \begin{cases} \rho\boldsymbol{\Lambda}_0^{-1}, & \\ \boldsymbol{\Lambda}_i^{-1} & \text{for } i = 1, \ldots, K, \end{cases}$$

where ρ satisfies $\rho\omega < 2$. Then (5.2) holds with $r_0 = \min(1, \rho)$, $r_1 = \max(1, \rho)$ and Assumption (S.3) holds with $C_2 = \rho\omega$.

Finally, we introduce a Clément type operator which will be used in the condition number estimates for $\mathbf{B}_a$ and $\mathbf{B}_m$ in Section C.

Lemma V.5. *There exists an interpolation operator $\mathcal{P}_{\mathrm{Q}_0} : H_0^2(\Omega) \to \mathrm{Q}_0$ such that*

$$|(\mathbf{I} - \mathcal{P}_{\mathrm{Q}_0})q|_{i,\Omega} \lesssim H^{2-i}|q|_{2,\Omega} \quad \textit{for all } q \in H_0^2(\Omega) \textit{ and } i = 0, 1, 2. \tag{5.12}$$

Proof. The proof will be done by construction. There exists a Clément type operator [26, 50] Π which maps $H^1(\Omega)$ onto its continuous piecewise linear subspace based on the mesh $\mathcal{T}_H$ and Π preserves the homogeneous boundary condition. For $T \in \mathcal{T}_H$, consider T to be a closed triangle which includes its own boundary. Let

$$S_T = \bigcup \{T_i | T_i \cap T \neq \emptyset, \;\; T_i \in \mathcal{T}_H\}.$$

Then, Π is stable in the sense that

$$\|\Pi w\|_{0,T} \lesssim \sum_{m=0}^{1} H^{m-1}|w|_{m,S_T} \quad \text{for all } w \in H^1(\Omega). \tag{5.13}$$

Let ϕ_j, $j = 1, \ldots, N$ be the basis of the Argyris finite element space Q_0. That is, ϕ_j is equal to 1 on the jth degree of freedom while vanishing on all the other degrees of freedom. The Argyris element is almost affine but not affine. However, by using the technique in the proof of Theorem 6.1.1 in [25], we can conclude that $\|\phi_j\|_{0,T} \lesssim H$ when the jth degree of freedom (dof) is the nodal value at a vertex or the moment on an edge, while $\|\phi_j\|_{0,T} \lesssim H^2$ when the jth degree of freedom is a first derivative at a vertex. Note that for $q \in H_0^2(\Omega)$, we have $\nabla q \in (H_0^1(\Omega))^2$. For $1 \leq j \leq N$, define linear operators $N_j:\ H_0^2(\Omega) \to \mathbb{R}$ by:

$$N_j(q) = \begin{cases} q(v), & \text{when the } j\text{th dof is the nodal value on vertex } v\,; \\ \Pi(\dfrac{\partial}{\partial x}q)(v), & \\ & \text{when the } j\text{th dof is a first derivative on vertex } v\,; \\ \Pi(\dfrac{\partial}{\partial y}q)(v), & \\ 0, & \text{when the } j\text{th dof is the second derivative on vertex } v\,; \\ \int_e \frac{\partial}{\partial \boldsymbol{n}} q\, ds, & \text{when the } j\text{th dof is the moment on edge } e\,. \end{cases}$$

Define the operator $\mathcal{P}_{\mathrm{Q}_0}:\ H_0^2(\Omega) \to \mathrm{Q}_0$ by

$$\mathcal{P}_{\mathrm{Q}_0} q = \sum_{j=1}^{N} N_j(q)\phi_j$$

Clearly $\mathcal{P}_{\mathrm{Q}_0}$ is well-defined and preserves the boundary condition. Another important observation is that $\mathcal{P}_{\mathrm{Q}_0} p = p$ for all $p \in P_1(\Omega)$. We will show that $\mathcal{P}_{\mathrm{Q}_0}$ defined as above satisfies Inequality (5.12).

First, we show that $\mathcal{P}_{\mathrm{Q}_0}$ is stable in the following sense:

$$|\mathcal{P}_{\mathrm{Q}_0} q|_{i,T} \le \sum_{m=0}^{2} H^{m-i} |q|_{m,S_T}, \quad \text{for } q \in H_0^2(\Omega),\ T \in \mathcal{T}_H,\ i = 0,1,2. \tag{5.14}$$

By the inverse inequality, we only need to prove Inequality (5.14) for $i = 0$.

Let v_k, $k = 1,2,3$ be the three vertices of T and e_k be the edge of T which is opposite to the vertex v_k. Then

$$\begin{aligned} \|\mathcal{P}_{\mathrm{Q}_0} q\|_{0,T} \le \sum_{j=1}^{N} |N_j(q)| \|\phi_j\|_{0,T} &\lesssim H \sum_{k=1}^{3} |q(v_k)| \\ &+ H^2 \sum_{k=1}^{3} |\Pi(\frac{\partial}{\partial x} q)(v_k)| + H^2 \sum_{k=1}^{3} |\Pi(\frac{\partial}{\partial y} q)(v_k)| + H \sum_{k=1}^{3} |\int_{e_k} \frac{\partial}{\partial \boldsymbol{n}} q\, ds|. \end{aligned}$$

By the Sobolev embedding theorem, Inequality (5.13) and the trace theorem, we have

$$\begin{aligned}
\sum_{k=1}^{3} |q(v_k)| &\lesssim \sum_{m=0}^{2} H^{m-1} |q|_{m,T}, \\
\sum_{k=1}^{3} |\Pi(\frac{\partial}{\partial x} q)(v_k)| + \sum_{k=1}^{3} |\Pi(\frac{\partial}{\partial y} q)(v_k)| &\lesssim \sum_{m=1}^{2} H^{m-2} |q|_{m,S_T}, \\
\sum_{k=1}^{3} |\int_{e_k} \frac{\partial}{\partial \boldsymbol{n}} q \, ds| &\lesssim H^{1/2} \| \frac{\partial}{\partial \boldsymbol{n}} q \|_{0,\partial T} \\
&\lesssim H^{1/2} (H^{-1/2} \|\nabla q\|_{0,T} + H^{1/2} \|\nabla q\|_{1,T}) \\
&\lesssim \sum_{m=1}^{2} H^{m-1} |q|_{m,T} \, .
\end{aligned}$$

The stability result (5.14) follows immediately by combining all the above inequalities.

Finally we prove Inequality (5.12). Let $q \in H_0^2(\Omega)$. By the Bramble-Hilbert Lemma, there exists a linear polynomial p such that

$$\|q - p\|_{i,S_T} \lesssim H^{2-i} |q|_{2,S_T} \qquad \text{for } i = 0, 1, 2.$$

Notice that $\mathcal{P}_{\mathrm{Q}_0} p = p$. By the triangle inequality and Inequality (5.14),

$$\begin{aligned}
|q - \mathcal{P}_{\mathrm{Q}_0} q|_{i,T} &\leq |q - p|_{i,T} + |\mathcal{P}_{\mathrm{Q}_0}(q - p)|_{i,T} \\
&\lesssim H^{2-i} |q|_{2,S_T}.
\end{aligned}$$

Inequality (5.12) follows from the above inequality and the limited overlapping property of $\{S_T\}$. $\square$

C. The condition number estimate

In this section, we prove that Assumptions (S.1) and (S.2) are true under the settings in Section B.

Denote $\mathbf{P}_{\boldsymbol{V}_0}$ to be the $\boldsymbol{L}^2$ orthogonal projection from $\boldsymbol{V}$ onto $\boldsymbol{V}_0$. Clearly,

$$\|\mathbf{P}_{\boldsymbol{V}_0}\boldsymbol{v}\|_{0,\Omega} \le \|\boldsymbol{v}\|_{0,\Omega} \quad \text{for all } \boldsymbol{v} \in \boldsymbol{V}. \tag{5.15}$$

Denote Π_{Q} to be the natural interpolation operator onto the Argyris finite element space defined on $\mathcal{T}_h$ associated with the degrees of freedom. It is not hard to see that $\Pi_{\mathrm{Q}}q$ is well defined as long as q is in $C^1(\Omega)$ and q has continuous second order derivatives on each node of the fine mesh. Furthermore, if we also have $q|_{\partial\Omega} = 0$ and $\nabla q|_{\partial\Omega} = \mathbf{0}$, then $\Pi_{\mathrm{Q}}q$ is in Q.

We construct a partition of unity $\{\theta_i\}_{i=1}^K$ using the Argyris finite element on the mesh $\mathcal{T}_h$ (without any boundary conditions). Specifically, we start with a smooth partition of unity, $\{\tilde{\theta}_i\}_{i=1}^K$ satisfying

$$(1)\ supp(\tilde{\theta}_i) \subset \overline{\Omega_i}\,; \qquad (2)\ |\tilde{\theta}_i|_{W^{j,\infty}(\Omega)} \le C\delta^{-j}, \quad j = 0, 1, 2.$$

We then define $\theta_i = \Pi_{\mathrm{Q}}\tilde{\theta}_i$. It easily follows that $\{\theta_i\}_{i=1}^K$ is a partition of unity satisfying

(1) $\theta_i|_T \in P_5(T)$ for any $T \in \mathcal{T}_h$;

(2) $\theta_i \subset C^1(\Omega)$ and θ_i has continuous second order derivatives on each vertex of the fine mesh;

(3) $|\theta_i|_{W^{j,\infty}(\Omega)} \le C\delta^{-j}, \quad j = 0, 1, 2.$

Clearly we have

$$\theta_i|_{\partial\Omega_i\backslash\partial\Omega} = 0, \quad \nabla\theta_i|_{\partial\Omega_i\backslash\partial\Omega} = \mathbf{0},$$
$$\mathbf{airy}\,\theta_i(v) = \mathbf{0} \quad \text{for all } v \in \Psi(\mathcal{T}_h) \cap (\partial\Omega_i\backslash\partial\Omega).$$

Hence for any $q \in \mathrm{Q}$, we have $\Pi_{\mathrm{Q}}(\theta_i q) \in \mathrm{Q}_i$. Furthermore, by the approximation property of the Argyris element (Theorem 6.1.1 in [25]) and an inverse inequality,

$$|\theta_i q - \Pi_{\mathrm{Q}}(\theta_i q)|_{2,\Omega}^2 \lesssim \sum_{T \in \mathcal{T}_h} (h^4 |\theta_i q|_{6,T})^2 \lesssim |\theta_i q|_{2,\Omega}^2 \quad \text{for all } q \in \mathrm{Q}.$$

Note that we can apply the inverse inequality here since $\theta_i q|_T$ is a polynomial of degree less than or equal to 10. Therefore we have

$$|\Pi_{\mathrm{Q}}(\theta_i q)|_{2,\Omega} \lesssim |\theta_i q|_{2,\Omega} \quad \text{for all } q \in \mathrm{Q}. \tag{5.16}$$

Lemma V.6. *Under the settings in Section B, Assumption (S.1) is true with* $C_0 = c(\frac{H^4}{\delta^4} + 1)$*, where* c *is a positive constant depending only on* ω *and* N_c*.*

Proof. The main idea of the proof is very similar to that used in the analysis given in [32]. It is based on the exact sequence given in Lemma III.4. The space $\boldsymbol{\Sigma}$ can be divided into two parts by the exact sequence. The decomposition in Assumption (S.1) will be constructed separately on the two different parts of $\boldsymbol{\Sigma}$.

For $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}$, define $\boldsymbol{\sigma}_0^g \in \boldsymbol{\Sigma}_0$ and $\boldsymbol{u}_0 \in \boldsymbol{V}_0$ such that

$$\begin{cases} (\boldsymbol{\sigma}_0^g, \boldsymbol{\tau}) + (\mathbf{div}\,\boldsymbol{\tau}, \boldsymbol{u}_0) = 0, & \text{for all } \boldsymbol{\tau} \in \boldsymbol{\Sigma}_0, \\ (\mathbf{div}\,\boldsymbol{\sigma}_0^g, \boldsymbol{v}) = (\mathbf{P}_{\boldsymbol{V}_0}\mathbf{div}\,\boldsymbol{\sigma}, \boldsymbol{v}), & \text{for all } \boldsymbol{v} \in \boldsymbol{V}_0. \end{cases}$$

For $i = 1, \ldots, K$, define $\boldsymbol{\sigma}_i^g \in \boldsymbol{\Sigma}(\mathcal{T}_h, \tilde{\Omega}_i, \partial\tilde{\Omega}_i)$ and $\boldsymbol{u}_i \in \boldsymbol{V}(\mathcal{T}_h, \tilde{\Omega}_i)$ such that

$$\begin{cases} (\boldsymbol{\sigma}_i^g, \boldsymbol{\tau})_{\tilde{\Omega}_i} + (\mathbf{div}\,\boldsymbol{\tau}, \boldsymbol{u}_i)_{\tilde{\Omega}_i} = 0 & \text{for all } \boldsymbol{\tau} \in \boldsymbol{\Sigma}(\mathcal{T}_h, \tilde{\Omega}_i, \partial\tilde{\Omega}_i), \\ (\mathbf{div}\,\boldsymbol{\sigma}_i^g, \boldsymbol{v})_{\tilde{\Omega}_i} = (\mathbf{div}\,\boldsymbol{\sigma} - \mathbf{P}_{\boldsymbol{V}_0}\mathbf{div}\,\boldsymbol{\sigma}, \boldsymbol{v})_{\tilde{\Omega}_i} & \text{for all } \boldsymbol{v} \in \boldsymbol{V}(\mathcal{T}_h, \tilde{\Omega}_i). \end{cases}$$

We need to show the above definitions are proper, i.e. the compatibility condition (2.9) is satisfied. Since $RM \subset \boldsymbol{V}_0 \subset \boldsymbol{V}$ and $\mathbf{div}\,\boldsymbol{\sigma}$ is orthogonal to RM in the $\boldsymbol{L}^2$ inner product, so clearly

$$(\mathbf{P}_{\boldsymbol{V}_0}\mathbf{div}\,\boldsymbol{\sigma}, \boldsymbol{v}) = (\mathbf{div}\,\boldsymbol{\sigma}, \boldsymbol{v}) = 0 \quad \text{for all } \boldsymbol{v} \in RM.$$

Thus $\boldsymbol{\sigma}_0^g$ is well defined. Since the boundary of $\tilde{\Omega}_i$ aligns with the coarse mesh, if $\boldsymbol{v}$ defined on $\tilde{\Omega}_i$ is in RM, then the extension of $\boldsymbol{v}$ by zero in $\Omega\backslash\tilde{\Omega}_i$ is in $\boldsymbol{V}_0$. Therefore,

$$\int_{\tilde{\Omega}_i} (\mathbf{div}\,\boldsymbol{\sigma} - \mathbf{P}_{\boldsymbol{V}_0}\mathbf{div}\,\boldsymbol{\sigma}) \cdot \boldsymbol{v}\, d\boldsymbol{x} = 0 \quad \text{for all } \boldsymbol{v} \in RM.$$

Hence $\boldsymbol{\sigma}_i^g$ is also well defined for $i = 1, \ldots, K$.

The moments of degree 0 and 1 of $\boldsymbol{\sigma}_i\boldsymbol{n}$ on each edge of the fine mesh on $\partial\tilde{\Omega}_i$ are zero. By Lemma III.6, we can extend $\boldsymbol{\sigma}_i^g$ to Ω_i by a divergence-free function in $\Omega_i\backslash\tilde{\Omega}_i$ which has nonzero degrees of freedom only on the vertices on $\partial\tilde{\Omega}_i$. The resulting function can be extended to Ω by zero outside Ω_i and yields a function, which is still denoted by $\boldsymbol{\sigma}_i^g$, in $\boldsymbol{\Sigma}_i$. By construction, $\mathbf{div}\,\boldsymbol{\sigma}_i^g = \mathbf{0}$ in $\Omega\backslash\tilde{\Omega}_i$. Since the mesh is quasi-uniform, a scaling argument shows that for $i = 1, \ldots, K$,

$$\|\boldsymbol{\sigma}_i^g\|_{\boldsymbol{H}(\mathbf{div},\Omega,\mathbb{S}_2)} \lesssim \|\boldsymbol{\sigma}_i^g\|_{\boldsymbol{H}(\mathbf{div},\tilde{\Omega}_i,\mathbb{S}_2)}.$$

By the above inequality and Lemma IV.1,

$$\begin{aligned}
\sum_{i=0}^{K} \boldsymbol{\Lambda}(\boldsymbol{\sigma}_i^g, \boldsymbol{\sigma}_i^g) &\lesssim \|\boldsymbol{\sigma}_0^g\|^2_{\boldsymbol{H}(\mathbf{div},\Omega,\mathbb{S}_2)} + \sum_{i=1}^{K} \|\boldsymbol{\sigma}_i^g\|^2_{\boldsymbol{H}(\mathbf{div},\tilde{\Omega}_i,\mathbb{S}_2)} \\
&\lesssim \|\mathbf{P}_{\boldsymbol{V}_0}\mathbf{div}\,\boldsymbol{\sigma}\|^2_{0,\Omega} + \sum_{i=1}^{K} \|\mathbf{div}\,\boldsymbol{\sigma} - \mathbf{P}_{\boldsymbol{V}_0}\mathbf{div}\,\boldsymbol{\sigma}\|^2_{0,\tilde{\Omega}_i} \\
&\lesssim \|\mathbf{div}\,\boldsymbol{\sigma}\|^2_{0,\Omega} \\
&\lesssim \boldsymbol{\Lambda}(\boldsymbol{\sigma}, \boldsymbol{\sigma}).
\end{aligned}$$

Next, consider

$$\boldsymbol{\sigma}^a = \boldsymbol{\sigma} - \mathbf{I}_0\boldsymbol{\sigma}_0^g - \sum_{i=1}^{K} \boldsymbol{\sigma}_i^g. \tag{5.17}$$

A simple calculation shows that $\mathbf{div}\,\boldsymbol{\sigma}^a = \mathbf{0}$. By the finite overlapping assumption and Lemma V.4, we know that

$$\begin{aligned}\mathbf{\Lambda}(\boldsymbol{\sigma}^a,\boldsymbol{\sigma}^a) &\lesssim \mathbf{\Lambda}(\boldsymbol{\sigma},\boldsymbol{\sigma}) + \omega\mathbf{\Lambda}(\boldsymbol{\sigma}_0^g,\boldsymbol{\sigma}_0^g) + N_c\sum_{i=1}^{K}\mathbf{\Lambda}(\boldsymbol{\sigma}_i^g,\boldsymbol{\sigma}_i^g)\\ &\lesssim (1+\omega+N_c)\mathbf{\Lambda}(\boldsymbol{\sigma},\boldsymbol{\sigma}).\end{aligned}$$

By the commutative diagram (5.9), for all $\boldsymbol{\tau}\in\mathbf{\Sigma}$ such that $\mathbf{div}\,\boldsymbol{\tau}=\mathbf{0}$, there exists a unique $p\in\mathrm{Q}$ which satisfies $\mathbf{airy}\,p=\boldsymbol{\tau}$. We define $\mathbf{airy}^{-1}\boldsymbol{\tau}=p$. Set

$$\begin{aligned}\boldsymbol{\sigma}_0^a &= \mathbf{airy}\,\mathcal{P}_{\mathrm{Q}_0}\mathbf{airy}^{-1}\boldsymbol{\sigma}^a,\\ \boldsymbol{\sigma}_i^a &= \mathbf{airy}\,\Pi_{\mathrm{Q}}(\theta_i\mathbf{airy}^{-1}(\boldsymbol{\sigma}^a-\mathbf{I}_0\boldsymbol{\sigma}_0^a)), \qquad \text{for } i=1,\ldots,K.\end{aligned}$$

The above definitions are proper since $\mathbf{div}\,\boldsymbol{\sigma}^a=\mathbf{0}$ and $\mathbf{div}\,(\boldsymbol{\sigma}^a-\mathbf{I}_0\boldsymbol{\sigma}_0^a)=\mathbf{0}$. Clearly

$$\boldsymbol{\sigma}^a=\sum_{i=0}^{K}\mathbf{I}_i\boldsymbol{\sigma}_i^a \tag{5.18}$$

while $\boldsymbol{\sigma}_i^a\in\mathbf{\Sigma}_i$ and $\mathbf{div}\,\boldsymbol{\sigma}_i^a=\mathbf{0}$ for $i=0,\ldots,K$. By Inequality (5.12),

$$\begin{aligned}\mathbf{\Lambda}(\boldsymbol{\sigma}_0^a,\boldsymbol{\sigma}_0^a) &= \|\mathbf{airy}\,\mathcal{P}_{\mathrm{Q}_0}\mathbf{airy}^{-1}\boldsymbol{\sigma}^a\|_{0,\Omega}^2 \lesssim |\mathcal{P}_{\mathrm{Q}_0}\mathbf{airy}^{-1}\boldsymbol{\sigma}^a|_{2,\Omega}\\ &\lesssim |\mathbf{airy}^{-1}\boldsymbol{\sigma}^a|_{2,\Omega} \lesssim \|\boldsymbol{\sigma}^a\|_{0,\Omega}^2 = \mathbf{\Lambda}(\boldsymbol{\sigma}^a,\boldsymbol{\sigma}^a).\end{aligned}$$

Let $\hat{q}=\mathbf{airy}^{-1}(\boldsymbol{\sigma}^a-\mathbf{I}_0\boldsymbol{\sigma}_0^a)$ and $q=\mathbf{airy}^{-1}\boldsymbol{\sigma}^a$. Then by the commutative diagram (5.9),

$$\hat{q}=\mathbf{airy}^{-1}(\boldsymbol{\sigma}^a-\mathbf{I}_0\mathbf{airy}\,\mathcal{P}_{\mathrm{Q}_0}q)=(\mathbf{I}-\mathcal{P}_{\mathrm{Q}_0})q+(\mathbf{I}-\mathcal{I}_0)\mathcal{P}_{\mathrm{Q}_0}q.$$

By Inequality (5.16), the assumptions on θ_i, Inequality (5.8) and Inequality (5.12),

$$\begin{aligned}\sum_{i=1}^{K} \mathbf{\Lambda}(\boldsymbol{\sigma}_i^a, \boldsymbol{\sigma}_i^a) &= \sum_{i=1}^{K} \|\mathbf{airy}\, \Pi_{\mathrm{Q}}(\theta_i \hat{q})\|_{0,\Omega}^2 \lesssim \sum_{i=1}^{K} |\theta_i \hat{q}|_{2,\Omega_i}^2 \\ &\lesssim \sum_{i=1}^{K} (\delta^{-4} |\hat{q}|_{0,\Omega_i}^2 + \delta^{-2} |\hat{q}|_{1,\Omega_i}^2 + |\hat{q}|_{2,\Omega_i}^2) \\ &\lesssim N_c (\frac{H^4}{\delta^4} + 1) |q|_{2,\Omega}^2 \\ &\lesssim N_c (\frac{H^4}{\delta^4} + 1) \|\boldsymbol{\sigma}^a\|_{0,\Omega}^2.\end{aligned}$$

Therefore we can conclude that

$$\sum_{i=0}^{K} \mathbf{\Lambda}(\boldsymbol{\sigma}_i^a, \boldsymbol{\sigma}_i^a) \lesssim (1 + \omega + N_c) N_c (\frac{H^4}{\delta^4} + 1) \mathbf{\Lambda}(\boldsymbol{\sigma}, \boldsymbol{\sigma}).$$

Finally, define $\boldsymbol{\sigma}_i = \boldsymbol{\sigma}_i^g + \boldsymbol{\sigma}_i^a$ for $i = 0, \ldots, K$. By (5.17) and (5.18), we have $\boldsymbol{\sigma} = \sum_{i=0}^{K} \mathbf{I}_i \boldsymbol{\sigma}_i$ while $\boldsymbol{\sigma}_i \in \mathbf{\Sigma}_i$ and

$$\sum_{i=0}^{K} \mathbf{\Lambda}(\boldsymbol{\sigma}_i, \boldsymbol{\sigma}_i) \leq 2(\sum_{i=0}^{K} \mathbf{\Lambda}(\boldsymbol{\sigma}_i^g, \boldsymbol{\sigma}_i^g) + \sum_{i=0}^{K} \mathbf{\Lambda}(\boldsymbol{\sigma}_i^a, \boldsymbol{\sigma}_i^a)) \leq c(\frac{H^4}{\delta^4} + 1) \mathbf{\Lambda}(\boldsymbol{\sigma}, \boldsymbol{\sigma}),$$

where c depends only on ω and N_c. This completes the proof of Lemma V.6. $\square$

Remark V.2. *In the case that* $\Gamma_D \neq \emptyset$, $\mathbf{P}_{\boldsymbol{V}_0} \mathbf{div}\, \boldsymbol{\sigma}$ *may not satisfy the compatibility condition (2.9). However,* $\mathbf{div}\, \boldsymbol{\sigma} - \mathbf{P}_{\boldsymbol{V}_0} \mathbf{div}\, \boldsymbol{\sigma}$ *still satisfies condition (2.9). Therefore, we can set* $\boldsymbol{\sigma}_0^g \in \mathbf{\Sigma}(\mathcal{T}_H, \Omega, \Gamma_T)$ *and* $\boldsymbol{\sigma}_i^g \in \mathbf{\Sigma}(\mathcal{T}_h, \Omega_i, \partial\Omega_i)$ *for* $i = 1, \ldots, K$, *which are well defined. Also,* $\mathcal{P}_{\mathrm{Q}_0}$ *should be constructed in a way such that it preserves the boundary condition on* Γ_T *(similar to the Scott-Zhang interpolation defined in Appendix A). The rest of the proof still holds.*

Remark V.3. *We have shown in the above theorem that* C_0 *is of order* $O(\frac{H^4}{\delta^4} + 1)$. *Recall that for classical second order elliptic problems, a similar result has been proved with* C_0 *of order* $O(\frac{H^2}{\delta^2} + 1)$. *In our proof the divergence free part is mapped to the fourth order Argyris finite element space, which brings* $\frac{H^4}{\delta^4}$ *to the result. It is not clear* *whether a sharper estimate can be proved for our problem.*

Lemma V.7. *Under the settings in Section B, Assumption (S.2) is true with* $C_1 = \omega + N_c$.

Proof. For simplicity, define the K-dimensional vectors $\bar{\boldsymbol{\sigma}}$ and $\bar{\boldsymbol{\tau}}$ by

$$(\bar{\boldsymbol{\sigma}})_i = \boldsymbol{\Lambda}(\boldsymbol{\sigma}_i, \boldsymbol{\sigma}_i)^{1/2} \quad \text{for } i = 1, \ldots, K,$$

$$(\bar{\boldsymbol{\tau}})_i = \boldsymbol{\Lambda}(\boldsymbol{\tau}_i, \boldsymbol{\tau}_i)^{1/2} \quad \text{for } i = 1, \ldots, K.$$

Denote $|\bar{\boldsymbol{\sigma}}|$ and $|\bar{\boldsymbol{\tau}}|$ to be the Euclidean norm of $\bar{\boldsymbol{\sigma}}$ and $\bar{\boldsymbol{\tau}}$ respectively. By the Schwarz inequality and the finite overlapping condition, we have

$$\sum_{i,j=1}^{K} \boldsymbol{\Lambda}(\boldsymbol{\sigma}_i, \boldsymbol{\tau}_j) \leq N_c |\bar{\boldsymbol{\sigma}}|\, |\bar{\boldsymbol{\tau}}|.$$

Therefore

$$\begin{aligned}
\sum_{i,j=0}^{K} \boldsymbol{\Lambda}(\mathbf{I}_i\boldsymbol{\sigma}_i, \mathbf{I}_j\boldsymbol{\tau}_j) &= \boldsymbol{\Lambda}(\mathbf{I}_0\boldsymbol{\sigma}_0, \mathbf{I}_0\boldsymbol{\tau}_0) + \boldsymbol{\Lambda}(\mathbf{I}_0\boldsymbol{\sigma}_0, \sum_{i=1}^{K}\boldsymbol{\tau}_i) \\
&\quad + \boldsymbol{\Lambda}(\sum_{i=1}^{K}\boldsymbol{\sigma}_i, \mathbf{I}_0\boldsymbol{\tau}_0) + \sum_{i,j=1}^{K}\boldsymbol{\Lambda}(\boldsymbol{\sigma}_i, \boldsymbol{\tau}_i) \\
&\leq \omega\boldsymbol{\Lambda}(\boldsymbol{\sigma}_0, \boldsymbol{\sigma}_0) + \omega^{1/2}\boldsymbol{\Lambda}(\boldsymbol{\sigma}_0, \boldsymbol{\sigma}_0)^{1/2} N_c^{1/2} |\bar{\boldsymbol{\tau}}| \\
&\quad + N_c^{1/2}|\bar{\boldsymbol{\sigma}}|\omega^{1/2}\boldsymbol{\Lambda}(\boldsymbol{\tau}_0, \boldsymbol{\tau}_0)^{1/2} + N_c|\bar{\boldsymbol{\sigma}}|\,|\bar{\boldsymbol{\tau}}| \\
&= \begin{pmatrix} \boldsymbol{\Lambda}(\boldsymbol{\sigma}_0, \boldsymbol{\sigma}_0)^{1/2} & |\bar{\boldsymbol{\sigma}}| \end{pmatrix} \begin{pmatrix} \omega & \omega^{1/2}N_c^{1/2} \\ \omega^{1/2}N_c^{1/2} & N_c \end{pmatrix} \begin{pmatrix} \boldsymbol{\Lambda}(\boldsymbol{\tau}_0, \boldsymbol{\tau}_0)^{1/2} \\ |\bar{\boldsymbol{\tau}}| \end{pmatrix} \\
&\leq (\omega + N_c) \left(\sum_{i=0}^{K} \boldsymbol{\Lambda}(\boldsymbol{\sigma}_i, \boldsymbol{\sigma}_i) \right)^{1/2} \left(\sum_{j=1}^{K} \boldsymbol{\Lambda}(\boldsymbol{\tau}_j, \boldsymbol{\tau}_j) \right)^{1/2}.
\end{aligned}$$

□

Combining Theorem V.1, Lemma V.6 and Lemma V.7, we can get a condition number estimate for $\mathbf{B}_a\boldsymbol{\Lambda}$. Notice that the condition number will depend on $\frac{H^4}{\delta^4} + 1$,

ω, N_c, r_0 and r_1, but not on h or K.

Recall that we have already shown that Assumption (S.3) holds under appropriate choices of $\mathbf{R}_i$ in Section B. Therefore, we also get a condition number estimate for $\mathbf{B}_m\boldsymbol{\Lambda}$. Again, the condition number will depend on $\frac{H^4}{\delta^4}+1$, ω, N_c, r_0 and r_1, but not on h or K.

CHAPTER VI

THE MULTIGRID PRECONDITIONER

In this section, we develop a multigrid preconditioner for the $\boldsymbol{H}(\mathbf{div})$ problem (4.11). First, we state the algorithm for the variable V-Cycle multigrid method and give an abstract condition number estimate under Assumptions (M.1) and (M.2) on the smoother. In Section B, we build a multigrid preconditioner with an additive Schwarz smoother for Problem (4.11). In Section C, we prove that the additive Schwarz smoother defined in Section B satisfies Assumptions (M.1) and (M.2).

A. Framework of a multigrid preconditioner

Assume that there exists a family of finite element spaces $\{\boldsymbol{\Sigma}_k\}_{k=1}^K$ such that $\boldsymbol{\Sigma}_k \subset \boldsymbol{H}_{0,\Gamma_T}(\mathcal{T}, \Omega, \mathbb{S}_2)$ for each k. Then, the $\boldsymbol{L}^2$ inner product $(\cdot,\cdot)$ and the $\boldsymbol{H}(\mathbf{div}\,, \Omega, \mathbb{S}_2)$ inner product $\boldsymbol{\Lambda}(\cdot,\cdot)$ are well defined on each $\boldsymbol{\Sigma}_k$. The spaces $\{\boldsymbol{\Sigma}_k\}_{k=1}^K$ may not be nested. However, we assume that there exists a series of interpolation operators $\mathbf{I}_k :\ \boldsymbol{\Sigma}_{k-1} \to \boldsymbol{\Sigma}_k$ for $k = 2, \ldots, K$.

In the remainder of this chapter, we use $\lesssim$ for "less than or equal to" up to a positive constant independent of the mesh size h and the level k.

For each k, define an operator $\boldsymbol{\Lambda}_k :\ \boldsymbol{\Sigma}_k \to \boldsymbol{\Sigma}_k$ by

$$(\boldsymbol{\Lambda}_k \boldsymbol{\sigma}, \boldsymbol{\tau}) = \boldsymbol{\Lambda}(\boldsymbol{\sigma}, \boldsymbol{\tau}) \quad \text{for all } \boldsymbol{\sigma}, \boldsymbol{\tau} \in \boldsymbol{\Sigma}_k.$$

Let $\mathbf{I}_k^t :\ \boldsymbol{\Sigma}_k \to \boldsymbol{\Sigma}_{k-1}$ be the $\boldsymbol{L}^2$-adjoint of $\mathbf{I}_k$ and let $\mathbf{P}_{k-1} :\ \boldsymbol{\Sigma}_k \to \boldsymbol{\Sigma}_{k-1}$ be the $\boldsymbol{\Lambda}$-adjoint of $\mathbf{I}_k$. It is clear that

$$\boldsymbol{\Lambda}_{k-1} \mathbf{P}_{k-1} = \mathbf{I}_k^t \boldsymbol{\Lambda}_k.$$

Assume that on each level k, there is a linear operator $\mathbf{R}_k :\ \boldsymbol{\Sigma}_k \to \boldsymbol{\Sigma}_k$ which is symmetric with respect to the $\boldsymbol{L}^2$ inner product and positive definite. We call $\mathbf{R}_k$ a *smoother*. Let m_k, $k = 2, \ldots, K$, be a series of positive numbers and assume that

$$\beta_0 m_k \leq m_{k-1} \leq \beta_1 m_k, \qquad \text{where } 1 < \beta_0 \leq \beta_1.$$

The variable V-cycle multigrid preconditioner $\mathbf{B}_k : \boldsymbol{\Sigma}_k \to \boldsymbol{\Sigma}_k$ is defined as follows:

Algorithm VI.1. *Set* $\mathbf{B}_1 = \boldsymbol{\Lambda}_1^{-1}$. *Assuming that* $\mathbf{B}_{k-1} : \boldsymbol{\Sigma}_{k-1} \to \boldsymbol{\Sigma}_{k-1}$ *is defined, define* $\mathbf{B}_k : \boldsymbol{\Sigma}_k \to \boldsymbol{\Sigma}_k$ *as follows: for* $\boldsymbol{g} \in \boldsymbol{\Sigma}_k$, *set* $\boldsymbol{\tau}^0 = \mathbf{0}$ *and define*

(1) $\boldsymbol{\tau}^l = \boldsymbol{\tau}^{l-1} + \mathbf{R}_k(\boldsymbol{g} - \boldsymbol{\Lambda}_k \boldsymbol{\tau}^{k-1})$ *for* $l = 1, \ldots, m_k$;

(2) $\boldsymbol{\sigma}^{m_k} = \boldsymbol{\tau}^{m_k} + \mathbf{I}_k \mathbf{B}_{k-1} \mathbf{I}_k^t(\boldsymbol{g} - \boldsymbol{\Lambda}_k \boldsymbol{\tau}^{m_k})$;

(3) $\boldsymbol{\sigma}^l = \boldsymbol{\sigma}^{l-1} + \mathbf{R}_k(\boldsymbol{g} - \boldsymbol{\Lambda}_k \boldsymbol{\sigma}^{l-1})$ *for* $l = m_k + 1, \ldots, 2m_k$.

Set $\mathbf{B}_k \boldsymbol{g} = \boldsymbol{\sigma}^{2m_k}$.

The general theorem giving a condition number estimate for $\mathbf{B}_k \boldsymbol{\Lambda}_k$ and its proof can be found in [19]. We state the theorem in the following:

Theorem VI.1. *Assume that*

(M.1) the spectrum of $\mathbf{I} - \mathbf{R}_k \boldsymbol{\Lambda}_k$ *lies inside the interval* $[0, 1)$;

(M.2) there exist a constant $0 < \alpha \leq 1$ *and a constant* C_p *independent of* k *such that for all* $\boldsymbol{\tau} \in \boldsymbol{\Sigma}_k$,

$$|\boldsymbol{\Lambda}_k((\mathbf{I} - \mathbf{I}_k \mathbf{P}_{k-1})\boldsymbol{\tau}, \boldsymbol{\tau})| \leq C_p^{2\alpha} \mathbf{R}_k(\boldsymbol{\Lambda}_k \boldsymbol{\tau}, \boldsymbol{\Lambda}_k \boldsymbol{\tau})^\alpha \boldsymbol{\Lambda}_k(\boldsymbol{\tau}, \boldsymbol{\tau})^{1-\alpha}.$$

Then, the preconditioner $\mathbf{B}_k$ *is symmetric and positive definite. Furthermore,* $\mathbf{B}_k$ *satisfies*

$$\frac{m_k^\alpha}{M + m_k^\alpha} \boldsymbol{\Lambda}(\boldsymbol{\tau}, \boldsymbol{\tau}) \leq \boldsymbol{\Lambda}(\mathbf{B}_k \boldsymbol{\Lambda}_k \boldsymbol{\tau}, \boldsymbol{\tau}) \leq \frac{M + m_k^\alpha}{m_k^\alpha} \boldsymbol{\Lambda}(\boldsymbol{\tau}, \boldsymbol{\tau}) \quad \text{for all } \boldsymbol{\tau} \in \boldsymbol{\Sigma}_k,$$

where M *is a sufficiently large positive constant depending only on* C_p *and* α.

B. A multigrid preconditioner for the $\boldsymbol{H}(\mathbf{div})$ problem

In this section, we construct a multigrid preconditioner for Problem (4.11) following the framework in Section A. We only consider the pure displacement boundary problem, that is, $\Gamma_t = \emptyset$, on convex polygonal domains. As noted in Chapter II, the solution has $\boldsymbol{H}^1$-regularity.

Let $\mathcal{T}_1$ be a unit-size coarse triangulation of Ω. Once we have the k-th level triangulation $\mathcal{T}_k$, define the $(k+1)$-st level mesh $\mathcal{T}_{k+1}$ by breaking each triangle in $\mathcal{T}_k$ into four triangles by connecting the midpoints of the edges. By repeating this process we get a series of nested meshes $\mathcal{T}_1, \mathcal{T}_1, \ldots, \mathcal{T}_K$. Denote the characteristic mesh size of $\mathcal{T}_k$ as h_k. Clearly, $h_k = \frac{1}{2}h_{k-1} = O(2^{-k})$. Define the finite element spaces

$$\mathrm{Q}_k = \mathrm{Q}(\mathcal{T}_k, \Omega), \qquad \boldsymbol{\Sigma}_k = \boldsymbol{\Sigma}(\mathcal{T}_k, \Omega), \qquad \boldsymbol{V}_k = \boldsymbol{V}(\mathcal{T}_k, \Omega).$$

Notice that $\mathrm{Q}_k \subset H^2(\Omega)$, $\boldsymbol{\Sigma}_k \subset \boldsymbol{H}(\mathbf{div}, \Omega, \mathbb{S}_2)$ and $\boldsymbol{V}_k \subset \boldsymbol{L}^2(\Omega, \mathbb{R}_2)$ for each k.

The bilinear form for the biharmonic problem is defined on $H^2(\Omega)$ by

$$\begin{aligned} \mathrm{A}(q,p) &= \int_\Omega \left(\frac{\partial^2 q}{\partial x^2}\frac{\partial^2 p}{\partial x^2} + 2\frac{\partial^2 q}{\partial x \partial y}\frac{\partial^2 p}{\partial x \partial y} + \frac{\partial^2 q}{\partial y^2}\frac{\partial^2 p}{\partial y^2}\right) d\boldsymbol{x} \\ &= (\mathbf{airy}\, q, \mathbf{airy}\, p). \end{aligned}$$

Define operators $\mathrm{A}_k : \mathrm{Q}_k \to \mathrm{Q}_k$ and $\boldsymbol{\Lambda}_k : \boldsymbol{\Sigma}_k \to \boldsymbol{\Sigma}_k$ by

$$\begin{aligned} (\mathrm{A}_k q, p) &= \mathrm{A}(q,p) && \text{for all } q, p \in \mathrm{Q}_k, \\ (\boldsymbol{\Lambda}_k \boldsymbol{\sigma}, \boldsymbol{\tau}) &= \boldsymbol{\Lambda}(\boldsymbol{\sigma}, \boldsymbol{\tau}) && \text{for all } \boldsymbol{\sigma}, \boldsymbol{\tau} \in \boldsymbol{\Sigma}_k. \end{aligned}$$

The spaces $\{\mathrm{Q}_k\}$ and $\{\boldsymbol{\Sigma}_k\}$ are non-nested, hence we need to define interpolation operators. Following the idea of (5.6) and (5.7), define $\tilde{q}$ and $\tilde{\boldsymbol{\tau}}$ by (5.6) with Q_0, Q,

$\mathbf{\Sigma}_0$ and $\mathbf{\Sigma}$ replaced by Q_{k-1}, Q_k, $\mathbf{\Sigma}_{k-1}$ and $\mathbf{\Sigma}_k$ respectively. Define

$$\begin{aligned} \mathcal{I}_k q &= q + \tilde{q} && \text{for all } q \in \mathrm{Q}_{k-1}, \\ \mathbf{I}_k \boldsymbol{\tau} &= \boldsymbol{\tau} + \tilde{\boldsymbol{\tau}} && \text{for all } \boldsymbol{\tau} \in \mathbf{\Sigma}_{k-1}. \end{aligned}$$

It is not hard to see that $\mathcal{I}_k$ maps Q_{k-1} to Q_k and $\mathbf{I}_k$ maps $\mathbf{\Sigma}_{k-1}$ to $\mathbf{\Sigma}_k$. Therefore $\tilde{q} \in H^2(\Omega)$ and $\tilde{\boldsymbol{\tau}} \in \boldsymbol{H}(\mathbf{div}\,, \Omega, \mathbb{S}_2)$.

Define $\mathbf{P}_{k-1}$ to be the $\mathbf{\Lambda}$-adjoint of $\mathbf{I}_k$ and define $\mathcal{P}_{k-1}$ to be the A-adjoint of $\mathcal{I}_k$. Similar to Lemma V.3 and Lemma V.4, we have the following results.

Lemma VI.1. *The following commutative diagram of exact sequences holds:*

$$\begin{array}{ccccccccccc} 0 & \longrightarrow & P_1(\Omega) & \longrightarrow & \mathrm{Q}_{k-1} & \xrightarrow{\mathbf{airy}} & \mathbf{\Sigma}_{k-1} & \xrightarrow{\mathbf{div}} & \boldsymbol{V}_{k-1} & \to & 0 \\ & & & & \downarrow \mathcal{I}_k & & \downarrow \mathbf{I}_k & & \downarrow id & & \\ 0 & \longrightarrow & P_1(\Omega) & \longrightarrow & \mathrm{Q}_k & \xrightarrow{\mathbf{airy}} & \mathbf{\Sigma}_k & \xrightarrow{\mathbf{div}} & \boldsymbol{V}_k & \to & 0 \end{array} \tag{6.1}$$

Lemma VI.2. *We have*

$$\mathbf{\Lambda}(\mathbf{I}_k \boldsymbol{\sigma}_{k-1}, \mathbf{I}_k \boldsymbol{\sigma}_{k-1}) \le \omega \mathbf{\Lambda}(\boldsymbol{\sigma}_{k-1}, \boldsymbol{\sigma}_{k-1}) \quad \textit{for all } \boldsymbol{\sigma}_{k-1} \in \mathbf{\Sigma}_{k-1},$$

where ω is independent of k. Consequently

$$\mathbf{\Lambda}(\mathbf{P}_{k-1} \boldsymbol{\sigma}_k, \mathbf{P}_{k-1} \boldsymbol{\sigma}_k) \le \omega \mathbf{\Lambda}(\boldsymbol{\sigma}_k, \boldsymbol{\sigma}_k) \quad \textit{for all } \boldsymbol{\sigma}_k \in \mathbf{\Sigma}_k.$$

One disadvantage of the interpolation $\mathbf{I}_k$ is that it has no approximation property, in general, for $\boldsymbol{\sigma} \in \mathbf{\Sigma}_{k-1}$. However, the following two lemmas indicate that $\mathbf{I}_k$ does give an "approximation" in some senses.

Lemma VI.3. *Let $\boldsymbol{\tau}_{k-1} \in \mathbf{\Sigma}_{k-1}$ be continuous piecewise linear on all components. Clearly $\boldsymbol{\tau}_{k-1} \in \mathbf{\Sigma}_k$ since it is continuous. We have*

$$((\mathbf{I} - \mathbf{I}_k)\boldsymbol{\sigma}_{k-1}, \boldsymbol{\tau}_{k-1}) = 0 \quad \textit{for all } \boldsymbol{\sigma}_{k-1} \in \mathbf{\Sigma}_{k-1}.$$

Proof: Let $T \in \mathcal{T}_{k-1}$ and v_i, $i = 1, 2, 3$, be the three midpoints of each edge of T. We note that $(\mathbf{I} - \mathbf{I}_k)\boldsymbol{\sigma}_{k-1}$ has nonzero degrees of freedom only on the nodal values at v_i, $i = 1, 2, 3$. By Lemma III.6, $(\mathbf{I} - \mathbf{I}_k)\boldsymbol{\sigma}_{k-1} = \mathbf{airy}\, q$ for some $q \in \mathrm{Q}(\mathcal{T}_k, T)$ which has nonzero degrees of freedom only on the second order derivatives on each v_i. Now, since $\boldsymbol{\sigma}_{k-1}$ has continuous normal components, we have $\mathbf{airy}\, q\, \boldsymbol{n}|_{\partial T} = \mathbf{0}$, i.e. $\frac{\partial^2}{\partial \boldsymbol{n} \partial \boldsymbol{s}} q = \frac{\partial^2}{\partial \boldsymbol{s}^2} q = 0$, where $\boldsymbol{n}$ is the outward normal vector and $\boldsymbol{s}$ is the normal tangential vector of T. It follows that both q and ∇q vanish on ∂T and integration by parts gives for continuous piecewise linear $\boldsymbol{\tau}_{k-1}$ that

$$((\mathbf{I} - \mathbf{I}_k)\boldsymbol{\sigma}_{k-1}, \boldsymbol{\tau}_{k-1}) = (\mathbf{airy}\, q, \boldsymbol{\tau}_{k-1}) = 0.$$

$\square$

Recall the definitions of the operators $\mathbf{div}^{-1}$ in (2.15) and $\mathbf{div}_{\mathcal{T}}^{-1}$ in (3.17). Denote $\mathbf{div}_k^{-1} = \mathbf{div}_{\mathcal{T}_k}^{-1}$.

Lemma VI.4. $\|(\mathbf{I} - \mathbf{I}_k)\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_{k-1}\|_{0,\Omega} \lesssim h_k \|\boldsymbol{v}_{k-1}\|_{0,\Omega}$ *for all* $\boldsymbol{v}_{k-1} \in \boldsymbol{V}_{k-1}$.

Proof: Notice that $(\mathbf{I} - \mathbf{I}_k)\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_{k-1}$ is divergence free. Therefore

$$((\mathbf{I} - \mathbf{I}_k)\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_{k-1}, \mathbf{div}^{-1}\boldsymbol{v}_{k-1}) = 0.$$

According to Lemma VI.3, for any $\boldsymbol{\tau}_{k-1} \in \boldsymbol{\Sigma}_{k-1}$ which is continuous and piecewise linear,

$$((\mathbf{I} - \mathbf{I}_k)\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_{k-1}, \boldsymbol{\tau}_{k-1}) = 0.$$

Let $\boldsymbol{\tau}_{k-1}$ be the $\boldsymbol{L}^2$ projection of $\mathbf{div}^{-1}\boldsymbol{v}_{k-1}$ into the space of continuous piecewise linear functions based on $\mathcal{T}_{k-1}$. Notice that $\mathbf{I}_k\boldsymbol{\tau}_{k-1} = \boldsymbol{\tau}_{k-1}$. By the regularity assumption and Lemma III.5,

$$\begin{aligned}
\|((\mathbf{I}-\mathbf{I}_k)\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_{k-1}\|_{0,\Omega}^2 &= ((\mathbf{I}-\mathbf{I}_k)\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_{k-1}, \mathbf{div}_{k-1}^{-1}\boldsymbol{v}_{k-1} - \mathbf{div}^{-1}\boldsymbol{v}_{k-1}) \\
&\quad - ((\mathbf{I}-\mathbf{I}_k)\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_{k-1}, \mathbf{I}_k(\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_{k-1} - \boldsymbol{\tau}_{k-1})) \\
&\lesssim \|(\mathbf{I}-\mathbf{I}_k)\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_{k-1}\|_{0,\Omega}(h_k\|\boldsymbol{v}_{k-1}\|_{0,\Omega} + \|\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_{k-1} - \boldsymbol{\tau}_{k-1}\|_{0,\Omega}) \\
&\lesssim \|(\mathbf{I}-\mathbf{I}_k)\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_{k-1}\|_{0,\Omega}(h_k\|\boldsymbol{v}_{k-1}\|_{0,\Omega} \\
&\quad + \|\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_{k-1} - \mathbf{div}^{-1}\boldsymbol{v}_{k-1}\|_{0,\Omega} + \|\mathbf{div}^{-1}\boldsymbol{v}_{k-1} - \boldsymbol{\tau}_{k-1}\|_{0,\Omega}) \\
&\lesssim h_k\|(\mathbf{I}-\mathbf{I}_k)\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_{k-1}\|_{0,\Omega}\|\boldsymbol{v}_{k-1}\|_{0,\Omega}.
\end{aligned}$$

This completes the proof of the lemma. $\square$

Finally, we construct a smoother $\mathbf{R}_k$ defined in terms of vertex based subspaces. Let $\mathcal{N}_k$ be the set of all nodes in $\mathcal{T}_k$. For each $v \in \mathcal{N}_k$, let $S_k(v)$ be the set of triangles in $\mathcal{T}_k$ meeting at the vertex v. The union of all triangles in $S_k(v)$ forms a sub-domain which we denote $\Omega_{k,v}$. Clearly $\{\Omega_{k,v}\}_{v\in\mathcal{N}_k}$ is an overlapping decomposition of Ω such that each $x \in \Omega$ is in at most three sub-domains in $\{\Omega_{k,v}\}_{v\in\mathcal{N}_k}$, which is denoted by $N_c = 3$. Define a decomposition of the spaces Q_k and $\boldsymbol{\Sigma}_k$ based on these sub-domains as follows: $\mathrm{Q}_{k,v}$ and $\boldsymbol{\Sigma}_{k,v}$ are the subspaces of functions in Q_k and $\boldsymbol{\Sigma}_k$ respectively, which have support contained in $\overline{\Omega_{k,v}}$. Let $\mathcal{P}_{k,v}: \mathrm{Q}_k \to \mathrm{Q}_{k,v}$ be the A projection, $\mathbf{P}_{k,v}: \boldsymbol{\Sigma}_k \to \boldsymbol{\Sigma}_{k,v}$ be the $\boldsymbol{\Lambda}$ projection and $\mathcal{I}_{k,v}^t: \mathrm{Q}_k \to \mathrm{Q}_{k,v}$, $\mathbf{I}_{k,v}^t: \boldsymbol{\Sigma}_k \to \boldsymbol{\Sigma}_{k,v}$ be $\boldsymbol{L}^2$ projections. Define $\mathrm{A}_{k,v}: \mathrm{Q}_{k,v} \to \mathrm{Q}_{k,v}$ and $\boldsymbol{\Lambda}_{k,v}: \boldsymbol{\Sigma}_{k,v} \to \boldsymbol{\Sigma}_{k,v}$ by

$$\begin{aligned}
(\mathrm{A}_{k,v}p, q) &= \mathrm{A}(p, q) && \text{for all } p, q \in \mathrm{Q}_{k,v}, \\
(\boldsymbol{\Lambda}_{k,v}\boldsymbol{\sigma}, \boldsymbol{\tau}) &= \boldsymbol{\Lambda}(\boldsymbol{\sigma}, \boldsymbol{\tau}) && \text{for all } \boldsymbol{\sigma}, \boldsymbol{\tau} \in \boldsymbol{\Sigma}_{k,v}.
\end{aligned}$$

Clearly, we have $\mathrm{A}_{k,v}\mathcal{P}_{k,v} = \mathcal{I}_{k,v}^t \mathrm{A}_k$ and $\mathbf{\Lambda}_{k,v}\mathbf{P}_{k,v} = \mathbf{I}_{k,v}^t \mathbf{\Lambda}_k$. Define

$$\begin{aligned} \mathcal{R}_k &= \rho \sum_{v\in\mathcal{N}_k} \mathcal{P}_{k,v}\mathrm{A}_k^{-1} = \rho \sum_{v\in\mathcal{N}_k} \mathrm{A}_{k,v}^{-1}\mathcal{I}_{k,v}^t, \\ \mathbf{R}_k &= \rho \sum_{v\in\mathcal{N}_k} \mathbf{P}_{k,v}\mathbf{\Lambda}_k^{-1} = \rho \sum_{v\in\mathcal{N}_k} \mathbf{\Lambda}_{k,v}^{-1}\mathbf{I}_{k,v}^t, \end{aligned} \tag{6.2}$$

where $\rho > 0$ is a scaling factor which only depends on the finite overlapping constant N_c. Similar to Lemma V.1, for all $\boldsymbol{\tau} \in \mathbf{\Sigma}_k$ we have

$$\mathbf{R}_k^{-1}(\boldsymbol{\tau}, \boldsymbol{\tau}) = \rho \inf_{\substack{\boldsymbol{\tau}_v \in \mathbf{\Sigma}_{k,v} \\ \sum_v \boldsymbol{\tau}_v = \boldsymbol{\tau}}} \sum_{v\in\mathcal{N}_k} \mathbf{\Lambda}(\boldsymbol{\tau}_v, \boldsymbol{\tau}_v). \tag{6.3}$$

Also, we mention that $\mathcal{R}_k$ is defined purely for theoretical analysis and only $\mathbf{R}_k$ will be used in the implementation. The implementation of $\mathbf{R}_k$ involves solving local problems on each $\Omega_{k,v}$.

Remark VI.1. *The above smoother $\mathbf{R}_k$ is constructed by using an additive Schwarz scheme. A multiplicative version of the smoother can be constructed based on the same space decomposition.*

C. The condition number estimate

In this section we prove that the smoother $\mathbf{R}_k$ satisfies Assumptions (M.1) and (M.2).

Lemma VI.5. *For $\rho \le 1/3$, the smoother $\mathbf{R}_k$ satisfies Assumption (M.1).*

Proof. The proof follows from the Schwartz inequality and the finite overlapping condition $N_c = 3$ (see [6]). $\square$

Lemma VI.6. *The smoother $\mathbf{R}_k$ satisfies Assumption (M.2).*

Proof. As shown in Section C of Chapter III, there exists a decomposition $\boldsymbol{\sigma}_k = \mathbf{airy}\, q_k + \mathbf{div}_k^{-1}\boldsymbol{v}_k$ for $\boldsymbol{\sigma}_k \in \boldsymbol{\Sigma}_K$, where $q_k \in \mathrm{Q}_k$ and $\boldsymbol{v}_k = \mathbf{div}\,\boldsymbol{\sigma}_k \in \boldsymbol{V}_k$. Define

$$\begin{aligned}
\boldsymbol{\sigma}_k^1 &= \mathbf{airy}\,(q_k - \mathcal{I}_k\mathcal{P}_{k-1}q_k),\\
\boldsymbol{\sigma}_k^2 &= \mathbf{I}_k(\mathbf{airy}\,\mathcal{P}_{k-1}q_k - \mathbf{P}_{k-1}\mathbf{airy}\,q_k),\\
\boldsymbol{\sigma}_k^3 &= \mathbf{div}_k^{-1}\boldsymbol{v}_k - \mathbf{I}_k\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_k,\\
\boldsymbol{\sigma}_k^4 &= \mathbf{I}_k(\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_k - \mathbf{P}_{k-1}\mathbf{div}_k^{-1}\boldsymbol{v}_k).
\end{aligned}$$

Notice that all $\boldsymbol{\sigma}_k^i$, $i = 1, 2, 3, 4$, are in $\boldsymbol{\Sigma}_k$ and $\boldsymbol{\sigma}_k^1$ is divergence free. Then

$$\begin{aligned}
|\boldsymbol{\Lambda}((\mathbf{I} - \mathbf{I}_k\mathbf{P}_{k-1})\boldsymbol{\sigma}_k, \boldsymbol{\sigma}_k)| &= |\boldsymbol{\Lambda}(\boldsymbol{\sigma}_k^1 + \boldsymbol{\sigma}_k^2 + \boldsymbol{\sigma}_k^3 + \boldsymbol{\sigma}_k^4, \boldsymbol{\sigma}_k)|\\
&\lesssim |\boldsymbol{\Lambda}(\boldsymbol{\sigma}_k^1, \mathbf{airy}\,q_k)| + \sum_{i=2}^{4} \mathbf{R}_k^{-1}(\boldsymbol{\sigma}_k^i, \boldsymbol{\sigma}_k^i)^{1/2}\mathbf{R}_k(\boldsymbol{\Lambda}_k\boldsymbol{\sigma}_k, \boldsymbol{\Lambda}_k\boldsymbol{\sigma}_k)^{1/2}.
\end{aligned} \tag{6.4}$$

We will show that

(I) $|\boldsymbol{\Lambda}(\boldsymbol{\sigma}_k^1, \mathbf{airy}\,q_k)| \lesssim \mathbf{R}_k(\boldsymbol{\Lambda}_k\boldsymbol{\sigma}_k, \boldsymbol{\Lambda}_k\boldsymbol{\sigma}_k)^{1/4}\boldsymbol{\Lambda}(\boldsymbol{\sigma}_k, \boldsymbol{\sigma}_k)^{3/4}$;

(II) $\mathbf{R}_k^{-1}(\boldsymbol{\sigma}_k^i, \boldsymbol{\sigma}_k^i) \lesssim \boldsymbol{\Lambda}(\boldsymbol{\sigma}_k, \boldsymbol{\sigma}_k)$ for $i = 2, 3, 4$.

Then, since Assumption (M.1) implies $\mathbf{R}_k(\boldsymbol{\Lambda}_k\boldsymbol{\sigma}_k, \boldsymbol{\Lambda}_k\boldsymbol{\sigma}_k) \leq \boldsymbol{\Lambda}(\boldsymbol{\sigma}_k, \boldsymbol{\sigma}_k)$, Assumption (M.2) will follow from (6.4), (I) and (II), with $\alpha = 1/4$.

To prove (I), first notice that for the biharmonic problem, we have ([19])

$$\frac{1}{\tilde{\lambda}_k}\|\mathrm{A}_k q_k\|_{0,\Omega}^2 \lesssim (\mathcal{R}_k\mathrm{A}_k q_k, \mathrm{A}_k q_k) \quad \text{for all } q_k \in \mathrm{Q}_k,$$

where $\tilde{\lambda}_k = O(h_k^{-4})$ is the largest eigenvalue of the operator A_k.

Theorem 14.1 in [19] states that

$$\mathrm{A}((\mathrm{I} - \mathcal{I}_k\mathcal{P}_{k-1})q_k, q_k) \lesssim (\mathrm{A}_k q_k, q_k)^{3/4}\left(\frac{\|\mathrm{A}_k q_k\|_{0,\Omega}^2}{\tilde{\lambda}_k}\right)^{1/4},$$

if Ω is a convex polygon. Therefore,

$$
\begin{aligned}
|\boldsymbol{\Lambda}(\boldsymbol{\sigma}_k^1, \mathbf{airy}\, q_k)| &= |\boldsymbol{\Lambda}(\mathbf{airy}\,(q_k - \mathcal{I}_k\mathcal{P}_{k-1}q_k), \mathbf{airy}\, q_k)| \\
&= |\mathrm{A}((\mathrm{I} - \mathcal{I}_k\mathcal{P}_{k-1})q_k, q_k)| \lesssim (\mathrm{A}_k q_k, q_k)^{3/4}\left(\frac{\|\mathrm{A}_k q_k\|_{0,\Omega}^2}{\tilde{\lambda}_k}\right)^{1/4} \\
&\lesssim \boldsymbol{\Lambda}(\boldsymbol{\sigma}_k, \boldsymbol{\sigma}_k)^{3/4}(\mathcal{R}_k\mathrm{A}_k q_k, \mathrm{A}_k q_k)^{1/4}.
\end{aligned}
$$

Thus, to prove (I), we only need to show that

$$
(\mathcal{R}_k\mathrm{A}_k q_k, \mathrm{A}_k q_k) \leq (\mathbf{R}_k\boldsymbol{\Lambda}_k\boldsymbol{\sigma}_k, \boldsymbol{\Lambda}_k\boldsymbol{\sigma}_k). \tag{6.5}
$$

Notice that by the definition of $\mathcal{R}_k$ and $\mathbf{R}_k$,

$$
\begin{aligned}
(\mathcal{R}_k\mathrm{A}_k q_k, \mathrm{A}_k q_k) &= \rho \sum_{v\in\mathcal{N}_k} \mathrm{A}(\mathcal{P}_{k,v}q_k, \mathcal{P}_{k,v}q_k), \\
(\mathbf{R}_k\boldsymbol{\Lambda}_k\boldsymbol{\sigma}_k, \boldsymbol{\Lambda}_k\boldsymbol{\sigma}_k) &= \rho \sum_{v\in\mathcal{N}_k} \boldsymbol{\Lambda}(\mathbf{P}_{k,v}\boldsymbol{\sigma}_k, \mathbf{P}_{k,v}\boldsymbol{\sigma}_k).
\end{aligned}
$$

Hence Inequality (6.5) will follow if for each $v \in \mathcal{N}_k$,

$$
\mathrm{A}(\mathcal{P}_{k,v}q_k, \mathcal{P}_{k,v}q_k) = \boldsymbol{\Lambda}(\mathbf{airy}\,(\mathcal{P}_{k,v}q_k), \mathbf{airy}\,(\mathcal{P}_{k,v}q_k)) \leq \boldsymbol{\Lambda}(\mathbf{P}_{k,v}\boldsymbol{\sigma}_k, \mathbf{P}_{k,v}\boldsymbol{\sigma}_k). \tag{6.6}
$$

Notice that for any $p \in \mathrm{Q}_{k,v}$,

$$
\begin{aligned}
\boldsymbol{\Lambda}(\mathbf{P}_{k,v}\boldsymbol{\sigma}_k, \mathbf{airy}\, p) &= (\boldsymbol{\sigma}_k, \mathbf{airy}\, p) = (\mathbf{airy}\, q_k, \mathbf{airy}\, p) \\
&= (\mathbf{airy}\, \mathcal{P}_{k,v}q_k, \mathbf{airy}\, p) = \boldsymbol{\Lambda}(\mathbf{airy}\, \mathcal{P}_{k,v}q_k, \mathbf{airy}\, p).
\end{aligned}
$$

This implies that $\mathbf{airy}\,(\mathcal{P}_{k,v}q_k)$ is the $\boldsymbol{\Lambda}$-projection of $\mathbf{P}_{k,v}\boldsymbol{\sigma}_k$ into the subspace $\mathbf{airy}\,(\mathrm{Q}_{k,v})$ of $\boldsymbol{\Sigma}_{k,v}$. Therefore, Inequality (6.6) is true. This completes the proof of (I).

Next, we prove (II). Notice that there exists a partition of unity $\{\theta_v\}_{v\in\mathcal{N}_k} \subset C(\Omega)$ which satisfies

$$
(1)\ \theta_v|_T \in P_1(T) \text{ for any } T \in \mathcal{T}_k; \qquad (2)\ supp(\theta_v) \subset \overline{\Omega_{k,v}};
$$
$$
(3)\ |\theta_v|_{W^{j,\infty}(\Omega)} \lesssim h_k^{-j}, \quad j = 0, 1.
$$

Specifically, it can be defined by interpolating a smooth partition of unity by using continuous piecewise linears.

Denote Π_k to be the natural interpolation operator onto $\boldsymbol{\Sigma}_k$ associated with the degrees of freedom. Clearly Π_k is linear and preserves $\boldsymbol{\sigma}_k \in \boldsymbol{\Sigma}_k$. Notice that for each $\boldsymbol{\sigma}_k^i$, $\Pi_k(\theta_v \boldsymbol{\sigma}_k^i)$ is a well defined function in $\boldsymbol{\Sigma}_{k,v}$ and $\boldsymbol{\sigma}_k^i = \sum_{v\in\mathcal{N}_k} \Pi_k(\theta_v \boldsymbol{\sigma}_k^i)$. Since the Arnold-Winther element is affine under the matrix Piola transformation (3.6), a simple scaling argument shows that

$$\|\Pi_k(\theta_v \boldsymbol{\tau})\|_{0,\Omega} \lesssim \|\theta_v \boldsymbol{\tau}\|_{0,\Omega}.$$

Also, by (3.8), it is easy to see that

$$\|\mathbf{div}\, \Pi_k(\theta_v \boldsymbol{\tau})\|_{0,\Omega} = \|\mathbf{P}_{\boldsymbol{V}_k} \mathbf{div}\,(\theta_v \boldsymbol{\tau})\|_{0,\Omega} \le \|\mathbf{div}\,(\theta_v \boldsymbol{\tau})\|_{0,\Omega},$$

where $\mathbf{P}_{\boldsymbol{V}_k}$ is the $\boldsymbol{L}^2$ projection onto $\boldsymbol{V}_k$.

By Equation (6.3), the inverse inequality and the properties of θ_v, for $i = 2, 3, 4$,

$$\begin{aligned}
\mathbf{R}_k^{-1}(\boldsymbol{\sigma}_k^i, \boldsymbol{\sigma}_k^i) &\le \rho \sum_{v\in\mathcal{N}_k} (\|\Pi_k(\theta_v \boldsymbol{\sigma}_k^i)\|_{0,\Omega_{k,v}}^2 + \|\mathbf{div}\, \Pi_k(\theta_v \boldsymbol{\sigma}_k^i)\|_{0,\Omega_{k,v}}^2) \\
&\lesssim \rho \sum_{v\in\mathcal{N}_k} (\|\theta_v \boldsymbol{\sigma}_k^i\|_{0,\Omega_{k,v}}^2 + \|\mathbf{div}\,(\theta_v \boldsymbol{\sigma}_k^i)\|_{0,\Omega_{k,v}}^2) \\
&\lesssim \rho h_k^{-2} \|\boldsymbol{\sigma}_k^i\|_{0,\Omega}^2 + \rho \|\mathbf{div}\, \boldsymbol{\sigma}_k^i\|_{0,\Omega}^2.
\end{aligned}$$

Hence the proof for (II) reduces to proving for $i = 2, 3, 4$ that

$$\begin{aligned}
\|\boldsymbol{\sigma}_k^i\|_{0,\Omega} &\lesssim h_k \|\boldsymbol{\sigma}_k\|_{\boldsymbol{H}(\mathbf{div}\,,\Omega,\mathbb{S}_2)}, \\
\|\mathbf{div}\, \boldsymbol{\sigma}_k^i\|_{0,\Omega} &\lesssim \|\boldsymbol{\sigma}_k\|_{\boldsymbol{H}(\mathbf{div}\,,\Omega,\mathbb{S}_2)}.
\end{aligned} \tag{6.7}$$

For $\boldsymbol{\sigma}_k^2$, clearly for any $\boldsymbol{\tau}_{k-1} = \mathbf{airy}\, p_{k-1} + \mathbf{div}_{k-1}^{-1}\boldsymbol{w}_{k-1} \in \boldsymbol{\Sigma}_{k-1}$,

$$
\begin{aligned}
&|\boldsymbol{\Lambda}(\mathbf{airy}\,\mathcal{P}_{k-1}q_k - \mathbf{P}_{k-1}\mathbf{airy}\, q_k, \boldsymbol{\tau}_{k-1})| \\
&\qquad = |(\mathbf{airy}\,\mathcal{P}_{k-1}q_k, \mathbf{airy}\, p_{k-1}) - (\mathbf{airy}\, q_k, \mathbf{I}_k\boldsymbol{\tau}_{k-1})| \\
&\qquad = |(\mathbf{airy}\, q_k, \mathbf{I}_k\mathbf{div}_{k-1}^{-1}\boldsymbol{w}_{k-1})| \\
&\qquad \le |(\mathbf{airy}\, q_k, (\mathbf{I}_k - \mathbf{I})\mathbf{div}_{k-1}^{-1}\boldsymbol{w}_{k-1})| + |(\mathbf{airy}\, q_k, \mathbf{div}_{k-1}^{-1}\boldsymbol{w}_{k-1} - \mathbf{div}^{-1}\boldsymbol{w}_{k-1})| \\
&\qquad \lesssim h_k\|\boldsymbol{\sigma}_k\|_{\boldsymbol{H}(\mathbf{div},\Omega,\mathbb{S}_2)}\|\boldsymbol{\tau}_{k-1}\|_{\boldsymbol{H}(\mathbf{div},\Omega,\mathbb{S}_2)}.
\end{aligned}
$$

We used the Schwartz inequality, Lemma III.5 and Lemma VI.4 for the last inequality above. Then, by setting $\boldsymbol{\tau}_{k-1} = \mathbf{airy}\,\mathcal{P}_{k-1}q_k - \mathbf{P}_{k-1}\mathbf{airy}\, q_k$ and using Lemma VI.2, we have

$$
\begin{aligned}
\|\boldsymbol{\sigma}_k^2\|_{\boldsymbol{H}(\mathbf{div},\Omega,\mathbb{S}_2)} &\lesssim \|\mathbf{airy}\,\mathcal{P}_{k-1}q_k - \mathbf{P}_{k-1}\mathbf{airy}\, q_k\|_{\boldsymbol{H}(\mathbf{div},\Omega,\mathbb{S}_2)} \\
&\lesssim h_k\|\boldsymbol{\sigma}_k\|_{\boldsymbol{H}(\mathbf{div},\Omega,\mathbb{S}_2)}.
\end{aligned}
$$

Therefore, $\boldsymbol{\sigma}_k^2$ satisfies (6.7).

Next, we consider $\boldsymbol{\sigma}_k^3$. Define $\mathbf{P}_{\boldsymbol{V}_{k-1}}$ to be the $\boldsymbol{L}^2$ projection onto $\boldsymbol{V}_{k-1}$. Then

$$
\|\mathbf{div}\,\boldsymbol{\sigma}_k^3\|_{0,\Omega} = \|\boldsymbol{v}_k - \mathbf{P}_{\boldsymbol{V}_{k-1}}\boldsymbol{v}_k\|_{0,\Omega} \le \|\boldsymbol{v}_k\|_{0,\Omega} \lesssim \|\boldsymbol{v}_k\|_{0,\Omega} \lesssim \|\boldsymbol{\sigma}_k\|_{H(\mathbf{div},\Omega,\mathbb{S}_2)}
$$

and by Lemma III.5 and Lemma VI.4,

$$
\begin{aligned}
\|\boldsymbol{\sigma}_k^3\|_{0,\Omega} &\lesssim \|\mathbf{div}_k^{-1}\boldsymbol{v}_k - \mathbf{div}^{-1}\boldsymbol{v}_k\|_{0,\Omega} + \|\mathbf{div}^{-1}\boldsymbol{v}_k - \mathbf{div}_{k-1}^{-1}\boldsymbol{v}_k\|_{0,\Omega} \\
&\qquad + \|(\mathbf{I} - \mathbf{I}_k)\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_k\|_{0,\Omega} \\
&\lesssim h_k\|\boldsymbol{v}_k\|_{0,\Omega} \lesssim h_k\|\boldsymbol{\sigma}_k\|_{H(\mathbf{div},\Omega,\mathbb{S}_2)}.
\end{aligned}
$$

Hence $\boldsymbol{\sigma}_k^3$ satisfies (6.7).

For $\boldsymbol{\sigma}_k^4$, let $\boldsymbol{\tau}_{k-1} \in \boldsymbol{\Sigma}_{k-1}$ be arbitrary. Then

$$
\begin{aligned}
|\boldsymbol{\Lambda}(\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_k - \mathbf{P}_{k-1}\mathbf{div}_k^{-1}\boldsymbol{v}_k, \boldsymbol{\tau}_{k-1})| &= |\boldsymbol{\Lambda}(\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_k, \boldsymbol{\tau}_{k-1}) - \boldsymbol{\Lambda}(\mathbf{div}_k^{-1}\boldsymbol{v}_k, \mathbf{I}_k\boldsymbol{\tau}_{k-1})| \\
&= |(\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_k, \boldsymbol{\tau}_{k-1}) - (\mathbf{div}_k^{-1}\boldsymbol{v}_k, \mathbf{I}_k\boldsymbol{\tau}_{k-1}) + (\mathbf{P}_{\boldsymbol{V}_{k-1}}\boldsymbol{v}_k - \boldsymbol{v}_k, \mathbf{div}\, \tau_{k-1})| \\
&= |(\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_k, \boldsymbol{\tau}_{k-1}) - (\mathbf{div}_k^{-1}\boldsymbol{v}_k, \mathbf{I}_k\boldsymbol{\tau}_{k-1})|.
\end{aligned}
\tag{6.8}
$$

Since $(\mathbf{div}^{-1}\boldsymbol{v}_k, (\mathbf{I}-\mathbf{I}_k)\boldsymbol{\tau}_{k-1})$ is zero, by (6.8), Lemma III.5 and Lemma VI.2, we have

$$
\begin{aligned}
|\boldsymbol{\Lambda}(\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_k - \mathbf{P}_{k-1}\mathbf{div}_k^{-1}\boldsymbol{v}_k, \boldsymbol{\tau}_{k-1})| &= |(\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_k - \mathbf{div}^{-1}\boldsymbol{v}_k, \boldsymbol{\tau}_{k-1}) \\
&\quad + (\mathbf{div}^{-1}\boldsymbol{v}_k - \mathbf{div}_k^{-1}\boldsymbol{v}_k, \mathbf{I}_k\boldsymbol{\tau}_{k-1})| \\
&\lesssim h_k\|\boldsymbol{\sigma}_k\|_{H(\mathbf{div}\,,\Omega,\mathbb{S}_2)}\|\boldsymbol{\tau}_{k-1}\|_{H(\mathbf{div}\,,\Omega,\mathbb{S}_2)}.
\end{aligned}
$$

Therefore, by setting $\boldsymbol{\tau}_{k-1} = \mathbf{div}_{k-1}^{-1}\boldsymbol{v}_k - \mathbf{P}_{k-1}\mathbf{div}_k^{-1}\boldsymbol{v}_k$ and using Lemma VI.2,

$$
\begin{aligned}
\|\boldsymbol{\sigma}_k^4\|_{\boldsymbol{H}(\mathbf{div}\,,\Omega,\mathbb{S}_2)} &\lesssim \|\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_k - \mathbf{P}_{k-1}\mathbf{div}_k^{-1}\boldsymbol{v}_k\|_{\boldsymbol{H}(\mathbf{div}\,,\Omega,\mathbb{S}_2)} \\
&\lesssim h_k\|\boldsymbol{\sigma}_k\|_{\boldsymbol{H}(\mathbf{div}\,,\Omega,\mathbb{S}_2)}.
\end{aligned}
$$

Therefore, $\boldsymbol{\sigma}_k^4$ satisfies (6.7).

Combining all the above shows that $\mathbf{R}_k$ satisfies Assumption (M.2) with a constant C_p independent of k. □

By Theorem VI.1, Lemma VI.5 and Lemma VI.6, the condition number of $\mathbf{B}_k\boldsymbol{\Lambda}_k$ is independent of the level k.

CHAPTER VII

NUMERICAL RESULTS

In this chapter, we present results from several numerical experiments. We start with an experiment on the approximation behavior of the Arnold-Winther element for the mixed elasticity problem. Then, experiments on the condition number estimates of preconditioned systems with the overlapping Schwarz preconditioner and the multigrid preconditioner are given.

We only consider homogeneous isotropic material (see Equation (2.4)) with Lamé coefficients $\mu = 0.5$ and $\lambda = 1$. In all the following experiments, Ω is the unit square $(0,1) \times (0,1)$. Let $\mathcal{T}_1$ be the mesh obtained by bisecting Ω into two triangles using its negatively sloped diagonal. For $k = 2, \ldots$, define $\mathcal{T}_k$ by breaking each triangle in the mesh $\mathcal{T}_{k-1}$ into four triangles by connecting the midpoints of the edges. Then, the characteristic mesh size of $\mathcal{T}_k$ is $h_k = 2^{1-k}$. For simplicity, all meshes in our experiments come from this family of meshes $\{\mathcal{T}_k\}$. Notice that each mesh in this family is totally decided by its size h_k. Therefore, in the remainder of this chapter, we use either the level k or the mesh size h_k to characterize each single mesh.

To observe the approximation behavior of the Arnold-Winther element, we consider the mixed problem (2.12) with homogeneous pure displacement boundary condition $\boldsymbol{u}|_{\partial\Omega} = \mathbf{0}$. The displacement field is set to be

$$\boldsymbol{u} = \begin{pmatrix} \sin(\pi x)\sin(\pi y) \\ \sin(\pi x)\sin(\pi y) \end{pmatrix}.$$

The Arnold-Winther element is used to discretize the problem. We solve the discrete problem (3.1) on each mesh $\mathcal{T}_k$, $k = 1, \ldots, 5$, and compare the discrete solution $(\boldsymbol{\sigma}_k, \boldsymbol{u}_k)$ with the nodal value interpolation, $(\tilde{\boldsymbol{\sigma}}_k, \tilde{\boldsymbol{u}}_k)$, of the exact solution $(\boldsymbol{\sigma}, \boldsymbol{u})$ on

$(\boldsymbol{\Sigma}(\mathcal{T}_k,\Omega), \boldsymbol{V}(\mathcal{T}_k,\Omega))$. In order to avoid introducing lower order approximation errors, the right hand side $(-\boldsymbol{f}, \boldsymbol{v})$ in (3.1) is calculated exactly instead of using numerical integration. By the Bramble-Hilbert lemma, the Schwarz Inequality and Theorem III.4, we expect the following approximation results:

$$\begin{aligned}
\|\tilde{\boldsymbol{\sigma}}_k - \boldsymbol{\sigma}_k\|_{0,\Omega} &\le ch_k^3, \\
\|\mathbf{div}\,\tilde{\boldsymbol{\sigma}}_k - \mathbf{div}\,\boldsymbol{\sigma}_k\|_{0,\Omega} &\le ch_k^2, \qquad (7.1)\\
\|\tilde{\boldsymbol{u}}_k - \boldsymbol{u}_k\|_{0,\Omega} &\le ch_k^2,
\end{aligned}$$

where c is a general constant independent of h_k. In Table 1, we give the numerical results, which appear to agree with (7.1). Notice that we even have $\|\mathbf{div}\,\tilde{\boldsymbol{\sigma}}_k - \mathbf{div}\,\boldsymbol{\sigma}_k\|_{0,\Omega} = 0$, which can be derived from the commutative diagram (3.9).

TABLE 1. Approximation behavior of the Arnold-Winther element.

k	1	2	3	4	5
$\|\tilde{\boldsymbol{\sigma}}_k - \boldsymbol{\sigma}_k\|_{0,\Omega}$	1.5875	0.2547	0.0337	0.0042	0.0005
$\|\mathbf{div}\,\tilde{\boldsymbol{\sigma}}_k - \mathbf{div}\,\boldsymbol{\sigma}_k\|_{0,\Omega}$	0	0	0	0	0
$\|\tilde{\boldsymbol{u}}_k - \boldsymbol{u}_k\|_{0,\Omega}$	0.5424	0.2664	0.0797	0.0208	0.0053

Next, we experiment with the overlapping Schwarz preconditioners. For simplicity, we only consider the pure traction boundary problem with homogeneous traction boundary condition $\boldsymbol{\sigma n}|_{\partial\Omega} = \mathbf{0}$. Both the mixed problem (3.1) and the $\boldsymbol{H}(\mathbf{div})$ problem (4.11) are solved by the preconditioned MINRES method. The overlapping Schwarz preconditioners $\mathbf{B}_a$ and $\mathbf{B}_m$ are used for Problem (4.11). The preconditioner $\mathcal{S}$, as defined in (4.9) with S_1 being $\mathbf{B}_a$ or $\mathbf{B}_m$ and S_2 being the identity operator, is used for Problem (3.1). For Problem (4.11) we set the exact solution to be

$$\boldsymbol{\sigma} = \begin{pmatrix} x(1-x) & 0 \\ 0 & y(1-y) \end{pmatrix}.$$

For Problem (3.1), we set the body force to be

$$\boldsymbol{g} = \begin{pmatrix} 1-3x^2 \\ 2y-1 \end{pmatrix},$$

which satisfies the compatibility condition (2.9). However, we do not know the exact solution for this problem.

We will report the condition number estimates for various meshes and sub-domains. For a symmetric positive definite matrix, we define the condition number as the ratio between the maximum eigenvalue and the minimum eigenvalue, while for a symmetric indefinite matrix, the condition number is defined as in (4.2). Since the 3-recurrence MINRES Algorithm is equivalent to the Lanczos procedure [48], the condition number estimate of a symmetric linear system can be calculated from its solving process using the MINRES Algorithm.

Following the notations in Chapter V, we denote H to be the size of the coarse mesh, h to be the size of the fine mesh and K to be the number of sub-domains. Recall that the overlapping sub-domain decomposition is obtained by extending a non-overlapping sub-domain decomposition by a distance of δ and the boundaries of the overlapping sub-domains have to align with the fine mesh. By the analysis in Chapter V, the condition numbers for $\mathbf{B}_a\boldsymbol{\Lambda}$ and $\mathbf{B}_m\boldsymbol{\Lambda}$ should depend on $\frac{H^4}{\delta^4}+1$ but not on h or K.

First, consider the case of fixed coarse mesh size $H = 1/2$ and sub-domains as shown in Figure 9. In this case, we have $K = 4$ and $\delta = 1/8$. In the experiment, we take fine mesh sizes to be $h = 1/8$, $1/16$, $1/32$ and report the condition number estimates of the unpreconditioned systems and the preconditioned systems for both Problem (4.11) and Problem (3.1) in Table 2. The condition number estimates using overlapping Schwarz preconditioners appear to be uniform with respect to h. We mention that the condition number estimates for Problem (3.1) and Problem (4.11) on the same mesh are not comparable since they are defined differently.

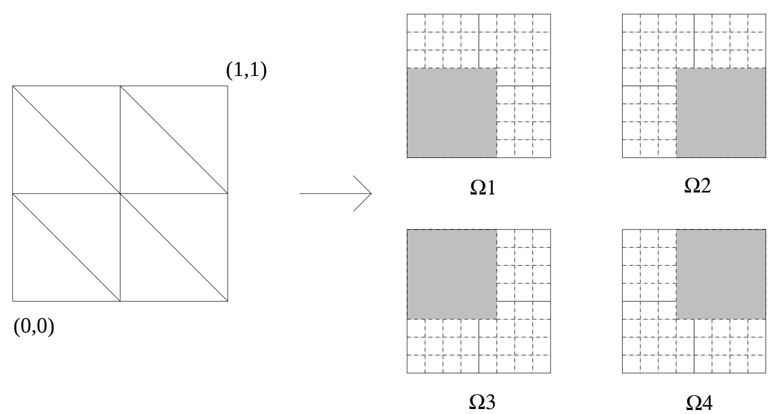

FIGURE 9. The coarse mesh and sub-domains with $H = 1/2$, $\delta = 1/8$ and $K = 4$.

TABLE 2. Condition number estimates with $H = 1/2$, $\delta = 1/8$ and $K = 4$.

h	$\boldsymbol{H}(\mathbf{div})$ problem (4.11)			Mixed problem (3.1)		
	No Prec.	Additive	Multiplicative	No Prec.	Additive	Multiplicative
1/8	3.4e+5	5.12	1.06	1.6e+3	5.78	1.86
1/16	1.4e+6	5.01	1.06	2.1e+3	5.66	1.88
1/32	5.5e+6	4.96	1.06	3.5e+3	5.19	1.89

For the $\boldsymbol{H}(\mathbf{div})$ problem (4.11), we also compute the condition number estimates for different K and h while the coarse mesh and the overlapping constant δ are fixed.

Set the coarse mesh size to be $H = 1/4$. We consider the cases of dividing Ω into $K = 4$, $K = 8$ and $K = 16$ non-overlapping rectangles whose boundary align with the coarse mesh $\mathcal{T}_H$. Extend each rectangle by a distance $\delta = 1/8$ to get overlapping decompositions of Ω. For example, in the case $K = 4$, the sub-domains will be the same as in Figure 9. The condition number estimates are given in Table 3 and they appear to be uniform with respect to both K and h.

TABLE 3. Condition number estimates for Problem (4.11) with $H = 1/4$ and $\delta = 1/8$.

h	Additive			Multiplicative		
	$K = 4$	$K = 8$	$K = 16$	$K = 4$	$K = 8$	$K = 16$
1/8	4.86	6.07	6.04	1.02	1.02	1.02
1/16	4.88	5.98	6.02	1.02	1.02	1.02

We also experimented on the case that both $H = 1/2$ and $K = 4$ are fixed, while δ and h change. The four sub-domains are obtained by dividing Ω into four equal squares and then extending each square by a distance $\delta = 1/4$, $1/8$, $1/16$ or $1/32$. In Table 4, the condition number estimates for various δ and h are given. Note that as suggested by Theorem V.1 and Lemma V.6, larger values of δ yields better preconditioners.

Finally, we report some numerical results for the multigrid preconditioner for Problem (4.11). Only pure displacement problems with homogeneous boundary conditions are considered. Random right hand sides are used in the experiments.

Consider the variable V-cycle multigrid preconditioner defined in Chapter VI with the number of smoothings satisfying $m_{k-1} = 2m_k$ and one smoothing on the finest level. Experiments show that the variable V-cycle multigrid preconditioner with

TABLE 4. Condition number estimates for Problem (4.11) with $K = 4$ and $H = 1/2$.

h	δ	Additive	Multiplicative
1/4	1/4	4.85	1.01
1/8	1/8	5.12	1.06
1/16	1/16	8.35	1.52
1/32	1/32	18.63	2.75

Richardson smoother, $\mathbf{B}_k^R$, does not work, as shown in Table 5. Here, the Richardson smoother is constructed by applying the Richardson method to the corresponding stiffness matrix $\underline{\underline{\Lambda}}_k$ on each level.

Next, we consider the variable V-cycle multigrid preconditioner $\mathbf{B}_k$ with the additive Schwarz smoother built on the vertex based subspaces, as defined in Chapter VI. The scaling factor ρ in Equation (6.2) is set to be $\frac{1}{3}$. We report the condition number estimates for $\mathbf{B}_k\mathbf{\Lambda}_k$ in Table 5, together with the condition number estimates for $\mathbf{B}_k^m\mathbf{\Lambda}_k$, where $\mathbf{B}_k^m$ is the variable V-cycle multigrid preconditioner using the multiplicative Schwarz smoother as discussed in Remark VI.1. These results indicate that $\mathbf{B}_k^m$ works better than $\mathbf{B}_k$, which is not surprising since multiplicative overlapping Schwarz methods have been observed to work better than additive overlapping Schwarz methods in numerous cases.

Further experiments also suggest that the V-cycle multigrid preconditioner $\mathbf{B}_k^V$ with the additive Schwarz smoother as in $\mathbf{B}_k$ is optimal for this test problem, as shown in Table 6, although we are still unable to explain that theoretically.

It is known that the multigrid preconditioner is more efficient than the overlapping Schwarz preconditioner in general. However, comparing the condition number estimates in Table 2 and Table 5, it appears that the multiplicative Schwarz preconditioner performs better than the multigrid preconditioner. This might be a result of large overlapping sizes, compared to the sub-domain size, in the overlapping Schwarz preconditioner and relatively small problems (the finest mesh is only $h = 1/32$).

TABLE 5. Condition number estimates for $\mathbf{\Lambda}_k$, $\mathbf{B}_k^R\mathbf{\Lambda}_k$, $\mathbf{B}_k\mathbf{\Lambda}_k$ and $\mathbf{B}_k^m\mathbf{\Lambda}_k$.

level	d.o.f.s	cond($\mathbf{\Lambda}_k$)	cond($\mathbf{B}_k^R\mathbf{\Lambda}_k$)	cond($\mathbf{B}_k\mathbf{\Lambda}_k$)	cond($\mathbf{B}_k^m\mathbf{\Lambda}_k$)
1	38	4.1507e+04			
2	115	1.5100e+05	1.7536e+04	4.52	3.10
3	395	6.0250e+05	7.0180e+04	4.49	3.19
4	1459	2.4105e+06	2.7784e+05	4.49	3.39
5	5603	9.6413e+06	1.0560e+06	4.45	3.41

TABLE 6. Condition number estimates for $\mathbf{B}_k^V\mathbf{\Lambda}_k$.

level	2	3	4	5
cond($\mathbf{B}_k^V\mathbf{\Lambda}_k$)	4.52	4.37	4.38	4.44

CHAPTER VIII

CONCLUSIONS

Our research provides a rigorous theoretical analysis on the mixed formulation (stress-displacement formulation) of the linear plane elasticity problem, its finite element discretization and some preconditioning techniques.

We prove the stability of the Arnold-Winther finite element spaces for pure traction boundary problems and mixed boundary problems under $\boldsymbol{H}^s$-regularity for $s > 1/2$. This generalizes the result given in [8], which is only for the pure displacement boundary problems. A Scott-Zhang type interpolation operator which preserves the homogeneous boundary condition $\boldsymbol{\tau n}|_{\Gamma_T} = 0$ is constructed and is essential to our proof of the stability of the Arnold-Winther finite element spaces.

Another goal was to develop efficient iterative solvers for the symmetric indefinite linear system which results from the Arnold-Winther finite element discretization of the mixed formulation. Several iterative solvers are discussed in Chapter IV. We concentrate on the preconditioned Minimum Residual Method and show that the problem of constructing a preconditioner for the mixed system can be reduced to the problem of constructing a preconditioner for the $\boldsymbol{H}(\mathbf{div})$ problem (4.11) on the Arnold-Winther finite element space. An overlapping Schwarz preconditioner and a multigrid preconditioner are developed for the $\boldsymbol{H}(\mathbf{div})$ problem and condition number estimates for the preconditioned systems are given in Chapter V and VI. For the overlapping Schwarz preconditioner, we prove that if the elasticity problem has $\boldsymbol{H}^s$-regularity for $s > 1/2$, then the condition number of the preconditioned system is independent of the fine mesh size h and the number of sub-domains K. For the multigrid preconditioner, we prove that if the elasticity problem has $\boldsymbol{H}^1$-regularity, then the condition number of the preconditioned system is independent of the number of levels k.

Since the operator $\boldsymbol{\Lambda}$ in the $\boldsymbol{H}(\mathbf{div})$ problem (4.11) is not uniformly elliptic, a Helmholtz-like decomposition is used in the analysis of the condition numbers for our preconditioners. The analysis is based on the observation that the space $\boldsymbol{H}_{0,\Gamma_T}(\mathbf{div}, \Omega, \mathbb{S}_2)$ can be decomposed into two orthogonal subspaces: the subspace of divergence free functions and its orthogonal compliment. As implied by Lemma II.4 and Lemma II.7, there exists a one-to-one correspondence between the divergence free subspace and the space $H^2_{0,\Gamma_T}(\Omega)/P_1(\Omega)$ via the Airy operator, while the orthogonal compliment of the divergence free subspace gains $\boldsymbol{H}^s$-regularity for $s > 1/2$. Similarly, on the discrete level, the finite element space $\boldsymbol{\Sigma}(\mathcal{T}, \Omega, \Gamma_T)$ can be decomposed into two orthogonal subspaces: the subspace of divergence free functions, where $\boldsymbol{\Lambda}$ behaves like an identity operator, and its orthogonal compliment, where $\boldsymbol{\Lambda}$ behaves like an ordinary second order differential operator. As shown in Lemma III.4, the divergence free subspace is related to the Argyris finite element space via the Airy operator. These two decompositions are essential to the analysis of our preconditioners.

There is still space for possible improvements to this research. For example, for the multigrid preconditioner, our analysis requires $\boldsymbol{H}^1$-regularity of the elasticity problem, which is only true if we have the pure displacement problem or the pure traction problem on a convex polygonal domain. There may be possible improvements to reduce this regularity requirement. Another example is that, as suggested by the numerical results in Chapter VII, the V-cycle multigrid also gives optimal convergence, although we are still unable to explain this theoretically. This is another direction for the future work.

REFERENCES

[1] R. A. Adams and J. F. Fournier. *Sobolev spaces.* Academic press, Boston, 2003.

[2] D. N. Arnold. Mixed finite element methods for elliptic problems. *Comput. Methods Appl. Mech. Engrg.*, **82**:281–300, 1990.

[3] D. N. Arnold, F. Brezzi, and J. J. Douglas. Peers: A new mixed finite element for plane elasticity. *Japan J. Appl. Math.*, **1**:347–367, 1984.

[4] D. N. Arnold, J. J. Douglas, and C. P. Gupta. A family of higher order mixed finite element methods for plane elasticity. *Numer. Math.*, **45**:1–22, 1984.

[5] D. N. Arnold and R. S. Falk. A new mixed formulation for elasticity. *Numer. Math.*, **53**:13–30, 1988.

[6] D. N. Arnold, R. S. Falk, and R. Winther. Preconditioning in H(div) and applications. *Math. Comp.*, **66**:957–984, 1997.

[7] D. N. Arnold, R. S. Falk, and R. Winther. Multigrid in H(div) and H(curl). *Numer. Math.*, **85**:197–217, 2000.

[8] D. N. Arnold and R. Winther. Mixed finite element for elasticity. *Numer. Math.*, **92**:401–419, 2002.

[9] D. N. Arnold and R. Winther. Nonconforming mixed elements for elasticity. *Math. Mod. Meth. Appl. Sci.*, **13**:295–307, 2003.

[10] K. J. Arrow, L. Hurwicz, and H. Uzawa. *Studies in linear and non-linear programming.* Stanford University Press, Stanford, 1958.

[11] C. Bacuta and J. H. Bramble. Regularity estimates for solutions of the equations of elasticity in convex plane polygonal domain. *preprint*, 2003.

[12] R. Bank, B. Welfert, and H. Yserentant. A preconditioning technique for indefinite systems resulting from mixed approximation of elliptic problems. *Numer. Math.*, **56**:645–666, 1990.

[13] J. H. Bramble. On the development of multigrid methods and their analysis. In *Proceedings of Symposia in Applied Mathematics* (W. Gautschi, eds.), Vol. 48, pages 5–19. Providence, Vancouver, 1994.

[14] J. H. Bramble, R. D. Lazarov, and J. E. Pasciak. Least-squares methods for linear elasticity based on a discrete negative norm. *Comput. Methods in Appl. Mech. Engrg.*, **152**:520–543, 2001.

[15] J. H. Bramble and J. E. Pasciak. A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems. *Math. Comp.*, **50**:1–17, 1988.

[16] J. H. Bramble and J. E. Pasciak. A domain decomposition technique for Stokes problems. *Appl. Numer. Math.*, **6**:251–261, 1990.

[17] J. H. Bramble, J. E. Pasciak, and A. T. Vassilev. Analysis of the inexact Uzawa algorithm for saddle point problems. *SIAM J. Numer. Anal.*, **34**:1072–1092, 1997.

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

[19] J. H. Bramble and X. Zhang. Handbook of numerical analysis. In *Handbook of numerical analysis* (P. Ciarlet and J.-L. Lions, eds.), Vol. VII, pages 173–415. North-Holland, Amsterdam, 2000.

[20] A. Brandt. Multilevel computations: Review and recent developments. In *Multigrid methods, Proceedings of the Third Copper Mountain Conference* (S. McCormick, ed.), pages 35–62. Marcel Dekker, New York, 1988.

[21] S. C. Brenner and L. R. Scott. *The mathematical theory of finite element methods*. Springer-Verlag, New York, 1994.

[22] F. Brezzi and M. Fortin. *Mixed and hybrid finite element methods*. Springer-Verlag, New York, 1991.

[23] Z. Cai and G. Starke. First-order system least squares for the stress-displacement formulation: linear elasticity. *SIAM J. Numer. Anal.*, **41**:715–730, 2003.

[24] Z. Cai and G. Starke. Least-squares methods for linear elasticity. *SIAM J. Numer. Anal.*, To appear.

[25] P. G. Ciarlet. *The finite element method for elliptic problems*. North-Holland, Amsterdam, 1978.

[26] P. Clément. Approximation by finite element functions using local regularization. *RAIRO Anal. Numér*, **R-2**:77–84, 1975.

[27] M. Dryja and O. B. Widlund. An additive variant of the Schwarz alternating method for the case of many subregions. *Tech. report, Courant Institute of Mathematical Sciences, New York*, 1987.

[28] M. Dryja and O. B. Widlund. Multilevel additive methods for elliptic finite element problems. *Tech. report, Courant Institute of Mathematical Sciences, New York*, 1990.

[29] M. Dryja and O. B. Widlund. Domain decomposition algorithms with small overlap. *SIAM J. Sci. Comput.*, **15**:604–620, 1994.

[30] T. Dupont and R. Scott. Polynomial approximation of functions in Sobolev spaces. *Math. Comp.*, **34**:441–463, 1980.

[31] H. C. Elman and G. H. Golub. Inexact and preconditioned Uzawa algorithms for saddle point problems. *SIAM J. Numer. Anal.*, **31**:1645–1661, 1994.

[32] R. E. Ewing and J. Wang. Analysis of the Schwarz algorithm for mixed finite element methods. *RAIRO Math. Modél. Anal. Numér.*, **26**:739–756, 1992.

[33] B. Fischer. *Polygonal based iteration methods for symmetric linear systems.* Wiley-Teubner, Leipzig, 1996.

[34] L. P. Franca, T. J. Hughes, A. F. Loula, and I. Miranda. A new family of stable elements for nearly incompressible elasticity based on a mixed Petrov-Galerkin finite element formulation. *Numer. Math.*, **53**:123–141, 1988.

[35] V. Girault and P. A. Raviart. *Finite Element Methods for Navier-Stokes equations.* Number 5 in Springer Series in Computational Mathematics. Springer-Verlag, New York, 1986.

[36] A. Greenbaum. *Iterative methods for solving linear systems.* SIAM, Philadelphia, 1997.

[37] P. Grisvard. *Elliptic problems in nonsmooth domains.* Pitman, Boston, 1985.

[38] P. Grisvard. Problémes aux limites dans les polygons. *Mode d'emploi. Bulletin de la Direction des Etudes et Recherches Series C Mathematiques, Informatique*, **1**:21–59, 1986.

[39] P. Grisvard. Singularités en elasticité. *Arch. Rat. Mech. Anal.*, **107**:157–180, 1989.

[40] P. Grisvard. *Singularities in boundary value problems.* Number 22 in Research notes in applied methematics. Springer-Verlag, New York, 1992.

[41] W. Hackbusch. *Multi-grid methods and applications.* Springer-Verlag, Berlin, 1985.

[42] R. Hiptmair. Multigrid method for H(div) in three dimensions. *ETNA*, **6**:133–152, 1997.

[43] R. Hiptmair and A. Toselli. Overlapping and multilevel Schwarz methods for vector valued elliptic problems in three dimensions. In *Parallel Solution of Partial Differential Equations* (P. B. und M. Luskin, eds), IMA Volumes in Mathematics und its Applications, pages 181–208. Springer-Verlag, Berlin, 2000.

[44] C. Johnson and B. Mercier. Some equilbrium finite element methods for two-dimensional elasticity problems. *Numer. Math.*, **30**:103–116, 1978.

[45] J. Lions and E. Magenes. *Non-homogeneous boundary value problems and applications*, Vol. I. Springer-Verlag, New York, 1972.

[46] M. F. Murphy, G. H. Golub, and A. J. Wathen. A note on preconditioning for indefinite linear systems. *SIAM J. Sci. Comput.*, **21**:1969–1972, 2000.

[47] J. Nečas. *Les Méthodes directes en théorie des équations elliptiques.* Masson, Paris, 1967.

[48] C. C. Paige and M. A. Saunders. Solution of sparse indefinite systems of linear equations. *SIAM J. Numer. Anal.*, **12**:617–629, 1975.

[49] T. Rusten and R. Winther. A preconditioned iterative method for saddlepoint problem. *SIAM J. Matrix Anal. Appl.*, **13**:887–904, 1992.

[50] L. R. Scott and S. Zhang. Finite element interpolation of nonsmooth function satisfying boundary conditions. *Math. Comp.*, **54**:483–493, 1990.

[51] W. S. Slaughter. *The linearized theory of elasticity.* Birkhāuser, Boston, 2002.

[52] B. F. Smith, P. E. Bjorstad, and W. D. Gropp. *Domain decomposition.* Cambridge University Press, Cambridge, 1996.

[53] J. Xu. Iterative methods by space decomposition and subspace correction. *SIAM Rev.*, **34**:581–613, 1992.

[54] X. Zhang. Multilevel Schwarz methods. *Numer. Math.*, **63**:521–539, 1992.

[55] W. Zulehner. Analysis of iterative methods for saddle point problems: a unified approach. *Math. Comp.*, **71**:479–505, 2001.

APPENDIX A
THE SCOTT-ZHANG TYPE OPERATOR R_H

We follow the idea in [50] in the construction of the Scott-Zhang type operator R_h. Let $\mathcal{T}$ be a quasi-uniform triangulation for the polygonal domain Ω with characteristic mesh size h. Define the finite element space of C^0-quadratics by

$$X_h = \{v \in C^0(\Omega) \text{ such that } v|_T \in P_2(T) \text{ for all } T \in \mathcal{T}\}.$$

The degrees of freedom are the nodal values on each vertex and the center point of each edge. We call those points "nodes". Let $\mathcal{N} = \{x_i\}_{i=1}^N$ be the set of all "nodes" and ϕ_i, $1 \leq i \leq N$, be the corresponding nodal basis function on x_i. Let $\mathcal{E}$ be the set of all edges of the triangles in $\mathcal{T}$.

We consider each triangle $T \in \mathcal{T}$ as a closed set which includes its boundary. We also consider each edge $e \in \mathcal{E}$ as a closed line segment which includes its ends. For each $T \in \mathcal{T}$, define

$$S_T = \cup\{T_i \in \mathcal{T} \mid T_i \cap T \neq \emptyset\}.$$

Since $\mathcal{T}$ is quasi-uniform, there are only a finite number of triangles in each S_T and $\{S_T\}_{T\in\mathcal{T}}$ satisfies the finite overlapping condition.

For each node $x_i \in \mathcal{N}$, we define a simplex K_i by the following:

- if $x_i \notin \overline{\Gamma_T}$, we choose $K_i = T$ where $T \in \mathcal{T}$ is any triangle such that $x_i \in T$ (the choice may not be unique);

- if $x_i \in \overline{\Gamma_T}$, we choose $K_i = e$ where $e \in \mathcal{E}$ is any edge such that $e \subset \overline{\Gamma_T}$ and $x \in e$ (the choice may not be unique).

Let $x_{i,j}$, $j = 1, \ldots, n_i$, be the nodes in K_i and assume that they are arranged in a way such that $x_{i,1} = x_i$. Let $\phi_{i,j}$, $j = 1, \ldots, n_i$, be the restriction to K_i of the basis function associated with the node $x_{i,j}$. Let $\{\psi_{i,j}\}_{j=1}^{n_i}$ be the $L^2(K_i)$-dual basis for $\{\phi_{i,j}\}_{j=1}^{n_i}$, i.e.

$$\int_{K_i} \phi_{i,j}\psi_{i,k}\, d\boldsymbol{x} = \delta_{jk} = \begin{cases} 1, & j = k, \\ 0, & j \neq k. \end{cases} \tag{A.1}$$

We define $\psi_i = \psi_{i,1}$. Notice that ψ_i is only defined on K_i, while ϕ_i is defined on Ω. The following lemma and its proof can be found in [50].

Lemma VIII.1. *For $i = 1, \ldots, N$, there exist $c > 0$ and $C > 0$ independent of h such that*

$$\|\phi_i\|_{0,\Omega} \leq ch, \tag{A.2}$$

$$\|\psi_i\|_{0,K_i} \leq Ch^{-\dim(K_i)/2}. \tag{A.3}$$

For $s > 1/2$, define $R_h : H^s(\Omega) \to X_h$ by

$$R_h v = \sum_{i=1}^{N} \phi_i \int_{K_i} \psi_i v\, d\boldsymbol{x}.$$

Clearly, R_h is well defined and preserves the homogeneous boundary condition on Γ_T. It is not hard to see that $R_h v = v$ for all $v \in X_h$.

Next, we prove the stability and the approximability for R_h.

Lemma VIII.2. *Let $1/2 < s \leq 1$ be a real number. For $v \in H^s(\Omega)$, we have*

$$\|R_h v\|_{0,T} \leq c(\|v\|_{0,S_T} + h^s|v|_{s,S_T}),$$

where c is a positive constant independent of h and T.

Proof. By a scaling argument and the Schwartz inequality,

$$\begin{aligned}
\|R_h v\|_{0,T} &\leq \sum_{x_i \in T} \left| \int_{K_i} \psi_i v \, d\boldsymbol{x} \right| \|\phi_i\|_{0,T} \\
&\leq ch \sum_{x_i \in T} \|\psi_i\|_{0,K_i} \|v\|_{0,K_i} \\
&\leq ch \sum_{x_i \in T} h^{-dim(K_i)/2} \|v\|_{0,K_i}.
\end{aligned}$$

If K_i is an edge, then K_i is an edge of a triangle $T_i \in \mathcal{T}$ such that $T_i \subset S_T$. Let $\hat{T}$ be a reference triangle and $\hat{K}$ be the corresponding edge of K_i in $\hat{T}$. By the trace theorem and a scaling argument,

$$h^{-dim(K_i)/2} \|v\|_{0,K_i} = h^{-1/2} \|v\|_{0,K_i} \quad \leq c(h^{-1} \|v\|_{0,T_i} + h^{s-1} |v|_{s,T_i}).$$

If K_i is a triangle $T_i \in \mathcal{T}$, clearly $T_i \subset S_T$. We have

$$h^{-dim(K_i)/2} \|v\|_{0,K_i} = h^{-1} \|v\|_{0,T_i}.$$

Combining all the above, we have

$$\|R_h v\|_{0,T} \leq ch \sum_{T \subset S_T} (h^{-1} \|v\|_{0,T} + h^{s-1} |v|_{s,T}) \leq c(\|v\|_{0,S_T} + h^s |v|_{s,S_T}).$$

□

Lemma VIII.3. *Let s and s' be two real numbers which satisfy $0 \leq s' \leq s \leq 3$. For all $v \in H^s(S_T)$, there exists $\xi \in P_2(S_T)$ such that*

$$\|v - \xi\|_{s',S_T} \leq ch^{s-s'} |v|_{s,S_T},$$

where c is a positive constant independent of h and T.

Proof. The result comes from the generalization of the Bramble-Hilbert lemma [30] and interpolation of Sobolev spaces. □

Lemma VIII.4. *Let $s \in (1/2, 3]$ be a real number and t satisfies $0 \le t \le min(s, 1)$. For $v \in H^s(\Omega)$, we have*

$$\|v - R_h v\|_{t,T} \le ch^{s-t}|v|_{s,S_T},$$

where c is a positive constant independent of h and T.

Proof. For $v \in H^s(\Omega)$, let ξ be defined as in Lemma VIII.3. Then

$$\begin{aligned}
\|v - R_h v\|_{t,T} &\le \|v - \xi\|_{t,T} + \|R_h(v - \xi)\|_{t,T} \\
&\le \|v - \xi\|_{t,T} + ch^{-t}(\|v - \xi\|_{0,S_T} + h^{\min(s,1)}|v - \xi|_{\min(s,1),S_T}) \\
&\le ch^{s-t}|v|_{s,S_T}.
\end{aligned}$$

□

VITA

Yanqiu Wang was born in Jiangdu, China on August 29, 1977 to Liulin Wang and Tongxiang Liu. She earned her B.S. degree in mathematics from Fudan University, China, in July 1997. She continued on her graduate study on numerical analysis in Fudan University, China and received the M.S. degree in mathematics, in July 2000. In the fall of 2000, she began her studies in the Department of Mathematics at Texas A&M University, and received her Ph.D. degree in August 2004.

Yanqiu Wang can be contacted by writing to the Department of Mathematics, Texas A&M University, College Station, TX 77843-3368, USA, or to the e-mail address: yqwang@math.tamu.edu.