A MULTIGRID METHOD FOR THE PSEUDOSTRESS FORMULATION OF STOKES PROBLEMS

ZHIQIANG CAI[*] AND YANQIU WANG[*]

Abstract. The purpose of this paper is to develop and analyze a multigrid solver for the finite element discretization of the pseudostress system associated with the differential operator $\mathcal{A} - \gamma \, \mathbf{grad} \, \mathbf{div}$ over 2×2 matrix-valued functions. This system is derived from the pseudostress-velocity formulation [11] of two-dimensional Stokes problems through the penalty method or natural time discretization for the respective steady or unsteady state problems. Here $\gamma > 0$ is a constant associated with either the penalty parameter or the time-step size and $\mathcal{A}$ is a singular map. In this paper, we develop a multigrid solver for the discrete problem using both the RT and the BDM finite elements. We show that the multigrid convergence rate is $O(\frac{1+\gamma^{-2}}{1+\gamma^{-2}+m})$, where m is the number of smoothings. This convergence rate is independent of the mesh size and the number of levels used in multigrid. Moreover, numerical results presented in this paper show that the multigrid convergence rate for the BDM elements does not depend on the parameter γ.

Key words. multigrid, mixed finite element, pseudostress-velocity formulation, Stokes problem

AMS subject classifications. 65N30, 65N55

1. Introduction. The original physical equations for incompressible Newtonian flows are the stress-velocity-pressure formulation [7, 15]. Recently, this formulation is gaining consistent attention because of the arising interest in non-Newtonian flows [6, 21]. For non-Newtonian flows, the stress can not be eliminated and, hence, a formulation containing the stress as independent variables is unavoidable. In addition, accurate direct calculation of the stress is important for computing the traction on a fluid-solid interface. Although the stress can be obtained in the velocity-pressure method by differentiating the velocity, this degrades the order of the approximation.

In this paper we are interested in mixed finite element methods for the Stokes problem. Notice that one disadvantage of using the stress-velocity-pressure formulation is that the symmetry constraint of the stress tensor poses extra difficulty in the discretization. For example, stable conforming mixed elements for symmetric metrices, e.g., [3, 5, 20], employ many degrees of freedom and, consequently, lead to large linear systems even for coarse meshes. There are different ways to avoid this difficulty. One way is to use the idea of pseudostress [11] and a pseudostress-velocity formulation. The pseudostress is non-symmetric and, hence, any stable pairs for the Darcy flows can be adapted to the pseudostress-velocity formulation. This is documented in [13]. Furthermore, the velocity can be eliminated through the penalty method for the steady-state problem or natural time discretization for the unsteady-state problem. So one only needs to solve numerically the pseudostress system associated with a differential operator $\mathcal{A} - \gamma \, \mathbf{grad} \, \mathbf{div}$ over 2×2 matrix-valued functions. Here, $\gamma > 0$ is a constant depending on either the penalty parameter or the time-step size, and $\mathcal{A}$ is a projection operator whose kernel intersects the kernel of the divergence operator. The purpose of this paper is to construct and analyze a multigrid solver for this pseudostress system. It must be pointed out that this problem is more delicate to handle than a normal $\boldsymbol{H}(\mathbf{div})$ problem because of the complicated kernel structure.

We also mention other methods to circumvent the difficulty posed by the symmetric stress. One is to use stabilization technique [7]. Another is to use the least-squares method, which do not require stable pairs of finite elements. Least-squares methods with fast multigrid solvers for the stress-velocity-pressure formulation and its variants, stress-velocity and pseudostress-velocity formulations, were studied in [11].

Multigrid methods are efficient for many problems. A vast amount of research has been done in this area, e.g., [8, 18, 22, 25]. However, the classical techniques for the multigrid method do not work for the $\boldsymbol{H}(\mathbf{div})$-type problem since standard smoothers can not damp out the highly oscillating part of divergence free components. Hence, special smoothers have to be developed in order to achieve optimal multigrid convergence rate [1, 2, 19, 24]. An additional difficulty in the multigrid analysis for the pseudostress system is the handling of the kernel of $\mathcal{A}$. To overcome the above difficulties, we employ a Helmholtz-like decomposition of the discrete space, which also utilizes the operator $\mathcal{A}$, so that the analysis can be done on separate subspaces. In this paper we consider multigrid solver for systems discretized using both the RT and the BDM finite elements. We show that the multigrid convergence rate is bounded above by $\rho = \frac{C(1+\gamma^{-2})}{C(1+\gamma^{-2})+2m} < 1$ where m is the number of smoothings and C is a constant independent of the mesh size and the number of levels used in multigrid. When $\gamma >> 1$, which is always true for the penalty method of the stationary problems as the penalty parameter goes to zero, the multigrid convergence for the pseudostress system can be considered independent of γ and, hence, the penalty parameter. When $\gamma << 1$, which corresponds to the small time-step size in the time-discretization of unsteady-state problems, ρ is close to one. Nevertheless, numerical results presented in this paper show that the multigrid convergence rate for the BDM elements does not depend on the parameter γ even if γ is small.

We give the outline of the paper as follows. In sections 2 and 3, we describe the pseudostress system of the Stokes problem and its finite element discretization. The key Helmholtz-type decomposition of the finite element space is given in Section 4. We construct a multigrid solver for the discrete $\boldsymbol{H}(\mathbf{div})$-type problem and give a convergence rate estimate in Section 5. Some supporting numerical results are presented in section 6. Finally, in Appendix A, we prove an essential inequality needed by the analysis in Section 4.

1.1. Notation. Let Ω be a convex polygon. Denote by $\mathbb{R}^2$ the set of two-dimensional vector-valued functions and by $\mathbb{M}_2$ the set of 2×2 matrix-valued functions. Throughout the paper, we adopt the convention that a Greek character denotes a 2×2 matrix and a bold Latin character in lower case denotes a vector. Let $\boldsymbol{\tau} = (\tau_{ij})_{2 \times 2} \in \mathbb{M}_2$ and $\boldsymbol{v} = (v_1, v_2)^t \in \mathbb{R}^2$, define

$$\mathrm{div}\,\boldsymbol{v} = \frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y}, \qquad \mathbf{div}\boldsymbol{\tau} = \begin{pmatrix} \frac{\partial \tau_{11}}{\partial x} + \frac{\partial \tau_{12}}{\partial y} \\ \frac{\partial \tau_{21}}{\partial x} + \frac{\partial \tau_{22}}{\partial y} \end{pmatrix},$$

$$\nabla \boldsymbol{v} = \begin{pmatrix} \frac{\partial v_1}{\partial x} & \frac{\partial v_1}{\partial y} \\ \frac{\partial v_2}{\partial x} & \frac{\partial v_2}{\partial y} \end{pmatrix}, \qquad \mathbf{curl}\,\boldsymbol{v} = \begin{pmatrix} -\frac{\partial v_1}{\partial y} & \frac{\partial v_1}{\partial x} \\ -\frac{\partial v_2}{\partial y} & \frac{\partial q_2}{\partial x} \end{pmatrix}.$$

Define the inner product between vectors and 2×2 matrices by:

$$\boldsymbol{u} \cdot \boldsymbol{v} = u_1\,v_1 + u_2\,v_2 \quad \text{and} \quad \boldsymbol{\sigma} : \boldsymbol{\tau} = \sum_{1 \le i,j \le 2} \sigma_{ij}\,\tau_{ij},$$

respectively. We use the usual notation $L^2(\Omega)$ to denote the set of square integrable functions on Ω and $H^s(\Omega)$, where $s > 0$ is a real number, to denote the Sobolev space on Ω with index s. Let $C_0^\infty(\Omega)$ be the space of infinitely differentiable functions on Ω with compact support. Denote by $H_0^s(\Omega)$ the closure of $C_0^\infty(\Omega)$ under the norm of $H^s(\Omega)$ and define $H^{-s}(\Omega) = (H_0^s(\Omega))'$, the dual space of $H_0^s(\Omega)$. It is natural to extend the above definitions to vector- and matrix-valued functions using product spaces. For example, $\boldsymbol{H}^s(\Omega, \mathbb{R}^2)$ and $\boldsymbol{H}^s(\Omega, \mathbb{M}_2)$ denote the Sobolev spaces over the set of vector- and 2×2 matrix-valued functions, respectively. Other notations, such as $\mathbf{L}^2(\Omega, \mathbb{R}^2)$, $\mathbf{L}^2(\Omega, \mathbb{M}_2)$, are defined in the same fashion. For simplicity, denote $\| \cdot \|_{s,\Omega}$ the Sobolev norm with index s over scalar-, vector-, or matrix-valued functions, depending on the type of the function. Similarly, denote by $| \cdot |_{s,\Omega}$ the Sobolev semi-norm and by $(\cdot, \cdot)$ the $\mathbf{L}^2$ inner product. Define the space

$$\boldsymbol{H}(\mathbf{div}, \Omega, \mathbb{M}_2) = \{\boldsymbol{\sigma} \in \mathbf{L}^2(\Omega, \mathbb{M}_2) \,|\, \mathbf{div}\,\boldsymbol{\sigma} \in \mathbf{L}^2(\Omega, \mathbb{R}^2)\}$$

with the norm $\|\boldsymbol{\sigma}\|^2_{\boldsymbol{H}(\mathbf{div},\Omega,\mathbb{M}_2)} = (\boldsymbol{\sigma}, \boldsymbol{\sigma}) + (\mathbf{div}\boldsymbol{\sigma}, \mathbf{div}\boldsymbol{\sigma})$.

2. Pseudostress System. Consider the two-dimensional steady-state Stokes problem

$$\begin{cases} -\Delta \boldsymbol{u} + \nabla p = -\boldsymbol{f} & \text{in } \Omega, \\ \operatorname{div} \boldsymbol{u} = 0 & \text{in } \Omega, \\ \boldsymbol{u}|_{\partial\Omega} = \mathbf{0}, \end{cases} \tag{2.1}$$

where $\boldsymbol{u}$ is the velocity and p is the pressure which satisfies the compatibility condition $\int_\Omega p \, d\boldsymbol{x} = 0$.

Let $\mathcal{A} : \mathbb{M}_2 \to \mathbb{M}_2$ be a fourth order tensor defined by

$$\mathcal{A}\boldsymbol{\tau} = \boldsymbol{\tau} - \frac{1}{2}(\operatorname{tr}\boldsymbol{\tau})\,\mathbf{I},$$

where $\operatorname{tr}\boldsymbol{\tau} = \tau_{11} + \tau_{22}$ is the trace of $\boldsymbol{\tau}$ and $\mathbf{I}$ is the 2×2 identity matrix. We immediately notice that $\mathcal{A}$ is a projection onto the trace-free subspace of $\mathbb{M}_2$ and $Ker(\mathcal{A}) = \{v\,\mathbf{I}\,|\,v \text{ is a scalar function}\}$. Define the pseudostress by

$$\boldsymbol{\sigma} = -p\,\mathbf{I} + \nabla\boldsymbol{u}.$$

Then the Stokes problem (2.1) can be rewritten in the pseudostress-velocity formulation:

$$\begin{cases} \mathbf{div}\boldsymbol{\sigma} = \boldsymbol{f} & \text{in } \Omega, \\ \mathcal{A}\boldsymbol{\sigma} - \nabla\boldsymbol{u} = \mathbf{0} & \text{in } \Omega, \\ \boldsymbol{u}|_{\partial\Omega} = \mathbf{0}. \end{cases} \tag{2.2}$$

The incompressible constraint $\operatorname{div} \boldsymbol{u} = 0$ is enforced through $\operatorname{div} \boldsymbol{u} = \operatorname{tr}(\nabla\boldsymbol{u}) = 0$. Notice that $\operatorname{tr}\boldsymbol{\sigma} = -2p$ also needs to satisfy the compatibility condition $\int_\Omega \operatorname{tr}\boldsymbol{\sigma}\,d\boldsymbol{x} = 0$.

From the well-known regularity estimate of problem (2.1) on convex polygons and the definition of the pseudostress, we have the following regularity results. Let $\boldsymbol{f} \in \mathbf{L}^2(\Omega)$, then the solutions to problem (2.1) and (2.2) satisfy $\boldsymbol{u} \in \boldsymbol{H}^2(\Omega) \cap \boldsymbol{H}_0^1(\Omega)$, $p \in H^1(\Omega)/\mathbb{R}$, $\boldsymbol{\sigma} \in \boldsymbol{H}^1(\Omega, \mathbb{M}_2)$ and

$$\|\boldsymbol{u}\|_{2,\Omega} \le c\|\boldsymbol{f}\|_{0,\Omega}, \quad \|p\|_{H^1(\Omega)/\mathbb{R}} \le c\|\boldsymbol{f}\|_{0,\Omega}, \quad \|\boldsymbol{\sigma}\|_{1,\Omega} \le c\|\boldsymbol{f}\|_{0,\Omega}, \tag{2.3}$$

where c is a positive constant independent of $\boldsymbol{f}$.

Denote $\mathbf{V} = \mathbf{L}^2(\Omega, \mathbb{R}^2)$ and

$$\boldsymbol{\Sigma} = \boldsymbol{H}(\mathbf{div}, \Omega, \mathbb{M}_2)/span\{\mathbf{I}\} = \{\boldsymbol{\tau} \in \boldsymbol{H}(\mathbf{div}, \Omega, \mathbb{M}_2) \,|\, \int_\Omega \mathrm{tr}\,\boldsymbol{\tau}\,d\boldsymbol{x} = 0\}.$$

Then we have the following variational form for Equation (2.2). *Given $\boldsymbol{f} \in \mathbf{V}$, find $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}$ and $\boldsymbol{u} \in \mathbf{V}$ such that*

$$\begin{cases} (\mathcal{A}\boldsymbol{\sigma}, \boldsymbol{\tau}) + (\mathbf{div}\boldsymbol{\tau}, \boldsymbol{u}) = 0, & \text{for all } \boldsymbol{\tau} \in \boldsymbol{\Sigma}, \\ (\mathbf{div}\boldsymbol{\sigma}, \boldsymbol{v}) = (\boldsymbol{f}, \boldsymbol{v}), & \text{for all } \boldsymbol{v} \in \mathbf{V}. \end{cases} \quad (2.4)$$

Notice that the velocity boundary condition becomes the nature boundary condition in the mixed formulation. According to [10], the existence and uniqueness of Problem (2.4) follows from the well known inf-sup condition $\sup_{\boldsymbol{\tau} \in \boldsymbol{H}^1(\Omega, \mathbb{M}_2)} \frac{(\mathbf{div}\boldsymbol{\tau}, \boldsymbol{v})}{\|\boldsymbol{\tau}\|_{1,\Omega}} \geq \kappa \|\boldsymbol{v}\|_{0,\Omega}$ for all $\boldsymbol{v} \in \mathbf{V}$, and the following lemma [12]:

LEMMA 2.1. *For all $\boldsymbol{\tau} \in \boldsymbol{\Sigma}$, we have*

$$\|\boldsymbol{\tau}\|_{0,\Omega}^2 \leq C_A(\|\mathcal{A}^{1/2}\boldsymbol{\tau}\|_{0,\Omega}^2 + \|\mathbf{div}\boldsymbol{\tau}\|_{-1,\Omega}^2),$$

where $C_A \geq 1$ is independent of $\boldsymbol{\tau}$.

Using the penalty method [10] on Problem (2.4) gives a new system

$$\begin{cases} (\mathcal{A}\boldsymbol{\sigma}_\gamma, \boldsymbol{\tau}) + (\mathbf{div}\boldsymbol{\tau}, \boldsymbol{u}_\gamma) = 0, & \text{for all } \boldsymbol{\tau} \in \boldsymbol{\Sigma}, \\ (\mathbf{div}\boldsymbol{\sigma}_\gamma, \boldsymbol{v}) - \frac{1}{\gamma}(\boldsymbol{u}_\gamma, \boldsymbol{v}) = (\boldsymbol{f}, \boldsymbol{v}), & \text{for all } \boldsymbol{v} \in \mathbf{V}. \end{cases} \quad (2.5)$$

where γ is a positive constant which is usually very large. One advantage of using the penalty method is that the mixed system (2.5) can be easily decoupled. Define a bilinear form over $\boldsymbol{H}(\mathbf{div}, \Omega, \mathbb{M}_2)$ by:

$$\boldsymbol{\Lambda}(\boldsymbol{\sigma}, \boldsymbol{\tau}) = (\mathcal{A}\boldsymbol{\sigma}, \boldsymbol{\tau}) + \gamma(\mathbf{div}\boldsymbol{\sigma}, \mathbf{div}\boldsymbol{\tau}) \quad \text{for all } \boldsymbol{\sigma}, \, \boldsymbol{\tau} \in \boldsymbol{\Sigma}.$$

By eliminating $\boldsymbol{u}_\gamma$, system (2.5) can be reduced to the following $\boldsymbol{H}(\mathbf{div})$ problem:

$$\boldsymbol{\Lambda}(\boldsymbol{\sigma}_\gamma, \boldsymbol{\tau}) = \gamma(\boldsymbol{f}, \mathbf{div}\boldsymbol{\tau}), \quad \text{for all } \boldsymbol{\tau} \in \boldsymbol{\Sigma}. \quad (2.6)$$

By Lemma 2.1, it is easy to see that $\boldsymbol{\Lambda}(\cdot, \cdot)$ is coercive over space $\boldsymbol{\Sigma}$. Hence problem (2.6) is well-posed. Then $\boldsymbol{u}_\gamma$ can be easily calculated through $\boldsymbol{u}_\gamma = \gamma(\mathbf{div}\boldsymbol{\sigma}_\gamma - \boldsymbol{f})$.

Finally, we mention about the case of time-dependent Stokes problems. It is not hard to see that the implicit time-discretization usually leads to the following pseudostress-velocity form:

$$\begin{cases} \mathbf{div}\boldsymbol{\sigma} - \frac{1}{\gamma}\boldsymbol{u} = \boldsymbol{f} & \text{in } \Omega, \\ \mathcal{A}\boldsymbol{\sigma} - \nabla\boldsymbol{u} = \mathbf{0} & \text{in } \Omega, \\ \boldsymbol{u}|_{\partial\Omega} = \mathbf{0}. \end{cases} \quad (2.7)$$

where $\gamma \ll 1$ is related to the time-step size. Again, problem (2.7) can easily be decoupled and will lead to the pseudostress problem (2.6). The only difference is that γ is small in this case.

3. Finite Element Approximation. In the remainder of this paper, we will focus on solving Problem (2.6) using the finite element method and a multigrid solver. To discretize the problem, we need a good finite element approximation to the space $\boldsymbol{\Sigma}$. There are several well-known conforming elements for the $H(div)$ space over the vector field $\mathbb{R}^2$. We will consider the lowest order RT element [23] and the lowest order BDM element [9]. It should be easy to extend the results to higher order elements.

Let $\mathcal{T}_h$ be a quasi-uniform triangulation of Ω with characteristic mesh size h. On each triangle $T \in \mathcal{T}$, denote $P_n(T)$ to be the set of all polynomials with degrees less than or equal to n. Define $RT_T = span\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right\}$ and $BDM_T = (P_1(T))^2$. The degrees of freedom for the RT element are the zeroth order moments of the normal components on each edge of T. The degrees of freedom for the BDM element are the zeroth order and the first order moments of the normal components on each edge of T.

We use two copies of RT/BDM element to approximate the space $\boldsymbol{H}(\mathbf{div}, \Omega, \mathbb{M}_2)$. Let

$$\boldsymbol{\Sigma}_T^{RT} = \{\boldsymbol{\tau} \in \mathbb{M}_2 \,|\, (\tau_{i1}, \tau_{i2}) \in RT_T \text{ for } i = 1, 2\},$$

$$\boldsymbol{\Sigma}_T^{BDM} = \{\boldsymbol{\tau} \in \mathbb{M}_2 \,|\, (\tau_{i1}, \tau_{i2}) \in BDM_T \text{ for } i = 1, 2\}.$$

Define the finite element spaces

$$\boldsymbol{\Sigma}_h^{RT} = \{\boldsymbol{\tau} \in \boldsymbol{\Sigma} \,|\, \boldsymbol{\tau}|_T \in \boldsymbol{\Sigma}_T^{RT}\} \quad \text{and} \quad \boldsymbol{\Sigma}_h^{BDM} = \{\boldsymbol{\tau} \in \boldsymbol{\Sigma} \,|\, \boldsymbol{\tau}|_T \in \boldsymbol{\Sigma}_T^{BDM}\}.$$

In later analysis, we use more of the common properties of the RT and the BDM elements than of their differences. Hence we will use $\boldsymbol{\Sigma}_h$ to denote either $\boldsymbol{\Sigma}_h^{RT}$ or $\boldsymbol{\Sigma}_h^{BDM}$. It is well known that the necessary continuity requirement for $\boldsymbol{\Sigma}_h$ to be in $\boldsymbol{\Sigma}$ is that any function in the discrete space should have continuous normal components across each edge of the mesh. Notice that $\boldsymbol{\Sigma}_h$ also inherits the constraint $\int_\Omega \operatorname{tr} \boldsymbol{\sigma} \, d\boldsymbol{x} = 0$ from $\boldsymbol{\Sigma}$.

Consider the discrete problem of (2.6): find $\boldsymbol{\sigma}_h \in \boldsymbol{\Sigma}_h$ such that

$$\boldsymbol{\Lambda}(\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) = \gamma\,(\boldsymbol{f}, \mathbf{div}\,\boldsymbol{\tau}_h) \quad \text{for all } \boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h. \tag{3.1}$$

Let $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma}_h$ be solutions to problem (2.4) and problem (3.1), respectively. The following error bound is established in [13] for the RT element with $\gamma = O(h^{-1})$:

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,\Omega} + \|\mathbf{div}\,(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_{0,\Omega} \lesssim h\,(\|\boldsymbol{\sigma}\|_{1,\Omega} + \|\boldsymbol{f}\|_{1,\Omega}).$$

Here and thereafter, we use $\lesssim$ to denote "less than or equal to" with a factor c independent of the mesh size h or other parameters appearing in the inequality. The above error bound for the BDM element can be obtained in a similar fashion.

Next we give the discretization of problem (2.4) and its error estimate, which will play an important role in later analysis. Define the space $\mathbf{V}_h \subset \mathbf{L}^2(\Omega, \mathbb{R}^2)$ as follows:

$$\mathbf{V}_h = \{(v_1, v_2)^t \,|\, v_i|_T \in P_0(T) \text{ for } i = 1, 2\}.$$

Since both the RT element and the BDM element satisfy the discrete inf-sup condition [10], we have $\mathbf{div}\boldsymbol{\Sigma}_h = \mathbf{V}_h$. Consider the discrete problem for system (2.4): given $\boldsymbol{f} \in \mathbf{L}^2(\Omega, \mathbb{R}^2)$, find $\boldsymbol{\sigma}_h \in \boldsymbol{\Sigma}_h$ and $\boldsymbol{u}_h \in \mathbf{V}_h$ such that

$$\begin{cases} (\mathcal{A}\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) + (\mathbf{div}\boldsymbol{\tau}_h, \boldsymbol{u}_h) = 0 & \text{for all } \boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h, \\ (\mathbf{div}\boldsymbol{\sigma}_h, \boldsymbol{v}_h) = (\boldsymbol{f}, \boldsymbol{v}_h) & \text{for all } \boldsymbol{v}_h \in \mathbf{V}_h. \end{cases} \tag{3.2}$$

By Lemma 2.1 and the discrete inf-sup condition of the RT/BDM element space, Problem (3.2) is well-posed and admits a unique solution.

Next, we will give the mixed finite element error estimate for problem (3.2). To do this, we need to show that $\boldsymbol{\Sigma}_h$, under the constraint $\int_\Omega \text{tr}\,\boldsymbol{\sigma}\,d\boldsymbol{x} = 0$, is a good approximation for $\boldsymbol{\Sigma}$. Denote by $\tilde{\Pi}_h$ the natural interpolation associated with the degrees of freedom onto $\boldsymbol{\Sigma}_h + span\{\mathbf{I}\}$. Let $\mathbf{P}_{\mathbf{V}_h}$ be the $\mathbf{L}^2$ projection onto $\mathbf{V}_h$. Define $\Pi_h : \boldsymbol{\Sigma} \to \boldsymbol{\Sigma}_h$ by

$$\Pi_h\,\boldsymbol{\tau} = \tilde{\Pi}_h\,\boldsymbol{\tau} - \frac{\int_\Omega (\text{tr}\,\tilde{\Pi}_h\,\boldsymbol{\tau})\,d\boldsymbol{x}}{2\,|\Omega|}\,\mathbf{I} \quad \text{for all } \boldsymbol{\tau} \in \boldsymbol{\Sigma}, \tag{3.3}$$

where $|\Omega| = \int_\Omega d\boldsymbol{x}$. Clearly, Π_h is a linear operator and $\Pi_h\,\boldsymbol{\tau} \in \boldsymbol{\Sigma}_h$. By the properties of the natural interpolation $\tilde{\Pi}_h$ [10] and the definition of Π_h, it is easy to prove [13] the following lemma:

LEMMA 3.1. *The commutative property* $\mathbf{div}\Pi_h = \mathbf{P}_{\mathbf{V}_h}\mathbf{div}$ *holds. Furthermore, for* $\boldsymbol{\tau} \in \boldsymbol{\Sigma} \cap \boldsymbol{H}^1(\Omega, \mathbb{M}_2)$, *we have*

$$\|\boldsymbol{\tau} - \Pi_h\boldsymbol{\tau}\|_{0,\Omega} \lesssim h|\boldsymbol{\tau}|_{1,\Omega}, \tag{3.4}$$

$$\|\mathbf{div}\boldsymbol{\tau} - \mathbf{div}(\Pi_h\boldsymbol{\tau})\|_{0,\Omega} \lesssim \|\mathbf{div}\boldsymbol{\tau}\|_{0,\Omega}. \tag{3.5}$$

The following lemma gives error estimates for the mixed finite element approximation (3.2).

LEMMA 3.2. *Let* $\boldsymbol{f} \in \mathbf{L}^2(\Omega, \mathbb{R}^2)$. *Define* $(\boldsymbol{\sigma}, \boldsymbol{u})$ *to be the solution of problem (2.4) and* $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h)$ *to be the solution of problem (3.2). We have*

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,\Omega} \lesssim h|\boldsymbol{\sigma}|_{1,\Omega},$$

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,\Omega} \lesssim h|\boldsymbol{u}|_{1,\Omega} + h|\boldsymbol{\sigma}|_{1,\Omega}.$$

Proof. By using the mixed finite element theory and the properties of operator Π_h, it is standard to prove that [10, 5]

$$\|\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_{0,\Omega} \lesssim h|\boldsymbol{\sigma}|_{1,\Omega},$$

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,\Omega} \lesssim h|\boldsymbol{u}|_{1,\Omega} + h|\boldsymbol{\sigma}|_{1,\Omega}.$$

Next, notice that $\mathbf{div}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) = \boldsymbol{f} - \mathbf{P}_{\mathbf{V}_h}\boldsymbol{f}$. By Lemma 2.1 and the approximation property of the $\mathbf{L}^2$ projection, we have

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,\Omega} \lesssim \|\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_{0,\Omega} + \|\mathbf{div}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_{-1,\Omega}$$

$$\lesssim h|\boldsymbol{\sigma}|_{1,\Omega} + \sup_{\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega, \mathbb{R}^2)} \frac{(\boldsymbol{f} - \mathbf{P}_{\mathbf{V}_h}\boldsymbol{f}, \boldsymbol{v})}{\|\boldsymbol{v}\|_{1,\Omega}}$$

$$= h|\boldsymbol{\sigma}|_{1,\Omega} + \sup_{\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega, \mathbb{R}^2)} \frac{(\boldsymbol{f}, \boldsymbol{v} - \mathbf{P}_{\mathbf{V}_h}\boldsymbol{v})}{\|\boldsymbol{v}\|_{1,\Omega}}$$

$$\lesssim h|\boldsymbol{\sigma}|_{1,\Omega} + h\|\boldsymbol{f}\|_{0,\Omega} = h|\boldsymbol{\sigma}|_{1,\Omega} + h\|\mathbf{div}\boldsymbol{\sigma}\|_{0,\Omega}$$

$$\lesssim h|\boldsymbol{\sigma}|_{1,\Omega}.$$

This completes the proof of the lemma. $\square$

4. A Decomposition in $\mathbf{\Sigma}_h$. In this section we introduce a Helmholtz-type decomposition of the discrete space $\mathbf{\Sigma}_h$. This decomposition and its properties borrowed many ideas from the decomposition presented in the work of Arnold, Falk and Winther on $H(\text{div})$ problems [1, 2]. The difference is that our decomposition need to use the projection operator $\mathcal{A}$, which makes the proof more complicated. For the convenience of readers, we give all details of the proof even if some parts are similar to the proof given in [1, 2].

Consider the mixed finite element problem (3.2). Let $\boldsymbol{\sigma}_h$ be the solution to (3.2) with the right hand side function $\boldsymbol{f} \in \mathbf{L}^2(\Omega, \mathbb{R}^2)$. Denote $\mathbf{div}_h^{-1}\boldsymbol{f} = \boldsymbol{\sigma}_h$. The operator $\mathbf{div}_h^{-1}$ maps $\mathbf{L}^2(\Omega, \mathbb{R}^2)$ to $\mathbf{\Sigma}_h$ and can be considered as a pseudo-inverse of the $\mathbf{div}$ operator. It is not hard to see that for all $\boldsymbol{\tau}_h \in \mathbf{\Sigma}_h$ such that $\mathbf{div}\boldsymbol{\tau}_h = \mathbf{0}$,

$$(\mathcal{A}\mathbf{div}_h^{-1}\boldsymbol{f}, \boldsymbol{\tau}_h) = -(\mathbf{div}\boldsymbol{\tau}_h, \boldsymbol{u}_h) = 0.$$

Define

$$\mathbf{\Sigma}_h^0 = \{\boldsymbol{\tau}_h \in \mathbf{\Sigma}_h \text{ such that } \mathbf{div}\boldsymbol{\tau}_h = \mathbf{0}\}, \tag{4.1}$$

$$\mathbf{\Sigma}_h^1 = \{\boldsymbol{\tau}_h \in \mathbf{div}_h^{-1}\mathbf{V}_h\}. \tag{4.2}$$

Then $\mathbf{\Sigma}_h^0$ and $\mathbf{\Sigma}_h^1$ are orthogonal under $\mathbf{\Lambda}(\cdot, \cdot)$. It is easy to see that $\mathbf{\Sigma}_h = \mathbf{\Sigma}_h^0 \oplus \mathbf{\Sigma}_h^1$.

Define the space $\tilde{\mathbf{U}}_h \subset \boldsymbol{H}^1(\Omega, \mathbb{R}^2)$ as follows: for RT element, let

$$\tilde{\mathbf{U}}_h = \{(q_1, q_2)^t \in \boldsymbol{H}^1(\Omega, \mathbb{R}^2) \,|\, q_1, q_2 \text{ are } P_1 \text{ polynomials on each } T \in \mathcal{T}\}.$$

and for BDM element, let

$$\tilde{\mathbf{U}}_h = \{(q_1, q_2)^t \in \boldsymbol{H}^1(\Omega, \mathbb{R}^2) \,|\, q_1, q_2 \text{ are } P_2 \text{ polynomials on each } T \in \mathcal{T}\}.$$

Define

$$\mathbf{U}_h = \tilde{\mathbf{U}}_h / span\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -y \\ x \end{pmatrix}\}.$$

It is well known that every $\boldsymbol{\tau}_h \in \mathbf{\Sigma}_h^0$ has a potential $\boldsymbol{q}_h \in \tilde{\mathbf{U}}_h$ such that $\boldsymbol{\tau}_h = \mathbf{curl}\,\boldsymbol{q}_h$ [10]. The constraint $\int_\Omega \operatorname{tr}\boldsymbol{\tau}_h \, d\boldsymbol{x} = 0$ further implies that $\boldsymbol{q}_h$ can be chosen from $\mathbf{U}_h$. In other words, $\mathbf{\Sigma}_h^0 = \mathbf{curl}\,(\mathbf{U}_h)$. Hence we have the following exact sequence:

$$0 \longrightarrow \mathbf{U}_h \xrightarrow{\mathbf{curl}} \mathbf{\Sigma}_h \xrightarrow{\mathbf{div}} \mathbf{V}_h \longrightarrow 0$$

For each $\boldsymbol{\tau}_h \in \mathbf{\Sigma}_h$, we have the decomposition $\boldsymbol{\sigma}_h = \mathbf{curl}\,\boldsymbol{q}_h + \mathbf{div}_h^{-1}\boldsymbol{v}_h$ where $\boldsymbol{q}_h \in \mathbf{U}_h$, $\boldsymbol{v}_h = \mathbf{div}\boldsymbol{\sigma}_h \in \mathbf{V}_h$ and the decomposition is orthogonal under both $(\mathcal{A}\cdot, \cdot)$ and $(\mathbf{div}\cdot, \mathbf{div}\cdot)$. This decomposition is usually referred to as the discrete Helmholtz decomposition.

Next, we study some properties related to the discrete Helmholtz decomposition. Define an operator $\mathbf{grad}_h : \mathbf{V}_h \to \mathbf{\Sigma}_h$ as follows:

$$(\mathbf{grad}_h\boldsymbol{v}_h, \boldsymbol{\tau}_h) = -(\mathbf{div}\boldsymbol{\tau}_h, \boldsymbol{v}_h), \quad \text{for all } \boldsymbol{\tau}_h \in \mathbf{\Sigma}_h.$$

The operator $\mathbf{grad}_h$ is actually the discrete dual operator of $\mathbf{div}$. Notice that for an arbitrary $\boldsymbol{\sigma}_h \in \mathbf{\Sigma}_h$, $\mathcal{A}\boldsymbol{\sigma}_h$ is not necessarily in $\mathbf{\Sigma}_h$ since it may not satisfy the

$\boldsymbol{H}(\mathbf{div}, \Omega, \mathbb{M}_2)$ continuity requirement on normal components. Therefore we need to define discrete operators $A_h : \boldsymbol{\Sigma}_h \to \boldsymbol{\Sigma}_h$ and $\boldsymbol{\Lambda}_h : \boldsymbol{\Sigma}_h \to \boldsymbol{\Sigma}_h$ by

$$(A_h \boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) = (\mathcal{A} \boldsymbol{\sigma}_h, \boldsymbol{\tau}_h), \quad \text{for all } \boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h,$$

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

It is equivalent to say that $A_h = \mathbf{P}_{\boldsymbol{\Sigma}_h} \mathcal{A}$ where $\mathbf{P}_{\boldsymbol{\Sigma}_h}$ is the $\mathbf{L}^2$ projection onto the finite dimensional space $\boldsymbol{\Sigma}_h$ and $\boldsymbol{\Lambda}_h \boldsymbol{\sigma}_h = A_h \boldsymbol{\sigma}_h - \gamma \mathbf{grad}_h \mathbf{div} \boldsymbol{\sigma}_h$. By the property of the projection operator $\mathbf{P}_{\boldsymbol{\Sigma}_h}$, we have

$$\|A_h \boldsymbol{\tau}_h\|_{0,\Omega} \lesssim \|\mathcal{A} \boldsymbol{\tau}_h\|_{0,\Omega} \qquad \text{for all } \boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h. \tag{4.3}$$

Let $\mathcal{T}_H$ be a quasi-uniform triangulation of Ω with characteristic mesh size H and let $\mathcal{T}_h$ be a refinement of $\mathcal{T}_H$ with characteristic mesh size h. We follow the notations previously introduced with the understanding that spaces and operators with subscript H are defined over the mesh $\mathcal{T}_H$. For example, $\boldsymbol{\Sigma}_H$ is defined on $\mathcal{T}_H$ in the same way that $\boldsymbol{\Sigma}_h$ is defined on $\mathcal{T}_h$. Besides, the RT/BDM finite element spaces are known to be nested, that is, $\boldsymbol{\Sigma}_H \subset \boldsymbol{\Sigma}_h$.

LEMMA 4.1. *Let $\boldsymbol{v}_h \in \mathbf{V}_h$. Define $\boldsymbol{\sigma}_h = \mathbf{div}_h^{-1} \boldsymbol{v}_h$ and $\boldsymbol{\sigma}_H = \mathbf{div}_H^{-1} \boldsymbol{v}_h$. Then*

$$\|\boldsymbol{\sigma}_h - \boldsymbol{\sigma}_H\|_{0,\Omega} \lesssim H \|\boldsymbol{v}_h\|_{0,\Omega},$$

$$\|\mathbf{div}(\boldsymbol{\sigma}_h - \boldsymbol{\sigma}_H)\|_{0,\Omega} \lesssim H \|\mathbf{grad}_h \boldsymbol{v}_h\|_{0,\Omega}.$$

Proof.   Let $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}$ satisfies

$$\begin{cases} (\mathcal{A}\boldsymbol{\sigma}, \boldsymbol{\tau}) + (\mathbf{div}\boldsymbol{\tau}, \boldsymbol{u}) = 0 & \text{for all } \boldsymbol{\tau} \in \boldsymbol{\Sigma}, \\ (\mathbf{div}\boldsymbol{\sigma}, \boldsymbol{w}) = (\boldsymbol{v}_h, \boldsymbol{w}) & \text{for all } \boldsymbol{w} \in \mathbf{L}^2(\Omega, \mathbb{R}^2). \end{cases}$$

By the regularity property (2.3), Lemma 3.2 and the fact that $h < H$, we have

$$\|\boldsymbol{\sigma}_h - \boldsymbol{\sigma}_H\|_{0,\Omega} \leq \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,\Omega} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_H\|_{0,\Omega} \lesssim H|\boldsymbol{\sigma}|_{1,\Omega} \lesssim H\|\boldsymbol{v}_h\|_{0,\Omega}.$$

Notice that $\mathbf{div}(\boldsymbol{\sigma}_h - \boldsymbol{\sigma}_H) = \boldsymbol{v}_h - \mathbf{P}_{\mathbf{V}_H}\boldsymbol{v}_h$, where $\mathbf{P}_{\mathbf{V}_H}$ is the $\mathbf{L}^2$ projection onto $\mathbf{V}_H$. Define $\tilde{\boldsymbol{\sigma}} \in \boldsymbol{\Sigma}$ by

$$\begin{cases} (\mathcal{A}\tilde{\boldsymbol{\sigma}}, \boldsymbol{\tau}) + (\mathbf{div}\boldsymbol{\tau}, \boldsymbol{u}) = 0 & \text{for all } \boldsymbol{\tau} \in \boldsymbol{\Sigma}, \\ (\mathbf{div}\tilde{\boldsymbol{\sigma}}, \boldsymbol{w}) = (\boldsymbol{v}_h - \mathbf{P}_{\mathbf{V}_H}\boldsymbol{v}_h, \boldsymbol{w}) & \text{for all } \boldsymbol{w} \in \mathbf{L}^2(\Omega, \mathbb{R}^2). \end{cases}$$

Recall that Π_h and Π_H satisfy $\mathbf{div}\Pi_h = \mathbf{P}_{\mathbf{V}_h}\mathbf{div}$ and $\mathbf{div}\Pi_H = \mathbf{P}_{\mathbf{V}_H}\mathbf{div}$. By using Lemma 3.1, the Schwarz inequality and the regularity result (2.3), we have

$$\|\boldsymbol{v}_h - \mathbf{P}_{\mathbf{V}_H}\boldsymbol{v}_h\|_{0,\Omega}^2 = (\mathbf{div}\tilde{\boldsymbol{\sigma}}, \boldsymbol{v}_h - \mathbf{P}_{\mathbf{V}_H}\boldsymbol{v}_h)$$

$$= (\mathbf{div}(\Pi_h\tilde{\boldsymbol{\sigma}} - \Pi_H\tilde{\boldsymbol{\sigma}}), \boldsymbol{v}_h) = -((\Pi_h\tilde{\boldsymbol{\sigma}} - \Pi_H\tilde{\boldsymbol{\sigma}}), \mathbf{grad}_h\boldsymbol{v}_h)$$

$$\leq \|\Pi_h\tilde{\boldsymbol{\sigma}} - \Pi_H\tilde{\boldsymbol{\sigma}}\|_{0,\Omega}\|\mathbf{grad}_h\boldsymbol{v}_h\|_{0,\Omega} \lesssim H|\tilde{\boldsymbol{\sigma}}|_{1,\Omega}\|\mathbf{grad}_h\boldsymbol{v}_h\|_{0,\Omega}$$

$$\lesssim H\|\boldsymbol{v}_h - \mathbf{P}_{\mathbf{V}_H}\boldsymbol{v}_h\|_{0,\Omega}\|\mathbf{grad}_h\boldsymbol{v}_h\|_{0,\Omega}.$$

This completes the proof of the lemma. $\square$

LEMMA 4.2. *Given $\boldsymbol{\sigma}_h \in \boldsymbol{\Sigma}_h$ satisfying $\boldsymbol{\Lambda}(\boldsymbol{\sigma}_h, \boldsymbol{\tau}_H) = 0$ for all $\boldsymbol{\tau}_H \in \boldsymbol{\Sigma}_H$. Let $\boldsymbol{\sigma}_h = \mathbf{curl}\, q_h + \mathbf{div}_h^{-1} v_h$, where $q_h \in \mathbf{U}_h$ and $v_h = \mathbf{div}\,\boldsymbol{\sigma}_h \in \mathbf{V}_h$, be the $\boldsymbol{\Lambda}$-orthogonal decomposition of $\boldsymbol{\sigma}_h$. Denote $\boldsymbol{\sigma}_h^1 = \mathbf{div}_h^{-1} v_h$. Then*

$$(\mathcal{A}\boldsymbol{\sigma}_h^1, \boldsymbol{\sigma}_h^1) \lesssim \gamma^{-1} H^2 \boldsymbol{\Lambda}(\boldsymbol{\sigma}_h, \boldsymbol{\sigma}_h), \tag{4.4}$$

$$(1 + \gamma H^{-2}) \|\boldsymbol{\sigma}_h^1\|_{0,\Omega}^2 \lesssim (1 + \gamma^{-2}) \boldsymbol{\Lambda}(\boldsymbol{\sigma}_h, \boldsymbol{\sigma}_h), \tag{4.5}$$

$$\|q_h\|_{0,\Omega}^2 \lesssim H^2 (\mathcal{A}\mathbf{curl}\, q_h, \mathbf{curl}\, q_h). \tag{4.6}$$

Proof. The proof of inequality (4.6) is a little long and hence will be put in the Appendix. Here we give the proof of inequalities (4.4) and (4.5). Define $\tilde{\boldsymbol{\sigma}}_h \in \boldsymbol{\Sigma}_h$ by

$$\boldsymbol{\Lambda}(\tilde{\boldsymbol{\sigma}}_h, \boldsymbol{\tau}_h) = (\mathcal{A}\boldsymbol{\sigma}_h^1, \boldsymbol{\tau}_h), \quad \text{for all } \boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h. \tag{4.7}$$

It is clear that

$$\boldsymbol{\Lambda}(\tilde{\boldsymbol{\sigma}}_h, \mathbf{curl}\, p_h) = (\mathcal{A}\boldsymbol{\sigma}_h^1, \mathbf{curl}\, p_h) = 0, \quad \text{for all } p_h \in \mathbf{U}_h.$$

This implies $\tilde{\boldsymbol{\sigma}}_h \in \boldsymbol{\Sigma}_h^1$. In other words, if we denote $\mathbf{div}\,\tilde{\boldsymbol{\sigma}}_h = \tilde{v}_h$, then $\tilde{\boldsymbol{\sigma}}_h = \mathbf{div}_h^{-1} \tilde{v}_h$. Define $\tilde{\boldsymbol{\sigma}}_H = \mathbf{div}_H^{-1} \tilde{v}_h$. By the Schwarz inequality,

$$\begin{aligned}
(\mathcal{A}\boldsymbol{\sigma}_h^1, \boldsymbol{\sigma}_h^1) &= \boldsymbol{\Lambda}(\tilde{\boldsymbol{\sigma}}_h, \boldsymbol{\sigma}_h^1) = \boldsymbol{\Lambda}(\tilde{\boldsymbol{\sigma}}_h, \boldsymbol{\sigma}_h) \\
&= \boldsymbol{\Lambda}(\tilde{\boldsymbol{\sigma}}_h - \tilde{\boldsymbol{\sigma}}_H, \boldsymbol{\sigma}_h) \leq \boldsymbol{\Lambda}(\tilde{\boldsymbol{\sigma}}_h - \tilde{\boldsymbol{\sigma}}_H, \tilde{\boldsymbol{\sigma}}_h - \tilde{\boldsymbol{\sigma}}_H)^{1/2} \boldsymbol{\Lambda}(\boldsymbol{\sigma}_h, \boldsymbol{\sigma}_h)^{1/2}.
\end{aligned} \tag{4.8}$$

We also notice that Equation (4.7) implies that $\boldsymbol{\Lambda}_h \tilde{\boldsymbol{\sigma}}_h = A_h \boldsymbol{\sigma}_h^1$. Furthermore, by setting $\boldsymbol{\tau}_h = \tilde{\boldsymbol{\sigma}}_h$ in (4.7) and using the Schwarz inequality, we can easily see that $\boldsymbol{\Lambda}(\tilde{\boldsymbol{\sigma}}_h, \tilde{\boldsymbol{\sigma}}_h) \leq (\mathcal{A}\boldsymbol{\sigma}_h^1, \boldsymbol{\sigma}_h^1)$. Hence by the fact that $\mathcal{A}$ is a projection, lemma 4.1, the triangular inequality and inequality (4.3), we have

$$\begin{aligned}
\boldsymbol{\Lambda}(\tilde{\boldsymbol{\sigma}}_h - \tilde{\boldsymbol{\sigma}}_H, \tilde{\boldsymbol{\sigma}}_h - \tilde{\boldsymbol{\sigma}}_H) &\leq \|\tilde{\boldsymbol{\sigma}}_h - \tilde{\boldsymbol{\sigma}}_H\|_{0,\Omega}^2 + \gamma \|\mathbf{div}(\tilde{\boldsymbol{\sigma}}_h - \tilde{\boldsymbol{\sigma}}_H)\|_{0,\Omega}^2 \\
&\lesssim H^2 \gamma^{-1} (\gamma \|\mathbf{div}\,\tilde{\boldsymbol{\sigma}}_h\|_{0,\Omega}^2 + \gamma^2 \|\mathbf{grad}_h \mathbf{div}\,\tilde{\boldsymbol{\sigma}}_h\|_{0,\Omega}^2) \\
&\lesssim H^2 \gamma^{-1} (\gamma \|\mathbf{div}\,\tilde{\boldsymbol{\sigma}}_h\|_{0,\Omega}^2 + \|A_h \tilde{\boldsymbol{\sigma}}_h\|_{0,\Omega}^2 + \|(A_h - \gamma \mathbf{grad}_h \mathbf{div})\tilde{\boldsymbol{\sigma}}_h\|_{0,\Omega}^2) \\
&\lesssim H^2 \gamma^{-1} (\boldsymbol{\Lambda}(\tilde{\boldsymbol{\sigma}}_h, \tilde{\boldsymbol{\sigma}}_h) + \|\boldsymbol{\Lambda}_h \tilde{\boldsymbol{\sigma}}_h\|_{0,\Omega}^2) \\
&\lesssim H^2 \gamma^{-1} ((\mathcal{A}\boldsymbol{\sigma}_h^1, \boldsymbol{\sigma}_h^1) + \|A_h \boldsymbol{\sigma}_h^1\|_{0,\Omega}^2) \\
&\lesssim H^2 \gamma^{-1} (\mathcal{A}\boldsymbol{\sigma}_h^1, \boldsymbol{\sigma}_h^1).
\end{aligned} \tag{4.9}$$

Combining (4.8) and (4.9) gives $(\mathcal{A}\boldsymbol{\sigma}_h^1, \boldsymbol{\sigma}_h^1) \lesssim H\gamma^{-1/2}(\mathcal{A}\boldsymbol{\sigma}_h^1, \boldsymbol{\sigma}_h^1)^{1/2} \boldsymbol{\Lambda}(\boldsymbol{\sigma}_h, \boldsymbol{\sigma}_h)^{1/2}$. So (4.4) holds.

Next, we prove (4.5). Notice that $\boldsymbol{\Lambda}(\boldsymbol{\sigma}_h, \boldsymbol{\tau}_H) = 0$ for all $\boldsymbol{\tau}_H \in \boldsymbol{\Sigma}_H$ implies

$$(\mathbf{div}\,\boldsymbol{\sigma}_h^1, \mathbf{div}\,\boldsymbol{\tau}_H) = -\gamma^{-1}(\mathcal{A}\boldsymbol{\sigma}_h^1, \boldsymbol{\tau}_H) - \gamma^{-1}(\mathcal{A}\mathbf{curl}\, q_h, \boldsymbol{\tau}_H). \tag{4.10}$$

By definition,

$$\|\mathbf{div}\,\boldsymbol{\sigma}_h^1\|_{-1,\Omega} = \sup_{v \in \boldsymbol{H}_0^1(\Omega, \mathbb{R}^2)} \frac{(\mathbf{div}\,\boldsymbol{\sigma}_h^1, v)}{\|v\|_{1,\Omega}}.$$

For $v \in \boldsymbol{H}_0^1(\Omega, \mathbb{R}^2)$, let v_H be the $\mathbf{L}^2$-projection of v onto $\mathbf{V}_H$. Then the following estimates are well-known:

$$\|v - v_H\|_{0,\Omega} \lesssim H\|v\|_{1,\Omega},$$

$$\|v_H\|_{0,\Omega} \leq \|v\|_{0,\Omega} \leq \|v\|_{1,\Omega}.$$

Define $\boldsymbol{\tau}_H = \mathbf{div}_H^{-1}\boldsymbol{v}_H$ and $\boldsymbol{\tau}_h = \mathbf{div}_h^{-1}\boldsymbol{v}_H$. We have $(\mathcal{A}\mathbf{curl}\,\boldsymbol{q}_h, \boldsymbol{\tau}_h) = 0$. Besides, by Lemma 3.2 and the regularity result (2.3), it is not hard to see that $\|\boldsymbol{\tau}_H\|_{0,\Omega} \lesssim \|\boldsymbol{v}_H\|_{0,\Omega}$. Combining the above and using Lemma 4.1 and Equation (4.10), we have

$$\sup_{\boldsymbol{v}\in\boldsymbol{H}_0^1(\Omega,\mathbb{R}^2)} \frac{(\mathbf{div}\boldsymbol{\sigma}_h^1, \boldsymbol{v})}{\|\boldsymbol{v}\|_{1,\Omega}} \leq \sup_{\boldsymbol{v}\in\boldsymbol{H}_0^1(\Omega,\mathbb{R}^2)} \frac{(\mathbf{div}\boldsymbol{\sigma}_h^1, \boldsymbol{v} - \boldsymbol{v}_H)}{\|\boldsymbol{v}\|_{1,\Omega}} + \sup_{\boldsymbol{v}\in\boldsymbol{H}_0^1(\Omega,\mathbb{R}^2)} \frac{(\mathbf{div}\boldsymbol{\sigma}_h^1, \boldsymbol{v}_H)}{\|\boldsymbol{v}\|_{1,\Omega}}$$

$$\lesssim H\|\mathbf{div}\boldsymbol{\sigma}_h^1\|_{0,\Omega} + \sup_{\boldsymbol{v}\in\boldsymbol{H}_0^1(\Omega,\mathbb{R}^2)} \frac{(\mathbf{div}\boldsymbol{\sigma}_h^1, \mathbf{div}\boldsymbol{\tau}_H)}{\|\boldsymbol{v}\|_{1,\Omega}}$$

$$= H\|\mathbf{div}\boldsymbol{\sigma}_h^1\|_{0,\Omega} + \sup_{\boldsymbol{v}\in\boldsymbol{H}_0^1(\Omega,\mathbb{R}^2)} \frac{-\gamma^{-1}(\mathcal{A}\boldsymbol{\sigma}_h^1, \boldsymbol{\tau}_H) - \gamma^{-1}(\mathcal{A}\mathbf{curl}\,\boldsymbol{q}_h, \boldsymbol{\tau}_H - \boldsymbol{\tau}_h)}{\|\boldsymbol{v}\|_{1,\Omega}}$$

$$\lesssim H\|\mathbf{div}\boldsymbol{\sigma}_h^1\|_{0,\Omega} + \gamma^{-1}(\mathcal{A}\boldsymbol{\sigma}_h^1, \boldsymbol{\sigma}_h^1)^{1/2} + \gamma^{-1}H(\mathcal{A}\mathbf{curl}\,\boldsymbol{q}_h, \mathbf{curl}\,\boldsymbol{q}_h)^{1/2}.$$

Using the above result, Lemma 2.1, inequality (4.4), the inequality $2\gamma^{-1} \leq 1+\gamma^{-2}$ and the fact that $H < O(1)$, we have

$$\|\boldsymbol{\sigma}_h^1\|_{0,\Omega}^2 \lesssim (\mathcal{A}\boldsymbol{\sigma}_h^1, \boldsymbol{\sigma}_h^1) + \|\mathbf{div}\boldsymbol{\sigma}_h^1\|_{-1,\Omega}^2$$
$$\lesssim (1+\gamma^{-2})\boldsymbol{\Lambda}(\boldsymbol{\sigma}_h^1, \boldsymbol{\sigma}_h^1) + \|\mathbf{div}\boldsymbol{\sigma}_h^1\|_{0,\Omega}^2 + \gamma^{-2}(\mathcal{A}\mathbf{curl}\,\boldsymbol{q}_h, \mathbf{curl}\,\boldsymbol{q}_h)$$
$$\lesssim (1+\gamma^{-2})\boldsymbol{\Lambda}(\boldsymbol{\sigma}_h, \boldsymbol{\sigma}_h),$$

$$\gamma H^{-2}\|\boldsymbol{\sigma}_h^1\|_{0,\Omega}^2 \lesssim (1+\gamma^{-2})\gamma H^{-2}(\mathcal{A}\boldsymbol{\sigma}_h^1, \boldsymbol{\sigma}_h^1) + \gamma\|\mathbf{div}\boldsymbol{\sigma}_h^1\|_{0,\Omega}^2 + \gamma^{-1}(\mathcal{A}\mathbf{curl}\,\boldsymbol{q}_h, \mathbf{curl}\,\boldsymbol{q}_h)$$
$$\lesssim (1+\gamma^{-2})\boldsymbol{\Lambda}(\boldsymbol{\sigma}_h, \boldsymbol{\sigma}_h).$$

Adding up the above inequalities gives the estimate (4.5). This completes the proof of the lemma. $\square$

5. A Multigrid Solver. For simplicity, we rewrite problem (3.1) in the following form: find $\boldsymbol{\sigma}_h \in \boldsymbol{\Sigma}_h$ such that

$$\boldsymbol{\Lambda}(\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) = (\boldsymbol{F}, \boldsymbol{\tau}_h), \quad \text{for all } \boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h. \tag{5.1}$$

where $\boldsymbol{F} \in \boldsymbol{\Sigma}_h$ is defined by $(\boldsymbol{F}, \boldsymbol{\tau}_h) = \gamma(\boldsymbol{f}, \mathbf{div}\boldsymbol{\tau}_h)$.

In this section, we construct and analyze a multigrid solver for the discrete problem (5.1). Let $\mathcal{T}_k$, $k = 1, \ldots, K$ be a series of nested meshes under uniform refinements and assume that $\mathcal{T}_1$ has characteristic mesh size $O(1)$. Denote the characteristic mesh size of $\mathcal{T}_k$ to be h_k. Define $\boldsymbol{\Sigma}_k = \boldsymbol{\Sigma}_{h_k}$, the RT/BDM finite element space based on $\mathcal{T}_k$. Notice that $\boldsymbol{\Sigma}_{k-1} \subset \boldsymbol{\Sigma}_k$. Similarly we define spaces $\mathbf{U}_k$ and $\mathbf{V}_k$. Denote $\mathbf{div}_k^{-1} = \mathbf{div}_{h_k}^{-1}$. Denote $\mathbf{Q}_{k-1} : \boldsymbol{\Sigma}_k \to \boldsymbol{\Sigma}_{k-1}$ and $\mathbf{P}_{k-1} : \boldsymbol{\Sigma}_k \to \boldsymbol{\Sigma}_{k-1}$ to be the $\mathbf{L}^2$-projection and the $\boldsymbol{\Lambda}$-projection respectively. Let $\mathbf{R}_k : \boldsymbol{\Sigma}_k \to \boldsymbol{\Sigma}_k$ be a symmetric positive definite operator, which will be defined later. Let m be the smoothing times on each level. Finally, denote A_k to be A_{h_k}.

We define a V-cycle multigrid iteration operator $Mg_k(\cdot, \cdot) : \boldsymbol{\Sigma}_k \times \boldsymbol{\Sigma}_k \to \boldsymbol{\Sigma}_k$ inductively as follows:

ALGORITHM 1. *Set $Mg_1(\boldsymbol{\tau}, \boldsymbol{F}) = \boldsymbol{\Lambda}_1^{-1}\boldsymbol{F}$. Assuming that $Mg_{k-1}(\cdot, \cdot)$ has been defined, define $Mg_k(\boldsymbol{\tau}, \boldsymbol{F})$ as follows: set $\boldsymbol{\tau}^0 = \boldsymbol{\tau}$ and define*
(1) $\boldsymbol{\tau}^l = \boldsymbol{\tau}^{l-1} + \mathbf{R}_k(\boldsymbol{F} - \boldsymbol{\Lambda}_k\boldsymbol{\tau}^{k-1})$ for $l = 1, \cdots, m$;
(2) $\boldsymbol{\sigma}^m = \boldsymbol{\tau}^m + \mathbf{B}_{k-1}\mathbf{Q}_{k-1}(\boldsymbol{F} - \boldsymbol{\Lambda}_k\boldsymbol{\tau}^m)$;

(3) $\boldsymbol{\sigma}^l = \boldsymbol{\sigma}^{l-1} + \mathbf{R}_k(\boldsymbol{F} - \boldsymbol{\Lambda}_k\boldsymbol{\sigma}^{l-1})$ *for $l = m+1, \cdots, 2m$;*
Set $Mg_k(\boldsymbol{\tau}, \boldsymbol{F}) = \boldsymbol{\sigma}^{2m}$.

The V-cycle multigrid iterations for solving Problem (5.1) over $\boldsymbol{\Sigma}_k$ is defined below. Given an initial guess $\boldsymbol{\sigma}^0 \in \boldsymbol{\Sigma}_k$, we defined a sequence approximating the solution $\boldsymbol{\sigma}_h$ by

$$\boldsymbol{\sigma}^{i+1} = Mg_k(\boldsymbol{\sigma}^i, \boldsymbol{F}), \qquad i = 0, 1, \dots.$$

The error propagation operator $\boldsymbol{\varepsilon} : \boldsymbol{\Sigma}_k \to \boldsymbol{\Sigma}_k$ for the above iteration is [8]

$$\boldsymbol{\sigma}_h - \boldsymbol{\sigma}^{i+1} = \boldsymbol{\varepsilon}(\boldsymbol{\sigma}_h - \boldsymbol{\sigma}^i) = Mg_k(\boldsymbol{\sigma}_h - \boldsymbol{\sigma}^i, 0).$$

The abstract multigrid theorem [8, 1] states that under the following two assumptions:

(M.1) The spectrum of $I - \mathbf{R}_k\boldsymbol{\Lambda}_k$ is in $[0, 1)$. That is

$$0 \le \boldsymbol{\Lambda}((I - \mathbf{R}_k\boldsymbol{\Lambda}_k)\boldsymbol{\tau}, \boldsymbol{\tau}) < \boldsymbol{\Lambda}(\boldsymbol{\tau}, \boldsymbol{\tau}) \qquad \text{for all } \boldsymbol{\tau} \in \boldsymbol{\Sigma}_k;$$

(M.2) There exists a positive constant C_P such that

$$\boldsymbol{\Lambda}((I - \mathbf{P}_{k-1})\boldsymbol{\tau}, \boldsymbol{\tau}) \le C_p\mathbf{R}_k(\boldsymbol{\Lambda}_k\boldsymbol{\tau}, \boldsymbol{\Lambda}_k\boldsymbol{\tau}) \quad \text{for all } \boldsymbol{\tau} \in \boldsymbol{\Sigma}_k;$$

the multigrid solver with $\mathbf{R}_k$ as smoother and m smoothing times on each level will have convergence rate

$$\boldsymbol{\Lambda}(\boldsymbol{\varepsilon}\boldsymbol{\tau}, \boldsymbol{\tau}) \le \frac{C_p}{C_p + 2m}\boldsymbol{\Lambda}(\boldsymbol{\tau}, \boldsymbol{\tau}), \quad \text{for all } \boldsymbol{\tau} \in \boldsymbol{\Sigma}_k. \tag{5.2}$$

In the remainder of this section, we will construct a smoother $\mathbf{R}_k$ which satisfies assumptions (M.1) and (M.2).

Let $\mathcal{V}_k$ be the set of all vertices in $\mathcal{T}_k$, including the boundary vertices. For each $\boldsymbol{x} \in \mathcal{V}_k$, define $\Omega_{k,\boldsymbol{x}}$ to be the interior of the union of all triangles in $\mathcal{T}_k$ which have $\boldsymbol{x}$ as a vertex. Let $\boldsymbol{\Sigma}_{k,\boldsymbol{x}}$ be the RT/BDM space defined on $\mathcal{T}_k$ such that each function in $\boldsymbol{\Sigma}_{k,\boldsymbol{x}}$ vanishes outside $\Omega_{k,\boldsymbol{x}}$. Notice that $\boldsymbol{\Sigma}_{k,\boldsymbol{x}} \subset (\boldsymbol{\Sigma}_k + span\{\mathbf{I}\})$ but $\boldsymbol{\Sigma}_{k,\boldsymbol{x}} \nsubseteq \boldsymbol{\Sigma}_k$ in general, since $\int_\Omega \text{tr}\,\boldsymbol{\tau}\,d\boldsymbol{x}$ may not vanish for $\boldsymbol{\tau} \in \boldsymbol{\Sigma}_{k,\boldsymbol{x}}$.

Define operator $\boldsymbol{\Lambda}_{k,\boldsymbol{x}} : \boldsymbol{\Sigma}_{k,\boldsymbol{x}} \to \boldsymbol{\Sigma}_{k,\boldsymbol{x}}$ by

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

Notice that $\boldsymbol{\Lambda}_{k,\boldsymbol{x}}$ is symmetric and $(\boldsymbol{\Lambda}_{k,\boldsymbol{x}}\boldsymbol{\sigma}, \boldsymbol{\sigma}) = 0$ if and only if $\boldsymbol{\sigma} = c\mathbf{I}$ where c is a constant. Since $c\mathbf{I} \notin \boldsymbol{\Sigma}_{k,\boldsymbol{x}}$ and $\boldsymbol{\Sigma}_{k,\boldsymbol{x}}$ is a finite dimensional space, it is easy to see that $\boldsymbol{\Lambda}_{k,\boldsymbol{x}}$ is symmetric positive definite over $\boldsymbol{\Sigma}_{k,\boldsymbol{x}}$, although the spectrum of $\boldsymbol{\Lambda}_{k,\boldsymbol{x}}$ may not have a uniform lower bound independent of all k and $\boldsymbol{x}$.

Let $\mathbf{Q}_{k,\boldsymbol{x}} : (\boldsymbol{\Sigma}_k + span\{\mathbf{I}\}) \to \boldsymbol{\Sigma}_{k,\boldsymbol{x}}$ and $\mathbf{P}_{k,\boldsymbol{x}} : (\boldsymbol{\Sigma}_k + span\{\mathbf{I}\}) \to \boldsymbol{\Sigma}_{k,\boldsymbol{x}}$ be the $\mathbf{L}^2$ projection and the $\boldsymbol{\Lambda}$ projection respectively. Here $\mathbf{P}_{k,\boldsymbol{x}}$ is well-defined since the operator $\boldsymbol{\Lambda}_{k,\boldsymbol{x}}$ is symmetric positive definite. Define operator $\tilde{\mathbf{R}}_k : \boldsymbol{\Sigma}_k \to (\boldsymbol{\Sigma}_k + span\{\mathbf{I}\})$ using an additive Schwarz scheme:

$$\tilde{\mathbf{R}}_k = \frac{1}{3}\sum_{\boldsymbol{x} \in \mathcal{V}_k} \mathbf{P}_{k,\boldsymbol{x}}\boldsymbol{\Lambda}_k^{-1} = \frac{1}{3}\sum_{\boldsymbol{x} \in \mathcal{V}_k} \boldsymbol{\Lambda}_{k,\boldsymbol{x}}^{-1}\mathbf{Q}_{k,\boldsymbol{x}}. \tag{5.3}$$

Define $\mathbf{R}_k : \boldsymbol{\Sigma}_k \to \boldsymbol{\Sigma}_k$ by

$$\mathbf{R}_k = \mathbf{Q}_{\boldsymbol{\Sigma}_k}\tilde{\mathbf{R}}_k, \tag{5.4}$$

where $\mathbf{Q}_{\boldsymbol{\Sigma}_k} : (\boldsymbol{\Sigma}_k + span\{\mathbf{I}\}) \to \boldsymbol{\Sigma}_k$ is the $\mathbf{L}^2$ projection. It is clear that $\mathbf{R}_k$ is symmetric positive definite over $\boldsymbol{\Sigma}_k$ under the $\mathbf{L}^2$ inner-product.

REMARK 1. *For simplicity, in this paper we will only analyze the smoother $\mathbf{R}_k$ as defined in (5.4). The analysis can easily be extended to the Hiptmair smoother [19], which is usually more time-efficient in terms of a single sweep of smoothing process, and the same multigrid convergence rate estimate holds.*

Next, we show that $\mathbf{R}_k$ satisfies assumptions (M.1) and (M.2). The proof of Assumption (M.1) follows directly from the Schwartz inequality and the fact that each triangle in $\mathcal{T}_k$ is covered by exactly three sub-domains in $\{\Omega_{k,\boldsymbol{x}}\}_{\boldsymbol{x}\in\mathcal{V}_k}$. In the following, we concentrate on the proof of Assumption (M.2).

First, we state the following result:

LEMMA 5.1. *Assume the following assumption holds:*

(R.1) *For all $\boldsymbol{\tau}_k \in \boldsymbol{\Sigma}_k$, there exists a decomposition $(\mathbf{I} - \mathbf{P}_{k-1})\boldsymbol{\tau}_k = \sum_{\boldsymbol{x}\in\mathcal{V}_k} \boldsymbol{\tau}_{k,\boldsymbol{x}}$ where $\boldsymbol{\tau}_{k,\boldsymbol{x}} \in \boldsymbol{\Sigma}_{k,\boldsymbol{x}}$ and a positive constant C_0 such that*

$$\sum_{\boldsymbol{x}\in\mathcal{V}_k} \boldsymbol{\Lambda}(\boldsymbol{\tau}_{k,\boldsymbol{x}}, \boldsymbol{\tau}_{k,\boldsymbol{x}}) \leq C_0 \boldsymbol{\Lambda}((\mathbf{I} - \mathbf{P}_{k-1})\boldsymbol{\tau}_k, \boldsymbol{\tau}_k);$$

Then (M.2) holds with $C_p = 3C_0$.

Proof. Assume (R.1) holds. By using the Schwarz inequality, We have

$$\boldsymbol{\Lambda}((\mathbf{I} - \mathbf{P}_{k-1})\boldsymbol{\tau}_k, \boldsymbol{\tau}_k) = \sum_{\boldsymbol{x}\in\mathcal{V}_k} \boldsymbol{\Lambda}(\boldsymbol{\tau}_{k,\boldsymbol{x}}, \boldsymbol{\tau}_k) = \sum_{\boldsymbol{x}\in\mathcal{V}_k} \boldsymbol{\Lambda}(\boldsymbol{\tau}_{k,\boldsymbol{x}}, \mathbf{P}_{k,\boldsymbol{x}}\boldsymbol{\tau}_k)$$

$$\leq \left(\sum_{\boldsymbol{x}\in\mathcal{V}_k} \boldsymbol{\Lambda}(\boldsymbol{\tau}_{k,\boldsymbol{x}}, \boldsymbol{\tau}_{k,\boldsymbol{x}})\right)^{1/2} \boldsymbol{\Lambda}^{1/2}(\sum_{\boldsymbol{x}\in\mathcal{V}_k} \mathbf{P}_{k,\boldsymbol{x}}\boldsymbol{\tau}_k, \boldsymbol{\tau}_k)$$

$$\leq \sqrt{3} C_0^{1/2} \boldsymbol{\Lambda}^{1/2}((\mathbf{I} - \mathbf{P}_{k-1})\boldsymbol{\tau}_k, \boldsymbol{\tau}_k) \boldsymbol{\Lambda}^{1/2}(\tilde{\mathbf{R}}_k\boldsymbol{\Lambda}_k\boldsymbol{\tau}_k, \boldsymbol{\tau}_k)$$

$$= \sqrt{3} C_0^{1/2} \boldsymbol{\Lambda}^{1/2}((\mathbf{I} - \mathbf{P}_{k-1})\boldsymbol{\tau}_k, \boldsymbol{\tau}_k) \boldsymbol{\Lambda}^{1/2}(\mathbf{R}_k\boldsymbol{\Lambda}_k\boldsymbol{\tau}_k, \boldsymbol{\tau}_k).$$

The last step comes from the fact that $span\{\mathbf{I}\}$ is in the kernel of $\boldsymbol{\Lambda}_k$. This completes the proof of the lemma. $\square$

Finally, we need to prove that condition (R.1) holds.

LEMMA 5.2. *Condition (R.1) holds with $C_0 = c(1 + \gamma^{-2})$, where c is a positive constant independent of k or γ.*

Proof. Let $\boldsymbol{\sigma}_k = (\mathbf{I} - \mathbf{P}_{k-1})\boldsymbol{\tau}_k$ where $\boldsymbol{\tau}_k \in \boldsymbol{\Sigma}_k$. We decompose $\boldsymbol{\sigma}_k$ into $\boldsymbol{\sigma}_k = \mathbf{curl}\, \boldsymbol{q}_k + \mathbf{div}_k^{-1}\boldsymbol{v}_k$ where $\boldsymbol{q}_k \in \mathbf{U}_k$ and $\boldsymbol{v}_k = \mathbf{div}\boldsymbol{\sigma}_k$. Denote $\boldsymbol{\sigma}_k^1 = \mathbf{div}_k^{-1}\boldsymbol{v}_k$. Clearly

$$\boldsymbol{\Lambda}(\mathbf{curl}\, \boldsymbol{q}_k, \mathbf{curl}\, \boldsymbol{q}_k) + \boldsymbol{\Lambda}(\boldsymbol{\sigma}_k^1, \boldsymbol{\sigma}_k^1) = \boldsymbol{\Lambda}(\boldsymbol{\sigma}_k, \boldsymbol{\sigma}_k).$$

Therefore we can consider the decomposition of $\mathbf{curl}\, \boldsymbol{q}_k$ and $\boldsymbol{\sigma}_k^1$ separately.

Next we need a partition of unity. Let $\theta_{\boldsymbol{x}}$ be a continuous piecewise linear function defined on $\mathcal{T}_k$ such that the support of each $\theta_{\boldsymbol{x}}$ is in $\Omega_{k,\boldsymbol{x}}$ and

$$|\theta_{\boldsymbol{x}}|_{W^{j,\infty}(\Omega)} \lesssim h_k^{-j}, \quad j = 0, 1,$$

where $W^{j,\infty}(\Omega)$ is the Sobolev space with index (j, ∞) [14]. Such kind of partition of unity exists and has been widely used in the analysis of overlapping Schwarz domain decomposition methods.

Clearly we have the decomposition $\mathbf{curl}\, \boldsymbol{q}_k = \sum_{\boldsymbol{x}\in\mathcal{V}_k} \mathbf{curl}\,(\Pi_{\mathbf{U}_k}(\theta_{\boldsymbol{x}}\boldsymbol{q}_k))$, where $\Pi_{\mathbf{U}_k}$ is the natural interpolation onto $\mathbf{U}_k$ associated with the degrees of freedom. By the standard scaling argument, it is not hard to see that

$$\|\mathbf{curl}\,(\Pi_{\mathbf{U}_k}(\theta_{\boldsymbol{x}}\boldsymbol{q}_k))\|_{0,\Omega} \lesssim \|\mathbf{curl}\,(\theta_{\boldsymbol{x}}\boldsymbol{q}_k)\|_{0,\Omega}.$$

Combining the above, Lemma 4.2 and the fact that $h_k = h_{k-1}/2$, we have

$$
\begin{aligned}
\sum_{\boldsymbol{x}\in\mathcal{V}_k} &\boldsymbol{\Lambda}(\mathbf{curl}\,(\Pi_{\mathbf{U}_k}(\theta_{\boldsymbol{x}}\boldsymbol{q}_k)),\mathbf{curl}\,(\Pi_{\mathbf{U}_k}(\theta_{\boldsymbol{x}}\boldsymbol{q}_k))) \\
&= \sum_{\boldsymbol{x}\in\mathcal{V}_k} (\mathcal{A}\mathbf{curl}\,(\Pi_{\mathbf{U}_k}(\theta_{\boldsymbol{x}}\boldsymbol{q}_k)),\mathbf{curl}\,(\Pi_{\mathbf{U}_k}(\theta_{\boldsymbol{x}}\boldsymbol{q}_k))) \\
&\leq \sum_{\boldsymbol{x}\in\mathcal{V}_k} \|\mathbf{curl}\,(\Pi_{\mathbf{U}_k}(\theta_{\boldsymbol{x}}\boldsymbol{q}_k))\|_{0,\Omega}^2 \leq \sum_{\boldsymbol{x}\in\mathcal{V}_k} \|\mathbf{curl}\,(\theta_{\boldsymbol{x}}\boldsymbol{q}_k)\|_{0,\Omega} \\
&\lesssim \|\mathbf{curl}\,\boldsymbol{q}_k\|_{0,\Omega}^2 + h_k^{-2}\|\boldsymbol{q}_k\|_{0,\Omega}^2 \lesssim (\mathcal{A}\mathbf{curl}\,\boldsymbol{q}_k,\mathbf{curl}\,\boldsymbol{q}_k) \\
&\leq \boldsymbol{\Lambda}(\boldsymbol{\sigma}_k,\boldsymbol{\sigma}_k).
\end{aligned}
\tag{5.5}
$$

Next, for $\boldsymbol{\sigma}_k^1$, we consider the decomposition $\boldsymbol{\sigma}_k^1 = \sum_{\boldsymbol{x}\in\mathcal{V}_k}\Pi_{\boldsymbol{\Sigma}_k}(\theta_{\boldsymbol{x}}\boldsymbol{\sigma}_k^1)$ where $\Pi_{\boldsymbol{\Sigma}_k}$ is the interpolation operator as defined in (3.3) on the mesh $\mathcal{T}_k$. This decomposition is well-defined since $\Pi_{\boldsymbol{\Sigma}_k}$ is a linear operator. By the standard scaling argument, we have

$$\|\Pi_{\boldsymbol{\Sigma}_k}(\theta_{\boldsymbol{x}}\boldsymbol{\sigma}_k^1)\|_{0,\Omega} \lesssim \|\theta_{\boldsymbol{x}}\boldsymbol{\sigma}_k^1\|_{0,\Omega}.$$

Therefore, by Lemma 4.2 and the fact that $2h_k = h_{k-1}$,

$$
\begin{aligned}
\sum_{\boldsymbol{x}\in\mathcal{V}_k} &\boldsymbol{\Lambda}(\Pi_{\boldsymbol{\Sigma}_k}(\theta_{\boldsymbol{x}}\boldsymbol{\sigma}_k^1),\Pi_{\boldsymbol{\Sigma}_k}(\theta_{\boldsymbol{x}}\boldsymbol{\sigma}_k^1)) \\
&\lesssim \sum_{\boldsymbol{x}\in\mathcal{V}_k} (\|\Pi_{\boldsymbol{\Sigma}_k}(\theta_{\boldsymbol{x}}\boldsymbol{\sigma}_k^1)\|_{0,\Omega}^2 + \gamma\|\mathbf{div}(\Pi_{\boldsymbol{\Sigma}_k}(\theta_{\boldsymbol{x}}\boldsymbol{\sigma}_k^1))\|_{0,\Omega}^2) \\
&\lesssim \|\boldsymbol{\sigma}_k^1\|_{0,\Omega}^2 + \gamma(\|\mathbf{div}\boldsymbol{\sigma}_k^1\|_{0,\Omega}^2 + h_k^{-2}\|\boldsymbol{\sigma}_k^1\|_{0,\Omega}^2) \\
&\lesssim (1+\gamma^{-2})\boldsymbol{\Lambda}(\boldsymbol{\sigma}_k,\boldsymbol{\sigma}_k).
\end{aligned}
\tag{5.6}
$$

Finally, we define the decomposition

$$\boldsymbol{\sigma}_k = \sum_{\boldsymbol{x}\in\mathcal{V}_k} (\mathbf{curl}\,(\Pi_{\mathbf{U}_k}(\theta_{\boldsymbol{x}}\boldsymbol{q}_k)) + \Pi_{\boldsymbol{\Sigma}_k}(\theta_{\boldsymbol{x}}\boldsymbol{\sigma}_k^1)).$$

The stability of this decomposition comes from equations (5.5) and (5.6). This completes the proof of the lemma. $\square$

By Inequality (5.2), Lemma 5.1 and Lemma 5.2, the convergence rate of the multigrid solver is $O(\frac{1+\gamma^{-2}}{1+\gamma^{-2}+m})$.

6. Numerical Results. We report some numerical results for the multigrid solver for the $\boldsymbol{H}(\mathbf{div})$-type problem (5.1). Let Ω be the unit square $(0,1)\times(0,1)$. Multigrid solver using a smoother as defined in Section 5 is used. The coarsest mesh is as shown in Figure 6.1. We solve the problem (5.1) on the 2nd, 3rd, 4th and 5th level respectively. The right hand side $\boldsymbol{F}\in\boldsymbol{\Sigma}_h$ is selected randomly.
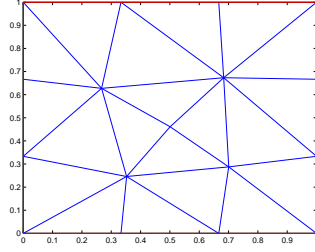
FIG. 6.1. *The coarsest mesh.*

Both the lowest order RT element and the lowest order BDM element are used in the discretization. The experiments are done for $\gamma = 10^{-4}$, 10^{-2}, 1, 10^2, 10^4. We tested the iteration numbers needed for reaching a relative residual 10^{-6}. The results are listed in Table 1 and Table 2.

TABLE 6.1
Iteration numbers using the RT element.

level	$\gamma = 10^{-4}$	$\gamma = 10^{-2}$	$\gamma = 1$	$\gamma = 10^2$	$\gamma = 10^4$
2	22	13	11	11	11
3	62	18	14	14	14
4	132	19	16	15	16
5	200	19	16	16	16

TABLE 6.2
Iteration numbers using the BDM element.

level	$\gamma = 10^{-4}$	$\gamma = 10^{-2}$	$\gamma = 1$	$\gamma = 10^2$	$\gamma = 10^4$
2	6	6	6	6	6
3	7	7	7	7	7
4	7	7	8	7	7
5	8	7	7	7	7

From the tables we can see that when γ is large, the iteration numbers seems to be independent of γ and the number of levels, which agrees with the theory. When γ is small, the theory suggests that the convergence rate deteriorates. From the numerical results we can see that this is true with the RT element. However, for the BDM element, the convergence rate seems to be still independent of γ.

Appendix A. Proof of Equation (4.6).

As stated in the beginning of this paper, we assume Ω is a convex polygon. Let $\boldsymbol{\sigma}_h$ and $\boldsymbol{q}_h$ be defined as in Lemma 4.2. Then $\boldsymbol{q}_h \in \boldsymbol{H}^1(\Omega, \mathbb{R}^2)/\mathbb{R}^2$ and for all $\boldsymbol{p}_H \in \tilde{\mathbf{U}}_H$, we have

$$(\mathcal{A}\mathbf{curl}\,\boldsymbol{q}_h, \mathbf{curl}\,\boldsymbol{p}_H) = \boldsymbol{\Lambda}(\mathbf{curl}\,\boldsymbol{q}_h, \mathbf{curl}\,\boldsymbol{p}_H) = \boldsymbol{\Lambda}(\boldsymbol{\sigma}_h, \mathbf{curl}\,\boldsymbol{p}_H) = 0. \qquad (A.1)$$

Denote $\boldsymbol{n}$ and $\boldsymbol{t}$ the unit outer normal vector and the unit tangential vector along

$\partial\Omega$. Define the curl operator for all $\boldsymbol{\tau} \in \boldsymbol{H}^1(\Omega, \mathbb{M}_2)$ by

$$\operatorname{curl}\boldsymbol{\tau} = \begin{pmatrix} -\frac{\partial\tau_{11}}{\partial y} + \frac{\partial\tau_{12}}{\partial x} \\ -\frac{\partial\tau_{21}}{\partial y} + \frac{\partial\tau_{22}}{\partial x} \end{pmatrix}.$$

First we notice that:

LEMMA A.1. *Assume there exists $\boldsymbol{\sigma} \in \boldsymbol{H}^1(\Omega, \mathbb{M}_2)$ satisfying $\boldsymbol{q}_h = \operatorname{curl}\boldsymbol{\sigma}$ and*
(S1) $\mathcal{A}\boldsymbol{\sigma} = \boldsymbol{\sigma}$ ($\boldsymbol{\sigma}$ is trace-free);
(S2) $(\boldsymbol{\sigma t})|_{\partial\Omega} = \boldsymbol{0}$;
(S3) $\|\boldsymbol{\sigma}\|_{1,\Omega} \lesssim \|\boldsymbol{q}_h\|_{0,\Omega}$.
Then Equation (4.6) holds.

Proof. Define $\boldsymbol{\sigma}^1$ by

$$\begin{cases} (\mathcal{A}\boldsymbol{\sigma}^1, \boldsymbol{\tau}) + (\mathbf{div}\boldsymbol{\tau}, \boldsymbol{u}) = 0, & \text{for all } \boldsymbol{\tau} \in \Sigma, \\ (\mathbf{div}\boldsymbol{\sigma}^1, \boldsymbol{v}) = (\mathbf{div}\boldsymbol{\sigma}, \boldsymbol{v}), & \text{for all } \boldsymbol{v} \in \mathbf{V}. \end{cases}$$

The regularity result (2.3) implies $\|\boldsymbol{\sigma}^1\|_{1,\Omega} \lesssim \|\mathbf{div}\boldsymbol{\sigma}\|_{0,\Omega}$. Since $\boldsymbol{\sigma} - \boldsymbol{\sigma}^1 \in \boldsymbol{H}^1(\Omega, \mathbb{M}_2)$ and $\mathbf{div}(\boldsymbol{\sigma} - \boldsymbol{\sigma}^1) = \boldsymbol{0}$, there must exist a potential function $\boldsymbol{p} \in \boldsymbol{H}^2(\Omega, \mathbb{R}^2)$ such that $\mathbf{curl}\,\boldsymbol{p} = \boldsymbol{\sigma} - \boldsymbol{\sigma}^1$ and

$$|\boldsymbol{p}|_{2,\Omega} \le \|\boldsymbol{\sigma}\|_{1,\Omega} + \|\boldsymbol{\sigma}^1\|_{1,\Omega} \lesssim \|\boldsymbol{\sigma}\|_{1,\Omega} \lesssim \|\boldsymbol{q}_h\|_{0,\Omega}.$$

Let $\boldsymbol{p}_H$ be the natural interpolation of $\boldsymbol{p}$ onto $\tilde{\mathbf{U}}_H$. Then by the assumptions on $\boldsymbol{\sigma}$, Equation (A.1) and integration by parts,

$$\|\boldsymbol{q}_h\|_{0,\Omega}^2 = (\operatorname{curl}\boldsymbol{\sigma}, \boldsymbol{q}_h) = -(\boldsymbol{\sigma}, \mathbf{curl}\,\boldsymbol{q}_h) = -(\mathcal{A}\boldsymbol{\sigma}, \mathbf{curl}\,\boldsymbol{q}_h)$$

$$= -(\mathcal{A}\mathbf{curl}\,\boldsymbol{p}, \mathbf{curl}\,\boldsymbol{q}_h) = -(\mathcal{A}\mathbf{curl}\,(\boldsymbol{p} - \boldsymbol{p}_H), \mathbf{curl}\,\boldsymbol{q}_h)$$

$$\lesssim H|\boldsymbol{p}|_{2,\Omega}\|\mathcal{A}\mathbf{curl}\,\boldsymbol{q}_h\|_{0,\Omega} \lesssim H\|\boldsymbol{q}_h\|_{0,\Omega}\|\mathcal{A}\mathbf{curl}\,\boldsymbol{q}_h\|_{0,\Omega},$$

which will give Equation (4.6). $\square$

In the remainder of this section, we will construct a $\boldsymbol{\sigma}$ that satisfies the assumption in Lemma A.1.

For any scalar function f, define $\mathbf{curl}\,f = (-\frac{\partial f}{\partial y}, \frac{\partial f}{\partial x})^t$. Then there exists a $\mathbf{L}^2$ orthogonal Helmholtz decomposition (see [16]) $\boldsymbol{q}_h = \mathbf{curl}\,\phi + \nabla\psi$, where $\phi \in H_0^1(\Omega)$ and $\psi \in H^1(\Omega)/\mathbb{R}$ satisfy $\|\phi\|_{1,\Omega} \le \|\boldsymbol{q}_h\|_{0,\Omega}$ and $\|\psi\|_{1,\Omega} \le \|\boldsymbol{q}_h\|_{0,\Omega}$. Then

$$\boldsymbol{q}_h = \operatorname{curl} \begin{pmatrix} \phi & \psi \\ -\psi & \phi \end{pmatrix} = \operatorname{curl} \begin{pmatrix} \phi & (\psi + C) \\ -(\psi + C) & \phi \end{pmatrix} + \operatorname{curl}\nabla\boldsymbol{u}$$

for all $C \in \mathbb{R}$ and $\boldsymbol{u} \in \boldsymbol{H}^2(\Omega, \mathbb{R}^2)$. Define

$$\boldsymbol{\sigma} = \begin{pmatrix} \phi & (\psi + C) \\ -(\psi + C) & \phi \end{pmatrix} + \nabla\boldsymbol{u}. \tag{A.2}$$

Clearly $\boldsymbol{\sigma} \in \boldsymbol{H}^1(\Omega, \mathbb{M}_2)$ and we only need to find C and $\boldsymbol{u}$ such that $\boldsymbol{\sigma}$ satisfies (S1), (S2) and (S3) in Lemma A.1, which can be translated to

$$\operatorname{div}\boldsymbol{u} = -2\phi, \qquad \frac{\partial\boldsymbol{u}}{\partial\boldsymbol{t}} = (\psi + C)\boldsymbol{n},$$

$$|C| \lesssim \|\boldsymbol{q}_h\|_{0,\Omega}, \qquad \|\boldsymbol{u}\|_{2,\Omega} \lesssim \|\boldsymbol{q}_h\|_{0,\Omega}. \tag{A.3}$$

Since $\boldsymbol{q}_h \in \boldsymbol{H}^1(\Omega, \mathbb{R}^2)/\mathbb{R}^2$ and $\phi \in H_0^1(\Omega)$, we have

$$
\begin{aligned}
\int_{\partial\Omega} \psi \boldsymbol{n} \, ds &= \int_\Omega \nabla \psi \, d\boldsymbol{x} = \int_\Omega (\boldsymbol{q}_h - \mathbf{curl}\, \phi) \, d\boldsymbol{x} \\
&= -\int_\Omega \mathbf{curl}\, \phi \, d\boldsymbol{x} = \int_{\partial\Omega} \phi \boldsymbol{t} \, ds = \boldsymbol{0}.
\end{aligned}
\tag{A.4}
$$

Recall that Ω is a convex polygon. Denote $\boldsymbol{z}_i$, $i = 1, \ldots, N$, the corners of Ω. To close the polygon, we set $\boldsymbol{z}_0 = \boldsymbol{z}_N$ or $\boldsymbol{z}_{N+1} = \boldsymbol{z}_1$. Denote Γ_i the edge of Ω connecting $\boldsymbol{z}_i$ and $\boldsymbol{z}_{i+1}$. Define $\boldsymbol{n}_i = (n_{1,i}, n_{2,i})^t$ and $\boldsymbol{t}_i = (n_{2,i}, -n_{1,i})^t$ to be the unit outer normal vector and the unit tangential vector on Γ_i respectively. According to the trace theorem on polygons (see Theorem 1.4.6 in [17]), we have

$$
\int_0^\varepsilon s^{-1} \left| -\psi(\boldsymbol{z}_i + s\boldsymbol{t}_i) + \psi(\boldsymbol{z}_i - s\boldsymbol{t}_{i-1}) \right|^2 ds \leq \infty,
\tag{A.5}
$$

for all $\varepsilon < min_{i=1,\ldots,N} |z_i - z_{i-1}|$.

Let s_0, s be points on $\partial\Omega$ and $\int_{s_0}^s dt$ be the line integral from s_0 to s along $\partial\Omega$ in counter-clockwise direction. Define

$$
C = \frac{-\int_\Omega 2\phi \, d\boldsymbol{x} - \int_{\partial\Omega} (\int_{s_0}^s \psi \boldsymbol{n} \, dt) \cdot \boldsymbol{n} \, ds}{\int_{\partial\Omega} (\int_{s_0}^s \boldsymbol{n} \, dt) \cdot \boldsymbol{n} \, ds},
\tag{A.6}
$$

and

$$
\boldsymbol{g}(s) = \int_{s_0}^s (\psi + C) \boldsymbol{n} \, dt.
\tag{A.7}
$$

By Equation (A.4) and the Green's formula, we immediately see that $\boldsymbol{g}$ is continuous on $\partial\Omega$. Furthermore, we clearly have $\boldsymbol{g}|_{\Gamma_i} \in \boldsymbol{H}^{3/2}(\Gamma_i, \mathbb{R}^2)$ and

$$
|C| \lesssim \|\phi\|_{0,\Omega} + \|\psi\|_{0,\partial\Omega} \lesssim \|\phi\|_{0,\Omega} + \|\psi\|_{1,\Omega} \lesssim \|\boldsymbol{q}_h\|_{0,\Omega},
$$

$$
\|\boldsymbol{g}\|_{3/2,\Gamma_i} \lesssim \|\psi\|_{1/2,\partial\Omega} + |C| \lesssim \|\phi\|_{0,\Omega} + \|\psi\|_{1,\Omega} \lesssim \|\boldsymbol{q}_h\|_{0,\Omega}.
$$

By (A.5), (A.6) and (A.7), it is not hard to see that $\int_\Omega -2\phi \, d\boldsymbol{x} = \int_{\partial\Omega} \boldsymbol{g} \cdot \boldsymbol{n} \, ds$ and

$$
\begin{aligned}
&\int_0^\varepsilon s^{-1} \left| \frac{\partial \boldsymbol{g}}{\partial \boldsymbol{t}}(z_i + s\boldsymbol{t}_i) \cdot \boldsymbol{n}_{i-1} - \frac{\partial \boldsymbol{g}}{\partial \boldsymbol{t}}(z_i - s\boldsymbol{t}_{i-1}) \cdot \boldsymbol{n}_i \right|^2 ds \\
&= \int_0^\varepsilon s^{-1} \left| (\psi + C)(z_i + s\boldsymbol{t}_i) - (\psi + C)(z_i - s\boldsymbol{t}_{i-1}) \right|^2 (\boldsymbol{n}_{i-1} \cdot \boldsymbol{n}_i)^2 ds \\
&< \infty,
\end{aligned}
$$

for all $\varepsilon < min_{i=1,\ldots,N} |z_i - z_{i-1}|$. Recall that $\phi \in H_0^1(\Omega)$. Combining all the above and according to Theorem 7.1 in [4], there exists $\boldsymbol{u} \in \boldsymbol{H}^2(\Omega, \mathbb{R}^2)$ such that $\operatorname{div} \boldsymbol{u} = -2\phi$, $\boldsymbol{u}|_{\partial\Omega} = \boldsymbol{g}$ and

$$
\|\boldsymbol{u}\|_{2,\Omega} \lesssim \|\phi\|_{1,\Omega} + \sum_{i=1,\ldots,N} \|\boldsymbol{g}\|_{3/2,\Gamma_i} \lesssim \|\phi\|_{1,\Omega} + \|\psi\|_{1,\Omega} \lesssim \|\boldsymbol{q}_h\|_{0,\Omega}.
$$

Clearly $\boldsymbol{u}$ defined above and C defined in (A.6) satisfy Equation (A.3), and hence $\boldsymbol{\sigma}$ defined as in (A.2) satisfies (S1), (S2) and (S3). This completes the proof for Equation (4.6).

REFERENCES

[1] D. N. Arnold, R. S. Falk, and R. Winther. Preconditioning in H(div) and applications. *Math. Comp.*, 66:957–984, 1997.

[2] D. N. Arnold, R. S. Falk, and R. Winther. Multigrid in H(div) and H(curl). *Numer. Math.*, 85:197–217, 2000.

[3] D. N. Arnold, J. Jim Douglas, and C. P. Gupta. A family of higher order mixed finite element methods for plane elasticity. *Numer. Math.*, 45:1–22, 1984.

[4] D. N. Arnold, L. R. Scott, and M. Vogelius. Regular inversion of the divergence operator with Dirichlet boundary conditions on a polygon. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.-Serie IV, XV,*, pages 169–192, 1988.

[5] D. N. Arnold and R. Winther. Mixed finite element for elasticity. *Numer. Math.*, 92:401–419, 2002.

[6] F. Baaijens. Mixed finite element methods for viscoelastic flow analysis: A review. *J. Non-Newtonian Fluid Mech.*, 79:361–385, 1998.

[7] M. Behr, L. France, and T. Tezduyar. Stabilized finite element methods for the velocity-pressure-stress formulation of incompressible flows. *Comput. Methods Appl. Mech. Engrg.*, 104:31–48, 1993.

[8] J. Bramble. *Multigrid methods*. Pitman Research Notes in Mathematics Series. Pitman, London, 1993.

[9] F. Brezzi, J. Douglas, and L. Marini. Two families of mixed finite elements for second order elliptic problem. *Numer. Math.*, 47:217–235, 1985.

[10] F. Brezzi and M. Fortin. *Mixed and hybrid finite element methods*. Springer-Verlag, New York, 1991.

[11] Z. Cai, B. Lee, and P. Wang, *Least-squares methods for incompressible Newtonian fluid flow: linear stationary problems*, SIAM J. Numer. Anal., 42:843–859, 2004.

[12] 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.

[13] Z. Cai, C. Tong, P. Vassilevski, and C. Wang, *Mixed finite element methods for incompressible flow: stationary Stokes equations*, manuscript, 2005.

[14] P. G. Ciarlet. *The finite element method for elliptic problems*. North-Holland, Amsterdam, 1978.

[15] M. Gerritsma and T. Phillips. Compatible spectral approximations for the velocity-stress-pressure formulation of the Stokes problem. *SIAM J. Sci. Comput.*, 20:1530–1550, 1999.

[16] 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.

[17] P. Grisvard. *Singularities in boundary value problems*. Number 22 in Research notes in applied mathematics. Springer-Verlag, New York, 1992.

[18] W. Hackbusch. *Multi-grid methods and applications*. Springer-Verlag, Berlin, 1985.

[19] R. Hiptmair. Multigrid method for H(div) in three dimensions. *ETNA*, 6:133–152, 1997.

[20] C. Johnson and B. Mercier. Some equilibrium finite element methods for two-dimensional elasticity problems. *Numer. Math.*, 30:103–116, 1978.

[21] R. Keunings. Advances in the computer modeling of the flow of polymeric liquids. *Comp. Fluid Dyn. J.*, 9:449–458, 2001.

[22] J. Mandel, S. McCormick, and R. Bank, *Variational multigrid theory*, Multigrid Methods (S. McCormick ed.), SIAM, Philadelphia, Penn, 131-178.

[23] P. Raviart and J. Thomas. A mixed finite element method for second order elliptic problems. In I. Galligani and E. Magenes, editors, *Mathematical aspects of the finite element method*, Lecture notes in Math. 606. Springer-Verlag, 1977.

[24] P. S. Vassilevski and J. Wang, *Multilevel iterative methods for mixed finite element discretizations of elliptic problems*, Numer. Math., 63:503–520, 1992.

[25] J. Xu. Iterative methods by space decomposition and subspace correction. *SIAM Rev.*, 34:581–613, 1992.