SIAM J. Numer. Anal.
Vol. 44, No. 2, pp. 478–493

A MULTIGRID PRECONDITIONER FOR THE MIXED FORMULATION OF LINEAR PLANE ELASTICITY*

JOSEPH E. PASCIAK† AND YANQIU WANG‡

Abstract. In this paper, we develop a multigrid preconditioner for the discrete system of linear equations that results from the mixed formulation of the linear plane elasticity problem using the Arnold–Winther elements. This, in turn, can be reduced to the problem of finding a multigrid preconditioner for the form $(\cdot,\cdot) + (\mathbf{div}\,\cdot, \mathbf{div}\,\cdot)$ in the symmetric matrix space resulting from Arnold–Winther elements. Since the form is not uniformly elliptic, a Helmholtz-type decomposition is essential. The Arnold–Winther finite element space gives rise to nonnested multilevel spaces adding difficulty to the analysis. We prove that for the variable V-cycle multigrid preconditioner, the condition number of the preconditioned system is independent of the number of levels. The results of numerical experiments are also presented.

Key words. multigrid, mixed finite element, linear elasticity

AMS subject classifications. 65N30, 65N55

DOI. 10.1137/040617820

1. Introduction.

Mixed finite element methods [7, 16] 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 (in this case the stress), which is often 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. 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. Most of this research was focused on developing stable pairs of mixed finite element spaces, and several different solutions have been proposed [5, 6, 26]. 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 [5, 26] resort to composite elements. Only recently did Arnold and Winther construct a stable pair of mixed finite elements [6] which did not use composite elements. The Arnold–Winther finite element spaces consist of piecewise polynomials over a triangular mesh tied together by degrees of freedom resulting in $\boldsymbol{H}(\mathbf{div})$ conforming symmetric approximation subspaces.

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 matrices [1, 4].

*Received by the editors October 28, 2004; accepted for publication (in revised form) July 18, 2005; published electronically March 15, 2006.
http://www.siam.org/journals/sinum/44-2/61782.html
†Department of Mathematics, Texas A&M University, College Station, TX 77843 (pasciak@math.tamu.edu).
‡Department of Mathematics, Purdue University, West Lafayette, IN 47907 (yqwang@math.purdue.edu).

Another way is to use the least-square formulation so that the classical discrete inf-sup condition is no longer needed [10, 17, 18]. Finally, other authors resort to the use of stabilizing techniques (see [22] and the references therein).

In this paper, we will focus on the lowest order Arnold–Winther finite element. The purpose is to develop and analyze a multigrid preconditioner for the resulting discrete system.

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. The first approach is to use Uzawa-type methods [9, 11, 20]. The second is the positive definite reformulation proposed by Bramble and Pasciak in [12] and [13]. The third is the preconditioned minimum residual method analyzed in [2, 27]. We adopt the idea of the preconditioned minimum residual method. An analysis similar to the one in [2] 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 matrix field.

In this paper, we construct and analyze a multigrid preconditioner for the $\boldsymbol{H}(\mathbf{div})$ problem. Multigrid methods provide efficient preconditioners for second order elliptic problems. A vast amount of research has been done in this area [15, 24, 29]. However, the classical techniques for the multigrid preconditioner do not work for the $\boldsymbol{H}(\mathbf{div})$ problem since the discrete operator which results from the $\boldsymbol{H}(\mathbf{div})$ problem is not uniformly elliptic. To deal with this difficulty, we follow the idea of using a Helmholtz-like decomposition [2, 3, 8, 21, 25] and decompose the Arnold–Winther finite element space into two orthogonal subspaces: the subspace of divergence-free functions and its orthogonal complement. Then, the analysis of our preconditioners can be done on these two subspaces separately. Our results show that for convex polygonal domains and 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.

The outline of the remainder of the paper is as follows. In section 2, we briefly introduce the mixed formulation of the elasticity problem, the Arnold–Winther mixed finite element for (2.3) and the technique for preconditioning a mixed system proposed in [2]. In section 3, the details of the multigrid preconditioner are explained, and the condition number of the preconditioned system is analyzed under certain assumptions on the smoother. In section 4, we construct a smoother and prove that it satisfies the assumptions stated in section 3. Finally, we give results of numerical experiments in section 5.

2. The mixed problem formulation, discretization, and preconditioning.

In this section, we first state the mixed form of the linear elasticity problem. Next, we introduce the Arnold–Winther elements of lowest order. Finally, we briefly describe the idea of preconditioning the mixed system introduced in [2] which reduces the preconditioning problem to one on $\boldsymbol{H}_0(\mathbf{div}\,, \Omega, \mathbb{S}_2)$ (defined below).

2.1. Mixed elasticity formulation.

Let Ω be a convex polygon in $\mathbb{R}^2$. We use the usual notation $H^s(\Omega)$, where s is a real number, to denote the Sobolev space defined on Ω [19]. For $s = 0$, the space is also denoted by $L^2(\Omega)$. Define $H_0^s(\Omega)$ to be the closure of $C_0^\infty(\Omega)$ under the $H^s(\Omega)$ norm.

Let $\mathbb{R}^2$ be the space of two-dimensional vector functions and $\mathbb{S}_2$ be the space of symmetric 2×2 matrix functions defined on Ω. Throughout the paper, we adopt the convention that bold Latin characters in lower case denote vectors and bold Greek characters denote 2×2 symmetric matrices. Let $\boldsymbol{\tau} = (\tau_{ij})_{1\le i,j\le 2} \in \mathbb{S}_2$, $\boldsymbol{v} = (v_i)_{1\le i\le 2} \in \mathbb{R}^2$, and q be a scalar function. Define $\operatorname{div} \boldsymbol{v} = \frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y}$ and

$$\mathbf{div}\,\boldsymbol{\tau} = \begin{pmatrix} \frac{\partial \tau_{11}}{\partial x} + \frac{\partial \tau_{12}}{\partial y} \\ \frac{\partial \tau_{21}}{\partial x} + \frac{\partial \tau_{22}}{\partial y} \end{pmatrix}, \qquad \mathbf{airy}\, q = \begin{pmatrix} \frac{\partial^2 q}{\partial y^2} & -\frac{\partial^2 q}{\partial x \partial y} \\ -\frac{\partial^2 q}{\partial x \partial y} & \frac{\partial^2 q}{\partial x^2} \end{pmatrix}. \tag{2.1}$$

Denote the inner product between vectors and the inner product between matrices by

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

We generalize the definition of the Sobolev space to the cases of vector functions and symmetric matrix functions. Define the spaces

$$\boldsymbol{H}^s(\Omega, \mathbb{R}^2) = (H^s(\Omega))^2, \qquad \boldsymbol{H}^s(\Omega, \mathbb{S}_2) = (H^s(\Omega))^3$$

with norms

$$\|\boldsymbol{v}\|_{\boldsymbol{H}^s(\Omega,\mathbb{R}^2)} = (\|v_1\|^2_{H^s(\Omega)} + \|v_2\|^2_{H^s(\Omega)})^{1/2},$$
$$\|\boldsymbol{\tau}\|_{\boldsymbol{H}^s(\Omega,\mathbb{S}_2)} = (\|\tau_{11}\|^2_{H^s(\Omega)} + 2\|\tau_{12}\|^2_{H^s(\Omega)} + \|\tau_{22}\|^2_{H^s(\Omega)})^{1/2}.$$

We define $\boldsymbol{L}^2(\Omega, \mathbb{R}^2)$ and $\boldsymbol{L}^2(\Omega, \mathbb{S}_2)$ in the same fashion. For simplicity, denote $\|\cdot\|_{s,\Omega}$ to be the H^s-norm over scalar, vector, or symmetric 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 Ω.

Define

$$\boldsymbol{H}_0(\mathbf{div}\,, \Omega, \mathbb{S}_2) = \{\boldsymbol{\tau} \in \boldsymbol{L}^2(\Omega, \mathbb{S}_2) : \ \mathbf{div}\,\boldsymbol{\tau} \in \boldsymbol{L}^2(\Omega, \mathbb{R}^2) \text{ and } \boldsymbol{\tau}\boldsymbol{n}|_{\partial\Omega} = \boldsymbol{0}\},$$

where $\boldsymbol{n}$ is the outward normal vector on $\partial\Omega$. The norm on $\boldsymbol{H}_0(\mathbf{div}\,, \Omega, \mathbb{S}_2)$ is defined to be

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

$\boldsymbol{H}_0(\mathbf{div}\,, \Omega, \mathbb{S}_2)$ is a Hilbert space with the inner product

$$\boldsymbol{\Lambda}(\boldsymbol{\sigma}, \boldsymbol{\tau}) = (\boldsymbol{\sigma}, \boldsymbol{\tau}) + (\mathbf{div}\,\boldsymbol{\sigma}, \mathbf{div}\,\boldsymbol{\tau}). \tag{2.2}$$

Next, we state the mixed formulation of the plane elasticity problem. We only consider the pure traction boundary problem [6, 16]: Find the stress $\boldsymbol{\sigma} \in \boldsymbol{H}_0(\mathbf{div}\,, \Omega, \mathbb{S}_2)$ and the displacement $\boldsymbol{u} \in \boldsymbol{L}^2(\Omega, \mathbb{R}^2)$ satisfying

$$\begin{cases} \int_\Omega \mathbb{A}\boldsymbol{\sigma} : \boldsymbol{\tau}\, d\boldsymbol{x} + \int_\Omega \mathbf{div}\,\boldsymbol{\tau} \cdot \boldsymbol{u}\, d\boldsymbol{x} = 0 & \text{for all } \boldsymbol{\tau} \in \boldsymbol{H}_0(\mathbf{div}\,, \Omega, \mathbb{S}_2), \\ \int_\Omega \mathbf{div}\,\boldsymbol{\sigma} \cdot \boldsymbol{v}\, d\boldsymbol{x} = \int_\Omega \boldsymbol{g} \cdot \boldsymbol{v}\, d\boldsymbol{x} & \text{for all } \boldsymbol{v} \in \boldsymbol{L}^2(\Omega, \mathbb{R}^2). \end{cases} \tag{2.3}$$

Here the fourth order compliance tensor $\mathbb{A}$ is bounded, symmetric, and uniformly positive definite the body force per unit volume $\boldsymbol{g}$ is in $\boldsymbol{L}^2(\Omega, \mathbb{R}^2)$. For (2.3) to be well posed, we need a compatibility condition on $\boldsymbol{g}$. Let

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

be the space of infinitesimal rigid motions. By Korn's inequality, one can see that for any $\boldsymbol{g} \in \boldsymbol{L}^2(\Omega, \mathbb{R}^2)/RM$ (the orthogonal complement of RM in $\boldsymbol{L}^2(\Omega, \mathbb{R}^2)$), system (2.3) has a unique solution in $\boldsymbol{H}_0(\mathbf{div}\,, \Omega, \mathbb{S}_2) \times \boldsymbol{L}^2(\Omega, \mathbb{R}^2)/RM$ [16].

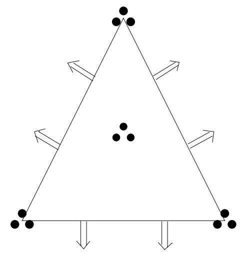

FIG. 2.1. *The Arnold–Winther finite element* $\boldsymbol{\Sigma}_T$.

2.2. Arnold–Winther elements. Let $\mathcal{T}$ be a quasi-uniform triangulation of Ω with characteristic mesh size h. On each triangle $T \in \mathcal{T}$ define

$$\begin{aligned}\boldsymbol{\Sigma}_T &= \{\text{symmetric matrices } \boldsymbol{\tau} \in (P_3(T))^3 \text{ such that } \mathbf{div}\, \boldsymbol{\tau} \in (P_1(T))^2\},\\ \boldsymbol{V}_T &= (P_1(T))^2,\end{aligned}$$

where $P_i(T)$ denotes the space consisting of polynomials of degree i or less. The degrees of freedom (dofs) for $\boldsymbol{\Sigma}_T$ are

- the nodal values of the three components of $\boldsymbol{\tau}(x)$ 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 dofs of $\boldsymbol{V}_T$ are given as the zeroth and first order moments on T. Figure 2.1 illustrates the dofs for $\boldsymbol{\Sigma}_T$. The finite element spaces on the mesh $\mathcal{T}$ and domain Ω are defined as follows:

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

The definition of $\boldsymbol{\Sigma}(\mathcal{T}, \Omega)$ implies that $\boldsymbol{\Sigma}(\mathcal{T}, \Omega) \subset \boldsymbol{H}_0(\mathbf{div}\,, \Omega, \mathbb{S}_2)$ (see [6, 16]). Note that the boundary condition $\boldsymbol{\tau}\boldsymbol{n}|_{\partial\Omega} = \mathbf{0}$ implies two linear relations among the three components of $\boldsymbol{\tau}$ on boundary nodes. Hence on the corner vertices where two boundary edges meet, we will have $\boldsymbol{\tau} = \mathbf{0}$. This fact was noticed by Arnold and Winther in [6]. Another immediate observation is that by Green's formula,

$$\mathbf{div}\, \boldsymbol{\tau} \in RM^{\perp_{\boldsymbol{V}(\mathcal{T},\Omega)}} \qquad \text{for all } \boldsymbol{\tau} \in \boldsymbol{\Sigma}(\mathcal{T}, \Omega).$$

The discrete elasticity problem can be written as follows: find $\boldsymbol{\sigma}_h \in \boldsymbol{\Sigma}(\mathcal{T}, \Omega)$ and $\boldsymbol{u}_h \in \boldsymbol{V}(\mathcal{T}, \Omega)$ such that

$$\begin{cases}(\mathbb{A}\boldsymbol{\sigma}_h, \boldsymbol{\tau}) + (\mathbf{div}\, \boldsymbol{\tau}, \boldsymbol{u}_h) = 0 & \text{for all } \boldsymbol{\tau} \in \boldsymbol{\Sigma}(\mathcal{T}, \Omega),\\ (\mathbf{div}\, \boldsymbol{\sigma}_h, \boldsymbol{v}) = (\boldsymbol{g}, \boldsymbol{v}) & \text{for all } \boldsymbol{v} \in \boldsymbol{V}(\mathcal{T}, \Omega).\end{cases} \tag{2.4}$$

Arnold and Winther have proved that the Arnold–Winther finite element spaces (without the essential boundary condition $\boldsymbol{\tau}\boldsymbol{n}|_{\partial\Omega} = \mathbf{0}$) satisfy the LBB condition [6]. In [28], it was proved that the Arnold–Winther finite element spaces $(\boldsymbol{\Sigma}(\mathcal{T}, \Omega), \boldsymbol{V}(\mathcal{T}, \Omega))$ (with the essential boundary condition $\boldsymbol{\tau}\boldsymbol{n}|_{\partial\Omega} = \mathbf{0}$) also satisfy the LBB condition.

Furthermore, the assumption on $\mathbb{A}$ implies that there exists a positive constant c such that

$$(\mathbb{A}\boldsymbol{\tau}, \boldsymbol{\tau}) \geq c\|\boldsymbol{\tau}\|^2_{\boldsymbol{H}(\mathbf{div}\,,\Omega,\mathbb{S}_2)} \quad \text{for all } \boldsymbol{\tau} \in \boldsymbol{H}_0(\mathbf{div}\,,\Omega,\mathbb{S}_2) \text{ with } \mathbf{div}\,\boldsymbol{\tau} = \mathbf{0}. \tag{2.5}$$

Combining these results shows that problem (2.4) has a unique solution (for compatible $\boldsymbol{g}$) in $(\boldsymbol{\Sigma}(\mathcal{T},\Omega), \boldsymbol{V}(\mathcal{T},\Omega)/RM)$. Furthermore, if $(\boldsymbol{\sigma}, \boldsymbol{u})$ is the solution of the weak problem (2.3) and $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h)$ is the solution of the discrete problem (2.4), we have the following error estimates [6, 28]:

$$\begin{aligned} \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,\Omega} m &\leq ch^m \|\boldsymbol{\sigma}\|_{m,\Omega}, && 1 \leq m \leq 3, \\ \|\mathbf{div}\,\boldsymbol{\sigma} - \mathbf{div}\,\boldsymbol{\sigma}_h\|_{0,\Omega}, &\leq ch^m \|\mathbf{div}\,\boldsymbol{\sigma}\|_{m,\Omega}, && 0 \leq m \leq 2, \\ \|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{L}^2(\Omega,\mathbb{R}^2)/RM}, &\leq ch^m \|\boldsymbol{u}\|_{m+1,\Omega}, && 1 \leq m \leq 2, \end{aligned} \tag{2.6}$$

where c is a positive constant independent of h.

Next, we introduce the Argyris element, which plays an important role in later analysis. Let Q_T denote the Argyris element [19] defined on T. It is a quintic element and the dofs are

- the function value on each vertex (three dofs), the first derivatives at each vertex (six dofs), and the second derivatives at each vertex (nine dofs);
- the moments of degree 0 of the normal derivative on each edges of T (three dofs).

Define the space

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

Clearly $\mathrm{Q}(\mathcal{T},\Omega) \subset H^2_0(\Omega)$.

Similar to the De Rham sequence, it is elementary to see that the following exact sequence holds [6, 28]:

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

Recall that operators in an exact sequence have the property that the range of the operator on the left equals the kernel of the operator on the right.

We can define an operator $\mathbf{div}^{-1} :\ \boldsymbol{L}^2(\Omega,\mathbb{R}^2)/RM \to \boldsymbol{H}_0(\mathbf{div}\,,\Omega,\mathbb{S}_2)/Ker(\mathbf{div}\,)$ as follows. For $\boldsymbol{v} \in \boldsymbol{L}^2(\Omega,\mathbb{R}^2)/RM$, let $\boldsymbol{\sigma} \in \boldsymbol{H}(\mathbf{div}\,,\Omega,\mathbb{S}_2)$ and $\boldsymbol{u} \in \boldsymbol{L}^2(\Omega,\mathbb{R}^2)$ satisfy

$$\begin{cases} (\boldsymbol{\sigma},\boldsymbol{\tau}) + (\mathbf{div}\,\boldsymbol{\tau},\boldsymbol{u}) = 0 & \text{for all } \boldsymbol{\tau} \in \boldsymbol{H}_0(\mathbf{div}\,,\Omega,\mathbb{S}_2), \\ (\mathbf{div}\,\boldsymbol{\sigma},\boldsymbol{w}) = (\boldsymbol{v},\boldsymbol{w}) & \text{for all } \boldsymbol{w} \in \boldsymbol{L}^2(\Omega,\mathbb{R}^2). \end{cases} \tag{2.7}$$

Since $\mathbf{div}$ maps $\boldsymbol{H}_0(\mathbf{div}\,,\Omega,\mathbb{S}_2)$ onto $\boldsymbol{L}^2(\Omega,\mathbb{R}^2)/RM$, system (2.7) admits a unique solution in $(\boldsymbol{H}_0(\mathbf{div}\,,\Omega,\mathbb{S}_2), \boldsymbol{L}^2(\Omega,\mathbb{R}^2)/RM)$ (see [16]). Then, set $\mathbf{div}^{-1}\boldsymbol{v} = \boldsymbol{\sigma}$. By definition, $\mathbf{div}^{-1}\boldsymbol{v}$ is orthogonal to any divergence free function in $\boldsymbol{H}_0(\mathbf{div}\,,\Omega,\mathbb{S}_2)$ under both the $\boldsymbol{L}^2$ inner product and the $\boldsymbol{H}_0(\mathbf{div}\,,\Omega,\mathbb{S}_2)$ inner product. Therefore, for all $\boldsymbol{\tau} \in \boldsymbol{H}_0(\mathbf{div}\,,\Omega,\mathbb{S}_2)$, we have a unique orthogonal decomposition

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

where $q \in H^2_0(\Omega)$ and $\boldsymbol{v} = \mathbf{div}\,\boldsymbol{\tau}$. Furthermore, we have the regularity result (see [23]),

$$\mathbf{div}^{-1}\boldsymbol{v} \in \boldsymbol{H}^1(\Omega,\mathbb{S}_2) \quad \text{and} \quad \|\mathbf{div}^{-1}\boldsymbol{v}\|_{1,\Omega} \leq c\|\boldsymbol{v}\|_{0,\Omega}, \tag{2.8}$$

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

Analogously, on the discrete level we have the following exact sequence:

$$0 \xrightarrow{\subset} \mathrm{Q}(\mathcal{T},\Omega) \xrightarrow{\mathbf{airy}} \boldsymbol{\Sigma}(\mathcal{T},\Omega) \xrightarrow{\mathbf{div}} \boldsymbol{V}(\mathcal{T},\Omega)/RM \rightarrow 0. \tag{2.9}$$

The exactness of this sequence for the Arnold–Winther finite element spaces follows from [6]. We define an operator $\mathbf{div}_{\mathcal{T}}^{-1} : \boldsymbol{L}^2(\Omega,\mathbb{R}^2)/RM \rightarrow \boldsymbol{\Sigma}(\mathcal{T},\Omega)/Ker(\mathbf{div})$ as follows. For $\boldsymbol{v} \in \boldsymbol{L}^2(\Omega,\mathbb{R}^2)/RM$, let $\boldsymbol{\sigma}_h \in \boldsymbol{\Sigma}(\mathcal{T},\Omega)$ and $\boldsymbol{u}_h \in \boldsymbol{V}(\mathcal{T},\Omega)$ satisfy

$$\begin{cases} (\boldsymbol{\sigma}_h,\boldsymbol{\tau}) + (\mathbf{div}\,\boldsymbol{\tau},\boldsymbol{u}_h) = 0 & \text{for all } \boldsymbol{\tau} \in \boldsymbol{\Sigma}(\mathcal{T},\Omega), \\ (\mathbf{div}\,\boldsymbol{\sigma}_h,\boldsymbol{w}) = (\boldsymbol{v},\boldsymbol{w}) & \text{for all } \boldsymbol{w} \in \boldsymbol{V}(\mathcal{T},\Omega). \end{cases} \tag{2.10}$$

Since the Arnold–Winther finite element spaces satisfy the LBB condition, the solution to (2.10) exists and is unique in $(\boldsymbol{\Sigma}(\mathcal{T},\Omega), \boldsymbol{V}(\mathcal{T},\Omega)/RM)$. Define $\mathbf{div}_{\mathcal{T}}^{-1}\boldsymbol{v} = \boldsymbol{\sigma}_h$. Then, for all $\boldsymbol{\tau} \in \boldsymbol{\Sigma}(\mathcal{T},\Omega)$, there exists a unique discrete orthogonal decomposition

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

where $q \in \mathrm{Q}(\mathcal{T},\Omega)$ and $\boldsymbol{v} = \mathbf{div}\,\boldsymbol{\tau}$.

By the approximation property (2.6) of the Arnold–Winther element and the regularity result (2.8), for all $\boldsymbol{v} \in \boldsymbol{L}^2(\Omega,\mathbb{R}^2)/RM$,

$$\|\mathbf{div}^{-1}\boldsymbol{v} - \mathbf{div}_{\mathcal{T}}^{-1}\boldsymbol{v}\|_{0,\Omega} \leq ch\|\mathbf{div}^{-1}\boldsymbol{v}\|_{1,\Omega} \leq ch\|\boldsymbol{v}\|_{0,\Omega}, \tag{2.11}$$

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

2.3. A block diagonal preconditioner for the mixed system. For simplicity, let $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\mathcal{T},\Omega)$ and $\boldsymbol{V} = \boldsymbol{V}(\mathcal{T},\Omega)/RM$. Let $\|\cdot\|_{\boldsymbol{\Sigma}}$ and $\|\cdot\|_{\boldsymbol{V}}$ be the norms on $\boldsymbol{\Sigma}$ and $\boldsymbol{V}$, respectively, i.e., $\|\cdot\|_{\boldsymbol{H}(\mathbf{div},\Omega,\mathbb{S}_2)}$ and $\|\cdot\|_{\boldsymbol{L}^2(\Omega,\mathbb{R}^2)}$. Let $\boldsymbol{\Sigma}^*$ and $\boldsymbol{V}^*$ be the dual spaces of $\boldsymbol{\Sigma}$ and $\boldsymbol{V}$ with dual norms $\|\cdot\|_{\boldsymbol{\Sigma}^*}$ and $\|\cdot\|_{\boldsymbol{V}^*}$ and $<\cdot,\cdot>$ denote the duality pairing. Define the operators

$$\begin{cases} \mathcal{A} : \boldsymbol{\Sigma} \rightarrow \boldsymbol{\Sigma}^*, & <\mathcal{A}\boldsymbol{\sigma},\boldsymbol{\tau}> = (\mathbb{A}\boldsymbol{\sigma},\boldsymbol{\tau}) \quad \text{for all } \boldsymbol{\tau} \in \boldsymbol{\Sigma}, \\ \mathcal{B} : \boldsymbol{\Sigma} \rightarrow \boldsymbol{V}^*, & <\mathcal{B}\boldsymbol{\sigma},\boldsymbol{v}> = (\mathbf{div}\,\boldsymbol{\sigma},\boldsymbol{v}) \quad \text{for all } \boldsymbol{v} \in \boldsymbol{V}. \end{cases}$$

Let $\mathcal{B}^t : \boldsymbol{V} \rightarrow \boldsymbol{\Sigma}^*$ be the adjoint of $\mathcal{B}$. Equation (2.4) can be rewritten as

$$\mathcal{M}\begin{pmatrix} \boldsymbol{\sigma} \\ \boldsymbol{u} \end{pmatrix} = \begin{pmatrix} \mathcal{A} & \mathcal{B}^t \\ \mathcal{B} & 0 \end{pmatrix}\begin{pmatrix} \boldsymbol{\sigma} \\ \boldsymbol{u} \end{pmatrix} = \begin{pmatrix} F \\ G \end{pmatrix}, \tag{2.12}$$

where $F \in \boldsymbol{\Sigma}^*$, $G \in \boldsymbol{V}^*$. The following lemma results from the LBB condition and (2.5). (See [16] for the proof.)

LEMMA 2.1. *The map* $(F,G) \rightarrow (\boldsymbol{\sigma},\boldsymbol{u})$ *defined by solving* (2.12) *with* $F \in \boldsymbol{\Sigma}^*$ *and* $G \in \boldsymbol{V}^*$ *is an isomorphism of* $\boldsymbol{\Sigma}^* \times \boldsymbol{V}^*$ *onto* $\boldsymbol{\Sigma} \times \boldsymbol{V}$ *and so*

$$c_0(\|F\|_{\boldsymbol{\Sigma}^*} + \|G\|_{\boldsymbol{V}^*}) \leq \|\boldsymbol{\sigma}\|_{\boldsymbol{\Sigma}} + \|\boldsymbol{u}\|_{\boldsymbol{V}} \leq c_1(\|F\|_{\boldsymbol{\Sigma}^*} + \|G\|_{\boldsymbol{V}^*}),$$

where c_0 *and* c_1 *are positive and independent of* h.

Our purpose is to find a preconditioner for the operator $\mathcal{M}$. By Lemma 2.1, we only need to find an operator $\mathcal{S} : \boldsymbol{\Sigma}^* \times \boldsymbol{V}^* \rightarrow \boldsymbol{\Sigma} \times \boldsymbol{V}$ such that $\|\mathcal{S}\|_{\mathcal{L}(\boldsymbol{\Sigma}^* \times \boldsymbol{V}^*, \boldsymbol{\Sigma} \times \boldsymbol{V})}$ and $\|\mathcal{S}^{-1}\|_{\mathcal{L}(\boldsymbol{\Sigma} \times \boldsymbol{V}, \boldsymbol{\Sigma}^* \times \boldsymbol{V}^*)}$ are bounded uniformly in h (see [2] for details). Indeed, we can consider an operator in the form $\mathcal{S} = \begin{pmatrix} \mathcal{S}_1 & 0 \\ 0 & \mathcal{S}_2 \end{pmatrix}$, where $\mathcal{S}_1 : \boldsymbol{\Sigma}^* \rightarrow \boldsymbol{\Sigma}$ and $\mathcal{S}_2 : \boldsymbol{V}^* \rightarrow \boldsymbol{V}$

and their inverses are bounded uniformly in h. Consider the following problem: find $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}$ such that

$$\boldsymbol{\Lambda}(\boldsymbol{\sigma}, \boldsymbol{\tau}) = F(\boldsymbol{\tau}) \qquad \text{for all } \ \boldsymbol{\tau} \in \boldsymbol{\Sigma}. \tag{2.13}$$

Clearly a good preconditioner for this problem will yield an ideal $\mathcal{S}_1$. Similarly, an ideal $\mathcal{S}_2$ will come from a good preconditioner for the following problem: find $\boldsymbol{u} \in \boldsymbol{V}$ such that

$$(\boldsymbol{u}, \boldsymbol{v}) = G(\boldsymbol{v}) \qquad \text{for all } \ \boldsymbol{v} \in \boldsymbol{V}. \tag{2.14}$$

The problem (2.14) is easy to solve efficiently. Indeed, we use the basis for $\boldsymbol{V}(\mathcal{T}, \Omega)$ in the implementation. (This, of course, provides a spanning set for $\boldsymbol{V}$.) First, we note that the functional G in original problem (2.12) is usually available as a functional $\widetilde{G}$ defined on $(\boldsymbol{V}(\mathcal{T}, \Omega))^*$ which vanishes on RM. This functional is naturally represented by its action on the basis functions for $\boldsymbol{V}(\mathcal{T}, \Omega)$ and provides the data for the first solve of (2.14). Subsequent solves of (2.14) involve this data plus the result of $\mathcal{B}$ applied to something in $\boldsymbol{\Sigma}$. Thus, at any step of the iteration, (2.14) will have to be solved with a known functional $\widetilde{G}$ on $(\boldsymbol{V}(\mathcal{T}, \Omega))^*$ which vanishes on RM. In this case, the solution of (2.14) coincides with the solution $\boldsymbol{u} \in \boldsymbol{V}(\mathcal{T}, \Omega)$ satisfying

$$(\boldsymbol{u}, \boldsymbol{v}) = \widetilde{G}(\boldsymbol{v}) \qquad \text{for all } \ \boldsymbol{v} \in \boldsymbol{V}(\mathcal{T}, \Omega). \tag{2.15}$$

The space $\boldsymbol{V}(\mathcal{T}, \Omega)$ consists of discontinuous linears on the triangles so the exact solution of (2.15) reduces to the inversion of a block diagonal matrix, with 3×3 diagonal blocks. Hence the problem of defining $\mathcal{S}$ reduces to the problem of constructing $\mathcal{S}_1$. In the remainder of this paper we will focus on constructing a multigrid preconditioner for problem (2.13).

3. The multigrid preconditioner. In this section, we construct and analyze a multigrid preconditioner for problem (2.13). To this end, let $\mathcal{T}_1$ be a unit-sized coarse triangulation of Ω. Subsequently finer triangulations are defined recursively. Given the kth 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. Repeating this process gives 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. We clearly have $h_k = \frac{1}{2} h_{k-1} = O(2^{-k})$. For simplicity of notation, in the rest of this paper, we use $\lesssim$ to denote "less then or equal to" with a factor c independent of k or h_k.

Denote the finite element spaces on the kth level by

$$\mathrm{Q}_k = \mathrm{Q}(\mathcal{T}_k, \Omega), \qquad \boldsymbol{\Sigma}_k = \boldsymbol{\Sigma}(\mathcal{T}_k, \Omega), \qquad \boldsymbol{V}_k = \boldsymbol{V}(\mathcal{T}_k, \Omega)/RM.$$

Notice that we have $\mathrm{Q}_k \subset H_0^2(\Omega)$, $\boldsymbol{\Sigma}_k \subset \boldsymbol{H}_0(\mathbf{div}\,, \Omega, \mathbb{S}_2)$, and $\boldsymbol{V}_k \subset \boldsymbol{L}^2(\Omega, \mathbb{R}^2)$ for each k.

The bilinear form for the biharmonic problem will play an important role in the following analysis. It is defined on $H_0^2(\Omega)$ by

$$\begin{aligned} \mathrm{A}(q, p) &= \int_\Omega \left(\frac{\partial^2 q}{\partial x^2} \frac{\partial^2 p}{\partial x^2} + 2 \frac{\partial^2 q}{\partial x \partial y} \frac{\partial^2 p}{\partial x \partial y} + \frac{\partial^2 q}{\partial y^2} \frac{\partial^2 p}{\partial y^2} \right) d\boldsymbol{x} \\ &= (\mathbf{airy}\, q, \mathbf{airy}\, p). \end{aligned} \tag{3.1}$$

Define operators $\mathrm{A}_k : \mathrm{Q}_k \to \mathrm{Q}_k$ and $\boldsymbol{\Lambda}_k : \boldsymbol{\Sigma}_k \to \boldsymbol{\Sigma}_k$ by

$$(\mathrm{A}_k q, p) = \mathrm{A}(q, p) \qquad \text{for all } q, p \in \mathrm{Q}_k,$$
$$(\boldsymbol{\Lambda}_k \boldsymbol{\sigma}, \boldsymbol{\tau}) = \boldsymbol{\Lambda}(\boldsymbol{\sigma}, \boldsymbol{\tau}) \qquad \text{for all } \boldsymbol{\sigma}, \boldsymbol{\tau} \in \boldsymbol{\Sigma}_k,$$

where the bilinear forms $\mathrm{A}(\cdot,\cdot)$ and $\boldsymbol{\Lambda}(\cdot,\cdot)$ were defined in (3.1) and (2.2), respectively.

The spaces $\{\mathrm{Q}_k\}$ and $\{\boldsymbol{\Sigma}_k\}$ are nonnested since, for example, a function $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_k$ is not necessarily continuous at the midpoints of the edges in the mesh $\mathcal{T}_k$ and a function $q \in \mathrm{Q}_k$ does not necessarily have continuous second order derivatives at the midpoints of the edges in the mesh $\mathcal{T}_k$. Hence we need to define interpolation operators $\mathcal{I}_k : \mathrm{Q}_{k-1} \to \mathrm{Q}_k$ and $\mathbf{I}_k : \boldsymbol{\Sigma}_{k-1} \to \boldsymbol{\Sigma}_k$. The easiest way to do this is by using the "local" nodal value interpolation on each triangle and then taking average on the discontinuous degrees of freedom at vertices.

Denote $\mathcal{N}_k$ to be the set of all nodes in the mesh $\mathcal{T}_k$. For any vertex $v \in \mathcal{N}_k$, let $S_{k-1}(v)$ be the set of all triangles in $\mathcal{T}_{k-1}$ which contain the vertex v and let $|S_{k-1}(v)|$ denote the number of triangles in $S_{k-1}(v)$. For $q \in \mathrm{Q}_{k-1}$ and $\boldsymbol{\tau} \in \boldsymbol{\Sigma}_{k-1}$, define the dofs for $\mathcal{I}_k q$ and $\mathbf{I}_k \boldsymbol{\tau}$ to be identical to those for q and $\boldsymbol{\tau}$ for all dofs excluding the second order derivatives at the vertices for $\mathcal{I}_k q$ and the nodal values at the vertices for $\mathbf{I}_k \boldsymbol{\tau}$. On the excluded dofs we use

$$\mathbf{airy}\,(\mathcal{I}_k q)(v) = \frac{1}{|S_{k-1}(v)|} \sum_{T_v \in S_{k-1}(v)} \mathbf{airy}\, q(v)|_{T_v} \quad \text{for } v \in \mathcal{N}_k,$$
$$\mathbf{I}_k \boldsymbol{\tau}(v) = \frac{1}{|S_{k-1}(v)|} \sum_{T_v \in S_{k-1}(v)} \boldsymbol{\tau}(v)|_{T_v} \qquad \text{for } v \in \mathcal{N}_k.$$

Combining the above gives the definition of $\mathcal{I}_k q$ and $\mathbf{I}_k \boldsymbol{\tau}$ on all dofs. We then have

$$\mathcal{I}_k q = q + \tilde{q} \qquad \text{for all } q \in \mathrm{Q}_{k-1},$$
$$\mathbf{I}_k \boldsymbol{\tau} = \boldsymbol{\tau} + \tilde{\boldsymbol{\tau}} \qquad \text{for all } \boldsymbol{\tau} \in \boldsymbol{\Sigma}_{k-1},$$

where $\tilde{q} \in H_0^2(\Omega)$ and $\tilde{\boldsymbol{\tau}} \in \boldsymbol{H}_0(\mathbf{div}\,, \Omega, \mathbb{S}_2)$ satisfy

$$\begin{aligned} \mathbf{airy}\,\tilde{q}(v)|_T &= \left(\frac{1}{|S_{k-1}(v)|} \sum_{T_v \in S_{k-1}(v)} \mathbf{airy}\, q(v)|_{T_v} \right) - \mathbf{airy}\, q(v)|_T, \\ \tilde{\boldsymbol{\tau}}(v)|_T &= \left(\frac{1}{|S_{k-1}(v)|} \sum_{T_v \in S_{k-1}(v)} \boldsymbol{\tau}(v)|_{T_v} \right) - \boldsymbol{\tau}(v)|_T \end{aligned} \tag{3.2}$$

at each vertex v of any triangle $T \in \mathcal{T}_k$ and vanish at all the other dofs. Define $\mathcal{P}_{k-1} : \mathrm{Q}_k \to \mathrm{Q}_{k-1}$ to be the A-adjoint of $\mathcal{I}_k$ and $\mathbf{P}_{k-1} : \boldsymbol{\Sigma}_k \to \boldsymbol{\Sigma}_{k-1}$ to be the $\boldsymbol{\Lambda}$-adjoint of $\mathbf{I}_k$.

LEMMA 3.1. *We have*

$$\boldsymbol{\Lambda}(\mathbf{I}_k \boldsymbol{\sigma}_{k-1}, \mathbf{I}_k \boldsymbol{\sigma}_{k-1}) \le \omega \boldsymbol{\Lambda}(\boldsymbol{\sigma}_{k-1}, \boldsymbol{\sigma}_{k-1}) \quad \textit{for all } \boldsymbol{\sigma}_{k-1} \in \boldsymbol{\Sigma}_{k-1},$$

where ω is independent of k. Consequently,

$$\boldsymbol{\Lambda}(\mathbf{P}_{k-1} \boldsymbol{\sigma}_k, \mathbf{P}_{k-1} \boldsymbol{\sigma}_k) \le \omega \boldsymbol{\Lambda}(\boldsymbol{\sigma}_k, \boldsymbol{\sigma}_k) \quad \textit{for all } \boldsymbol{\sigma}_k \in \boldsymbol{\Sigma}_k.$$

Proof. The proof follows from a standard scaling argument, the definition of $\mathbf{I}_k$, and the quasi-uniformity of the mesh. ∎

We have the following two lemmas concerning the interpolation operators $\mathcal{I}_k$ and $\mathbf{I}_k$ from [28].

LEMMA 3.2. *Let T be a triangle and v_i, $i = 1, 2, 3$, be its vertices. Let $\boldsymbol{\tau}_i$, $i = 1, 2, 3$, be given constant symmetric matrices. Define $q \in \mathrm{Q}_T$ and $\boldsymbol{\tau} \in \mathbf{\Sigma}_T$ such that*

$$\begin{aligned} \mathbf{airy}\, q(v_i) &= \boldsymbol{\tau}_i \qquad \text{for } i = 1, 2, 3, \\ \boldsymbol{\tau}(v_i) &= \boldsymbol{\tau}_i \qquad \text{for } i = 1, 2, 3, \end{aligned}$$

while vanishing on all the other dofs. Then $\mathbf{airy}\, q = \boldsymbol{\tau}$.

LEMMA 3.3. *The following commutative diagram of exact sequences holds:*

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathrm{Q}_{k-1} & \xrightarrow{\mathbf{airy}} & \mathbf{\Sigma}_{k-1} & \xrightarrow{\mathbf{div}} & \boldsymbol{V}_{k-1}/RM & \rightarrow & 0 \\ & & \downarrow\mathcal{I}_k & & \downarrow\mathbf{I}_k & & \downarrow id & & \\ 0 & \longrightarrow & \mathrm{Q}_k & \xrightarrow{\mathbf{airy}} & \mathbf{\Sigma}_k & \xrightarrow{\mathbf{div}} & \boldsymbol{V}_k/RM & \rightarrow & 0. \end{array} \tag{3.3}$$

It is not our goal to study the general approximation properties of the interpolation operator $\mathbf{I}_k$. Instead, for the multigrid analysis, we require the specific results obtained in the following two lemmas.

LEMMA 3.4. *Let $\boldsymbol{\tau}_{k-1}$ be piecewise linear with respect to $\mathcal{T}_{k-1}$ on all components. Then,*

$$((\mathbf{I} - \mathbf{I}_k)\boldsymbol{\sigma}_{k-1}, \boldsymbol{\tau}_{k-1}) = 0 \quad \textit{for all } \boldsymbol{\sigma}_{k-1} \in \mathbf{\Sigma}_{k-1}.$$

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 $(\mathbf{I} - \mathbf{I}_k)\boldsymbol{\sigma}_{k-1}$ restricted to T is in $\mathbf{\Sigma}(\mathcal{T}_k, T)$ and has nonzero dofs only on the nodal values at v_i, $i = 1, 2, 3$. On each of the four finer triangles T_i, $i = 1, \ldots, 4$, making up T, we have

$$(\mathbf{I} - \mathbf{I}_k)\boldsymbol{\sigma}_{k-1} = \mathbf{airy}\, q_i$$

for q_i as defined in Lemma 3.2. By construction, these q_i share the same nodal values at v_j, $j = 1, 2, 3$, and thus the function q whose restriction is q_i on T_i is in $Q(\mathcal{T}_k, T) \subset C^1(T)$. Now, since $\boldsymbol{\sigma}_{k-1}$ has continuous normal components, we have

$$(\mathbf{I} - \mathbf{I}_k)\boldsymbol{\sigma}_{k-1}\, \boldsymbol{n}|_{\partial T} = \mathbf{airy}\, q\, \boldsymbol{n}|_{\partial T} = \mathbf{0},$$

i.e., $\frac{\partial^2 q}{\partial \boldsymbol{n} \partial \boldsymbol{s}} = \frac{\partial^2 q}{\partial \boldsymbol{s}^2} = 0$, 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 are continuous across ∂T_i. Thus, integration by parts gives that for any linear function f on T,

$$((\mathbf{I} - \mathbf{I}_k)\boldsymbol{\sigma}_{k-1}, f)_{L^2(T)} = (\mathbf{airy}\, q, f)_{L^2(T)} = 0.$$

This completes the proof of the lemma. ∎

LEMMA 3.5. *There exists a positive constant c such that for all $\boldsymbol{v}_{k-1} \in \boldsymbol{V}_{k-1}$,*

$$\|(\mathbf{I} - \mathbf{I}_k)\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_{k-1}\|_{0,\Omega} \le c h_k \|\boldsymbol{v}_{k-1}\|_{0,\Omega}.$$

Here $\mathbf{div}_k^{-1} = \mathbf{div}_{\mathcal{T}_k}^{-1}$ as defined by (2.10).

Proof. Notice that $(\mathbf{I} - \mathbf{I}_k)\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_{k-1}$ is divergence free by Lemma 3.2. 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 3.4, for any $\boldsymbol{\tau}_{k-1} \in \boldsymbol{\Sigma}_{k-1}$ which is continuous and piecewise linear with respect to $\mathcal{T}_{k-1}$,

$$((\mathbf{I}-\mathbf{I}_k)\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_{k-1}, \boldsymbol{\tau}_{k-1}) = 0.$$

Let $\boldsymbol{\tau}_{k-1}$ be the 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 result (2.8) and the approximation result (2.11),

$$\begin{aligned}\|(\mathbf{I}-\mathbf{I}_k)\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_{k-1}\|_{0,\Omega}^2 &= ((\mathbf{I}-\mathbf{I}_k)\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_{k-1}, \mathbf{div}_{k-1}^{-1}\boldsymbol{v}_{k-1} - \mathbf{div}^{-1}\boldsymbol{v}_{k-1}) \\ &\quad - ((\mathbf{I}-\mathbf{I}_k)\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_{k-1}, \mathbf{I}_k(\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_{k-1} - \boldsymbol{\tau}_{k-1})) \\ &\lesssim \|(\mathbf{I}-\mathbf{I}_k)\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_{k-1}\|_{0,\Omega}(h_k\|\boldsymbol{v}_{k-1}\|_{0,\Omega} + \|\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_{k-1} - \boldsymbol{\tau}_{k-1}\|_{0,\Omega}).\end{aligned}$$

Thus,

$$\begin{aligned}\|(\mathbf{I}-\mathbf{I}_k)\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_{k-1}\|_{0,\Omega} &\lesssim h_k\|\boldsymbol{v}_{k-1}\|_{0,\Omega} + \|\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_{k-1} - \mathbf{div}^{-1}\boldsymbol{v}_{k-1}\|_{0,\Omega} \\ &\quad + \|\mathbf{div}^{-1}\boldsymbol{v}_{k-1} - \boldsymbol{\tau}_{k-1}\|_{0,\Omega} \\ &\lesssim h_k\|\boldsymbol{v}_{k-1}\|_{0,\Omega}.\end{aligned}$$

This completes the proof of the lemma. ◻

Now we state the variable V-cycle multigrid preconditioner. Let $\mathbf{R}_k : \boldsymbol{\Sigma}_k \to \boldsymbol{\Sigma}_k$ be a symmetric and positive definite linear operator which we call a smoother. A construction for $\mathbf{R}_k$ will be given in the next section. Let m_k, the number of smoothing steps on the kth level, satisfy

$$\beta_0 m_k \le m_{k-1} \le \beta_1 m_k, \qquad \text{where } 1 < \beta_0 \le \beta_1.$$

The choice of $\beta_0 = \beta_1 = 2$ is typical. Denote $\mathbf{I}_k^t : \boldsymbol{\Sigma}_k \to \boldsymbol{\Sigma}_{k-1}$ to be the $\boldsymbol{L}^2$-adjoint of $\mathbf{I}_k$, i.e.,

$$(\mathbf{I}_k^t\boldsymbol{\sigma}_k, \boldsymbol{\tau}_{k-1}) = (\boldsymbol{\sigma}_k, \mathbf{I}_k\boldsymbol{\tau}_{k-1}) \quad \text{for all } \boldsymbol{\tau}_{k-1} \in \boldsymbol{\Sigma}_{k-1}.$$

The variable V-cycle multigrid preconditioner $\mathbf{B}_k : \boldsymbol{\Sigma}_k \to \boldsymbol{\Sigma}_k$ is defined inductively as follows.

ALGORITHM 1. *Set* $\mathbf{B}_1 = \boldsymbol{\Lambda}_1^{-1}$. *Assuming that* $\mathbf{B}_{k-1} : \boldsymbol{\Sigma}_{k-1} \to \boldsymbol{\Sigma}_{k-1}$ *has been 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* $l = 1, \dots, m_k$;
(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* $l = m_k + 1, \dots, 2m_k$;

Set $\mathbf{B}_k\boldsymbol{g} = \boldsymbol{\sigma}^{2m_k}$.

Remark 1. It appears that one needs to solve linear systems involving the mass matrix for the computation of $\mathbf{I}_k^t$ and $\boldsymbol{\Lambda}_k$ in the above algorithm. This Gram matrix inversion is avoided in the implementation because of the judicious choice of $\mathbf{R}_k$. For these and other implementation issues, see [15, 28].

The following theorem and its proof is a straightforward variation of Theorem 7.4 in [15].

THEOREM 3.6. *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* $0 < \alpha \le 1$ *and a constant* C_p *independent of* k *such that for all* $\boldsymbol{\tau} \in \boldsymbol{\Sigma}_k$,

$$|\boldsymbol{\Lambda}((\mathbf{I} - \mathbf{I}_k\mathbf{P}_{k-1})\boldsymbol{\tau}, \boldsymbol{\tau})| \le C_p^{2\alpha}(\mathbf{R}_k\boldsymbol{\Lambda}_k\boldsymbol{\tau}, \boldsymbol{\Lambda}_k\boldsymbol{\tau})^\alpha \boldsymbol{\Lambda}(\boldsymbol{\tau}, \boldsymbol{\tau})^{1-\alpha}.$$

Then, the preconditioner $\mathbf{B}_k$ *is symmetric and positive definite. Furthermore,* $\mathbf{B}_k$ *satisfies*

$$\left(\frac{m_k^\alpha}{M+m_k^\alpha}\right)\mathbf{\Lambda}(\boldsymbol{\tau},\boldsymbol{\tau}) \le \mathbf{\Lambda}(\mathbf{B}_k\mathbf{\Lambda}_k\boldsymbol{\tau},\boldsymbol{\tau}) \le \left(\frac{M+m_k^\alpha}{m_k^\alpha}\right)\mathbf{\Lambda}(\boldsymbol{\tau},\boldsymbol{\tau}) \quad \text{for all } \boldsymbol{\tau}\in\mathbf{\Sigma}_k,$$

where M *is a sufficiently large positive constant depending only on* C_p *and* α.

In the next section, we will construct an additive smoother and prove it satisfies assumptions (M.1) and (M.2).

4. An additive Schwarz smoother. Recall that $\mathcal{N}_k$ denotes the set of all nodes in the triangulation $\mathcal{T}_k$ (including the boundary nodes) and $S_k(v)$ denotes the set of triangles in $\mathcal{T}_k$ meeting at the vertex v for each $v \in \mathcal{N}_k$. The (interior of the) union of all triangles in $S_k(v)$ forms a subdomain 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 subdomains in $\{\Omega_{k,v}\}_{v\in\mathcal{N}_k}$.

Let $\mathrm{Q}_{k,v}$ and $\mathbf{\Sigma}_{k,v}$ be the subspace of functions in Q_k and $\mathbf{\Sigma}_k$, respectively, which have support contained in $\overline{\Omega}_{k,v}$. It is easy to see that the span of $\{\mathrm{Q}_{k,v}\}$ (respectively, $\mathbf{\Sigma}_{k,v}$) is all of Q_k (respectively, $\mathbf{\Sigma}_k$). Let $\mathcal{P}_{k,v} : \mathrm{Q}_k \to \mathrm{Q}_{k,v}$ be the A-projection, $\mathbf{P}_{k,v} : \mathbf{\Sigma}_k \to \mathbf{\Sigma}_{k,v}$ be the $\mathbf{\Lambda}$-projection, and $\mathcal{I}^t_{k,v} : \mathrm{Q}_k \to \mathrm{Q}_{k,v}$, $\mathbf{I}^t_{k,v} : \mathbf{\Sigma}_k \to \mathbf{\Sigma}_{k,v}$ be the $\boldsymbol{L}^2$-projections. Define $\mathrm{A}_{k,v} : \mathrm{Q}_{k,v} \to \mathrm{Q}_{k,v}$ and $\mathbf{\Lambda}_{k,v} : \mathbf{\Sigma}_{k,v} \to \mathbf{\Sigma}_{k,v}$ by

$$\begin{aligned} (\mathrm{A}_{k,v}p, q) &= \mathrm{A}(p,q) && \text{for all } p,q \in \mathrm{Q}_{k,v},\\ (\mathbf{\Lambda}_{k,v}\boldsymbol{\sigma},\boldsymbol{\tau}) &= \mathbf{\Lambda}(\boldsymbol{\sigma},\boldsymbol{\tau}) && \text{for all } \boldsymbol{\sigma},\boldsymbol{\tau} \in \mathbf{\Sigma}_{k,v}. \end{aligned}$$

Clearly, we have $\mathrm{A}_{k,v}\mathcal{P}_{k,v} = \mathcal{I}^t_{k,v}\mathrm{A}_k$ and $\mathbf{\Lambda}_{k,v}\mathbf{P}_{k,v} = \mathbf{I}^t_{k,v}\mathbf{\Lambda}_k$. Define

$$\begin{aligned} \mathcal{R}_k &= \rho \sum_{v\in\mathcal{N}_k} \mathcal{P}_{k,v}\mathrm{A}_k^{-1} = \rho \sum_{v\in\mathcal{N}_k} \mathrm{A}_{k,v}^{-1}\mathcal{I}^t_{k,v},\\ \mathbf{R}_k &= \rho \sum_{v\in\mathcal{N}_k} \mathbf{P}_{k,v}\mathbf{\Lambda}_k^{-1} = \rho \sum_{v\in\mathcal{N}_k} \mathbf{\Lambda}_{k,v}^{-1}\mathbf{I}^t_{k,v}, \end{aligned} \tag{4.1}$$

where $\rho > 0$ is a scaling factor which will only depend on the finite overlapping constant, e.g., $\rho = 1/3$. It is well known (see [30]) that since $\{\mathbf{\Sigma}_{k,v}\}$ spans $\mathbf{\Sigma}_k$, $\mathbf{R}_k$ is invertible and satisfies

$$(\mathbf{R}_k^{-1}\boldsymbol{\tau},\boldsymbol{\tau}) = \rho^{-1} \inf_{\substack{\boldsymbol{\tau}_v\in\mathbf{\Sigma}_{k,v}\\ \sum_v \boldsymbol{\tau}_v=\boldsymbol{\tau}}} \sum_{v\in\mathcal{N}_k} \mathbf{\Lambda}(\boldsymbol{\tau}_v,\boldsymbol{\tau}_v) \qquad \text{for all } \boldsymbol{\tau}\in\mathbf{\Sigma}_k. \tag{4.2}$$

Also, we note that $\mathcal{R}_k$ is defined purely for theoretical analysis and only $\mathbf{R}_k$ appears in the implementation. The implementation of $\mathbf{R}_k$ involves solving local problems on each $\Omega_{k,v}$.

Remark 2. 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.

In the remainder of this section, we prove that the smoother $\mathbf{R}_k$ satisfies assumptions (M.1) and (M.2). These results are gathered in the next two lemmas.

LEMMA 4.1. *For* $\rho \le 1/3$, *the smoother* $\mathbf{R}_k$ *satisfies assumption* (M.1).

Proof. The proof follows from the Cauchy–Schwarz inequality and the finite overlapping condition (see, e.g., [14]). ∎

LEMMA 4.2. *The smoother* $\mathbf{R}_k$ *satisfies assumption* (M.2).

Proof. As shown in section 2, 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 $\boldsymbol{v}_k = \mathbf{div}\,\boldsymbol{\sigma}_k \in \boldsymbol{V}_k/RM$. By Lemma 3.3,

$$(\mathbf{I} - \mathbf{I}_k\mathbf{P}_{k-1})\boldsymbol{\sigma}_k = \sum_{i=1}^{4} \boldsymbol{\sigma}_k^i,$$

where

$$\begin{aligned}
\boldsymbol{\sigma}_k^1 &= \mathbf{airy}\,(q_k - \mathcal{I}_k\mathcal{P}_{k-1}q_k),\\
\boldsymbol{\sigma}_k^2 &= \mathbf{I}_k(\mathbf{airy}\,\mathcal{P}_{k-1}q_k - \mathbf{P}_{k-1}\mathbf{airy}\,q_k),\\
\boldsymbol{\sigma}_k^3 &= \mathbf{div}_k^{-1}\boldsymbol{v}_k - \mathbf{I}_k\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_k,\\
\boldsymbol{\sigma}_k^4 &= \mathbf{I}_k(\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_k - \mathbf{P}_{k-1}\mathbf{div}_k^{-1}\boldsymbol{v}_k).
\end{aligned}$$

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. Thus

$$\begin{aligned}
|\boldsymbol{\Lambda}((\mathbf{I} - \mathbf{I}_k\mathbf{P}_{k-1})\boldsymbol{\sigma}_k, \boldsymbol{\sigma}_k)| &= |\boldsymbol{\Lambda}(\boldsymbol{\sigma}_k^1 + \boldsymbol{\sigma}_k^2 + \boldsymbol{\sigma}_k^3 + \boldsymbol{\sigma}_k^4, \boldsymbol{\sigma}_k)|\\
&\lesssim |\boldsymbol{\Lambda}(\boldsymbol{\sigma}_k^1, \mathbf{airy}\,q_k)| + \sum_{i=2}^{4}(\mathbf{R}_k^{-1}\boldsymbol{\sigma}_k^i, \boldsymbol{\sigma}_k^i)^{1/2}(\mathbf{R}_k\boldsymbol{\Lambda}_k\boldsymbol{\sigma}_k, \boldsymbol{\Lambda}_k\boldsymbol{\sigma}_k)^{1/2}.
\end{aligned} \tag{4.3}$$

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 $(\mathbf{R}_k\boldsymbol{\Lambda}_k\boldsymbol{\sigma}_k, \boldsymbol{\Lambda}_k\boldsymbol{\sigma}_k) \le \boldsymbol{\Lambda}(\boldsymbol{\sigma}_k, \boldsymbol{\sigma}_k)$, assumption (M.2) with $\alpha = 1/4$ will follow from (4.3), (I), and (II).

To prove (I), first notice that for the biharmonic problem, we have (see [15])

$$\frac{1}{\tilde{\lambda}_k}\|\mathrm{A}_k q_k\|_{0,\Omega}^2 \lesssim (\mathcal{R}_k\mathrm{A}_k q_k, \mathrm{A}_k q_k) \quad \text{for all } q_k \in \mathrm{Q}_k,$$

where $\tilde{\lambda}_k = O(h_k^{-4})$ is the largest eigenvalue of the operator A_k.

Theorem 14.1 in [15] states that if Ω is a convex polygon, then

$$\mathrm{A}((\mathrm{I} - \mathcal{I}_k\mathcal{P}_{k-1})q_k, q_k) \lesssim (\mathrm{A}_k q_k, q_k)^{3/4}\left(\frac{\|\mathrm{A}_k q_k\|_{0,\Omega}^2}{\tilde{\lambda}_k}\right)^{1/4}.$$

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

$$(\mathcal{R}_k\mathrm{A}_k q_k, \mathrm{A}_k q_k) \le (\mathbf{R}_k\boldsymbol{\Lambda}_k\boldsymbol{\sigma}_k, \boldsymbol{\Lambda}_k\boldsymbol{\sigma}_k). \tag{4.4}$$

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 (4.4) will follow if for each $v \in \mathcal{N}_k$,

$$\text{(4.5)} \quad \mathrm{A}(\mathcal{P}_{k,v}q_k, \mathcal{P}_{k,v}q_k) = \boldsymbol{\Lambda}(\mathbf{airy}\,(\mathcal{P}_{k,v}q_k), \mathbf{airy}\,(\mathcal{P}_{k,v}q_k)) \le \boldsymbol{\Lambda}(\mathbf{P}_{k,v}\boldsymbol{\sigma}_k, \mathbf{P}_{k,v}\boldsymbol{\sigma}_k).$$

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, (4.5) follows. This completes the proof of (I).

Next, we prove (II). For each $v \in \mathcal{N}_k$ let θ_v denote the piecewise continuous linear basis function associated with v. Clearly, $\sum_v \theta_v$ gives a partition of unity on Ω which satisfies

$$\begin{aligned}&(1)\ \theta_v|_T \in P_1(T) \text{ for any } T \in \mathcal{T}_k; \qquad (2)\ \mathrm{supp}(\theta_v) \subset \overline{\Omega}_{k,v};\\ &(3)\ |\theta_v|_{W^{j,\infty}(\Omega)} \lesssim h_k^{-j}, \quad j = 0, 1.\end{aligned}$$

Let Π_k denote the natural interpolation operator onto $\boldsymbol{\Sigma}_k$ associated with the dofs. 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 [6], a simple scaling argument shows that

$$\text{(4.6)} \qquad \|\Pi_k(\theta_v\boldsymbol{\tau})\|_{0,\Omega} \lesssim \|\theta_v\boldsymbol{\tau}\|_{0,\Omega}.$$

Also, it has been shown in [6] that $\mathbf{div}\,\Pi_k = \mathbf{P}_{\boldsymbol{V}_k}\mathbf{div}$, where $\mathbf{P}_{\boldsymbol{V}_k}$ is the $\boldsymbol{L}^2$ projection onto $\boldsymbol{V}_k$. Therefore

$$\|\mathbf{div}\,\Pi_k(\theta_v\boldsymbol{\tau})\|_{0,\Omega} = \|\mathbf{P}_{\boldsymbol{V}_k}\mathbf{div}\,(\theta_v\boldsymbol{\tau})\|_{0,\Omega} \le \|\mathbf{div}\,(\theta_v\boldsymbol{\tau})\|_{0,\Omega}.$$

By (4.2), (4.6), an inverse inequality, and the properties of θ_v, for $i = 2, 3, 4$,

$$\begin{aligned}(\mathbf{R}_k^{-1}\boldsymbol{\sigma}_k^i, \boldsymbol{\sigma}_k^i) &\le \rho^{-1} \sum_{v \in \mathcal{N}_k} (\|\Pi_k(\theta_v\boldsymbol{\sigma}_k^i)\|_{0,\Omega_{k,v}}^2 + \|\mathbf{div}\,\Pi_k(\theta_v\boldsymbol{\sigma}_k^i)\|_{0,\Omega_{k,v}}^2)\\ &\lesssim \sum_{v \in \mathcal{N}_k} (\|\theta_v\boldsymbol{\sigma}_k^i\|_{0,\Omega_{k,v}}^2 + \|\mathbf{div}\,(\theta_v\boldsymbol{\sigma}_k^i)\|_{0,\Omega_{k,v}}^2)\\ &\lesssim h_k^{-2}\|\boldsymbol{\sigma}_k^i\|_{0,\Omega}^2 + \|\mathbf{div}\,\boldsymbol{\sigma}_k^i\|_{0,\Omega}^2.\end{aligned}$$

Hence the proof for (II) reduces to proving for $i = 2, 3, 4$ that

$$\text{(4.7)} \qquad \begin{aligned}\|\boldsymbol{\sigma}_k^i\|_{0,\Omega} &\lesssim h_k\|\boldsymbol{\sigma}_k\|_{\boldsymbol{H}(\mathbf{div}\,,\Omega,\mathbb{S}_2)},\\ \|\mathbf{div}\,\boldsymbol{\sigma}_k^i\|_{0,\Omega} &\lesssim \|\boldsymbol{\sigma}_k\|_{\boldsymbol{H}(\mathbf{div}\,,\Omega,\mathbb{S}_2)}.\end{aligned}$$

For $\boldsymbol{\sigma}_k^2$ and any $\boldsymbol{\tau}_{k-1} = \mathbf{airy}\,p_{k-1} + \mathbf{div}_{k-1}^{-1}\boldsymbol{w}_{k-1} \in \boldsymbol{\Sigma}_{k-1}$,

$$\begin{aligned}&|\boldsymbol{\Lambda}(\mathbf{airy}\,\mathcal{P}_{k-1}q_k - \mathbf{P}_{k-1}\mathbf{airy}\,q_k, \boldsymbol{\tau}_{k-1})|\\ &\qquad = |(\mathbf{airy}\,\mathcal{P}_{k-1}q_k, \mathbf{airy}\,p_{k-1}) - (\mathbf{airy}\,q_k, \mathbf{I}_k\boldsymbol{\tau}_{k-1})|.\end{aligned}$$

Now

$$(\mathbf{airy}\,\mathcal{P}_{k-1}q_k, \mathbf{airy}\,p_{k-1}) = (\mathbf{airy}\,q_k, \mathbf{I}_k\mathbf{airy}\,p_{k-1})$$

so

$$
\begin{aligned}
|\mathbf{\Lambda}(\mathbf{airy}\,\mathcal{P}_{k-1}q_k - \mathbf{P}_{k-1}\mathbf{airy}\,q_k, \boldsymbol{\tau}_{k-1})| &= |(\mathbf{airy}\,q_k, \mathbf{I}_k\mathbf{div}_{k-1}^{-1}\boldsymbol{w}_{k-1})| \\
&\leq |(\mathbf{airy}\,q_k, (\mathbf{I}_k - \mathbf{I})\mathbf{div}_{k-1}^{-1}\boldsymbol{w}_{k-1})| + |(\mathbf{airy}\,q_k, \mathbf{div}_{k-1}^{-1}\boldsymbol{w}_{k-1} - \mathbf{div}^{-1}\boldsymbol{w}_{k-1})| \\
&\lesssim h_k\|\boldsymbol{\sigma}_k\|_{\boldsymbol{H}(\mathbf{div},\Omega,\mathbb{S}_2)}\|\boldsymbol{\tau}_{k-1}\|_{\boldsymbol{H}(\mathbf{div},\Omega,\mathbb{S}_2)}.
\end{aligned}
$$

We used the Cauchy–Schwarz inequality, (2.11), and Lemma 3.5 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 3.1, 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 (4.7).

Next, we consider $\boldsymbol{\sigma}_k^3$. Define $\mathbf{P}_{\boldsymbol{V}_{k-1}}$ to be the $\boldsymbol{L}^2$ projection onto $\boldsymbol{V}_{k-1}/RM$. Then

$$
\|\mathbf{div}\,\boldsymbol{\sigma}_k^3\|_{0,\Omega} = \|\boldsymbol{v}_k - \mathbf{P}_{\boldsymbol{V}_{k-1}}\boldsymbol{v}_k\|_{0,\Omega} \leq \|\boldsymbol{v}_k\|_{0,\Omega} \lesssim \|\boldsymbol{\sigma}_k\|_{\boldsymbol{H}(\mathbf{div},\Omega,\mathbb{S}_2)}
$$

and by (2.11), Lemma 3.5, and the fact that $h_{k-1} = 2h_k$,

$$
\begin{aligned}
\|\boldsymbol{\sigma}_k^3\|_{0,\Omega} &\lesssim \|\mathbf{div}_k^{-1}\boldsymbol{v}_k - \mathbf{div}^{-1}\boldsymbol{v}_k\|_{0,\Omega} + \|\mathbf{div}^{-1}\boldsymbol{v}_k - \mathbf{div}_{k-1}^{-1}\boldsymbol{v}_k\|_{0,\Omega} \\
&\quad + \|(\mathbf{I} - \mathbf{I}_k)\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_k\|_{0,\Omega} \\
&\lesssim h_k\|\boldsymbol{v}_k\|_{0,\Omega} \lesssim h_k\|\boldsymbol{\sigma}_k\|_{\boldsymbol{H}(\mathbf{div},\Omega,\mathbb{S}_2)}.
\end{aligned}
$$

Hence $\boldsymbol{\sigma}_k^3$ satisfies (4.7).

For $\boldsymbol{\sigma}_k^4$, let $\boldsymbol{\tau}_{k-1} \in \boldsymbol{\Sigma}_{k-1}$ be arbitrary. Then

(4.8)
$$
\begin{aligned}
|\mathbf{\Lambda}(\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_k - \mathbf{P}_{k-1}\mathbf{div}_k^{-1}\boldsymbol{v}_k, \boldsymbol{\tau}_{k-1})| &= |\mathbf{\Lambda}(\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_k, \boldsymbol{\tau}_{k-1}) - \mathbf{\Lambda}(\mathbf{div}_k^{-1}\boldsymbol{v}_k, \mathbf{I}_k\boldsymbol{\tau}_{k-1})| \\
&= |(\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_k, \boldsymbol{\tau}_{k-1}) - (\mathbf{div}_k^{-1}\boldsymbol{v}_k, \mathbf{I}_k\boldsymbol{\tau}_{k-1}) + (\mathbf{P}_{\boldsymbol{V}_{k-1}}\boldsymbol{v}_k - \boldsymbol{v}_k, \mathbf{div}\,\tau_{k-1})| \\
&= |(\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_k, \boldsymbol{\tau}_{k-1}) - (\mathbf{div}_k^{-1}\boldsymbol{v}_k, \mathbf{I}_k\boldsymbol{\tau}_{k-1})|.
\end{aligned}
$$

Since $(\mathbf{div}^{-1}\boldsymbol{v}_k, (\mathbf{I} - \mathbf{I}_k)\boldsymbol{\tau}_{k-1})$ is zero, by (4.8), (2.11), and Lemma 3.1, we have

$$
\begin{aligned}
|\mathbf{\Lambda}(\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_k - \mathbf{P}_{k-1}\mathbf{div}_k^{-1}\boldsymbol{v}_k, \boldsymbol{\tau}_{k-1})| &= |(\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_k - \mathbf{div}^{-1}\boldsymbol{v}_k, \boldsymbol{\tau}_{k-1}) \\
&\quad + (\mathbf{div}^{-1}\boldsymbol{v}_k - \mathbf{div}_k^{-1}\boldsymbol{v}_k, \mathbf{I}_k\boldsymbol{\tau}_{k-1})| \\
&\lesssim h_k\|\boldsymbol{\sigma}_k\|_{\boldsymbol{H}(\mathbf{div},\Omega,\mathbb{S}_2)}\|\boldsymbol{\tau}_{k-1}\|_{\boldsymbol{H}(\mathbf{div},\Omega,\mathbb{S}_2)}.
\end{aligned}
$$

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 3.1 gives

$$
\begin{aligned}
\|\boldsymbol{\sigma}_k^4\|_{\boldsymbol{H}(\mathbf{div},\Omega,\mathbb{S}_2)} &\lesssim \|\mathbf{div}_{k-1}^{-1}\boldsymbol{v}_k - \mathbf{P}_{k-1}\mathbf{div}_k^{-1}\boldsymbol{v}_k\|_{\boldsymbol{H}(\mathbf{div},\Omega,\mathbb{S}_2)} \\
&\lesssim h_k\|\boldsymbol{\sigma}_k\|_{\boldsymbol{H}(\mathbf{div},\Omega,\mathbb{S}_2)}.
\end{aligned}
$$

Therefore, $\boldsymbol{\sigma}_k^4$ satisfies (4.7).

Combining all the above shows that $\mathbf{R}_k$ satisfies assumption (M.2) with a constant C_p independent of k. ∎

TABLE 5.1
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	dofs	cond($\mathbf{\Lambda}_k$)	cond($\mathbf{B}_k^R\mathbf{\Lambda}_k$)	cond($\mathbf{B}_k\mathbf{\Lambda}_k$)	cond($\mathbf{B}_k^m\mathbf{\Lambda}_k$)
2	115	1.58e+04	6.37e+03	3.43	2.66
3	395	7.19e+04	3.90e+04	4.09	3.15
4	1459	2.97e+05	1.67e+05	4.23	3.41
5	5603	1.20e+06	6.82e+05	4.24	3.53

TABLE 5.2
Condition number estimates for $\mathbf{B}_k^V\mathbf{\Lambda}_k$.

level	2	3	4	5
cond($\mathbf{B}_k^V\mathbf{\Lambda}_k$)	3.43	4.03	4.20	4.22

5. Numerical results. We report some numerical results for the multigrid preconditioners for the $\boldsymbol{H}(\mathbf{div})$ problem (2.13). Let Ω be the unit square $(0,1)\times(0,1)$. We solve problem (2.13) by the preconditioned conjugate gradient method (PCG). The right-hand side is selected randomly.

Three different multigrid preconditioners are considered. For variable V-cycle preconditioners, we use $\beta_0 = \beta_1 = 2$ and one smoothing on the finest grid. First, we consider the variable V-cycle multigrid preconditioner with Richardson smoother (denoted by $\mathbf{B}_k^R$). Secondly, we experiment on the variable V-cycle multigrid preconditioner $\mathbf{B}_k$ with the additive Schwarz smoother built on the vertex-based subspaces, as defined in section 4. The scaling factor ρ in (4.1) is set to be $\frac{1}{3}$. Finally, we consider the variable V-cycle multigrid preconditioner $\mathbf{B}_k^m$ using the multiplicative Schwarz smoother as discussed in Remark 2. For all three preconditioners, we set the first level mesh by bisecting Ω using its negatively sloped diagonal.

Experiments show that $\mathbf{B}_k^R$ does not work well, as shown in Table 5.1. We report the condition number estimates for $\mathbf{B}_k\mathbf{\Lambda}_k$ in Table 5.1, together with the condition number estimates for $\mathbf{B}_k^m\mathbf{\Lambda}_k$. Both appear to be bounded independently of k. These results also 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 for many other applications.

Further experiments also suggest that the V-cycle multigrid preconditioner $\mathbf{B}_k^V$ with the additive Schwarz smoother as in $\mathbf{B}_k$ and one smoothing on each level is also optimal for this test problem (see Table 5.2). We are unable to explain this theoretically.

REFERENCES

[1] D. N. ARNOLD, F. BREZZI, AND J. J. DOUGLAS, *Peers: A new mixed finite element for plane elasticity*, Japan J. Appl. Math., 1 (1984), pp. 347–367.
[2] D. N. ARNOLD, R. S. FALK, AND R. WINTHER, *Preconditioning in H(div) and applications*, Math. Comp., 66 (1997), pp. 957–984.
[3] D. N. ARNOLD, R. S. FALK, AND R. WINTHER, *Multigrid in H(div) and H(curl)*, Numer. Math., 85 (2000), pp. 197–217.
[4] D. N. ARNOLD AND R. S. FALK, *A new mixed formulation for elasticity*, Numer. Math., 53 (1988), pp. 13–30.
[5] 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 (1984), pp. 1–22.
[6] D. N. ARNOLD AND R. WINTHER, *Mixed finite element for elasticity*, Numer. Math., 92 (2002), pp. 401–419.

[7] D. N. Arnold, *Mixed finite element methods for elliptic problems*, Comput. Methods Appl. Mech. Engrg., 82 (1990), pp. 281–300.
[8] T. M. Austin, T. A. Manteuffel, and S. McCormick, *A robust multilevel approach for minimizing H(div)-dominated functionals in an H1-conforming finite element space*, Numer. Linear Algebra Apps., 11 (2004), pp. 115–140.
[9] R. Bank, B. Welfert, and H. Yserentant, *A preconditioning technique for indefinite systems resulting from mixed approximation of elliptic problems*, Numer. Math., 56 (1990), pp. 645–666.
[10] J. H. Bramble, R. D. Lazarov, and J. E. Pasciak, *Least-squares methods for linear elasticity based on a discrete negative norm*, Comput. Methods Appl. Mech. Engrg., 152 (2001), pp. 520–543.
[11] 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 (1997) pp. 1072–1092.
[12] J. H. Bramble and J. E. Pasciak, *A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems*, Math. Comp., 50 (1988), pp. 1–17.
[13] J. H. Bramble and J. E. Pasciak, *A domain decomposition technique for Stokes problems*, Appl. Numer. Math., 6 (1990), pp. 251–261.
[14] J. H. Bramble and J. E. Pasciak, *The analysis of smoothers for multigrid algorithms*, Math. Comp., 58 (1992), pp. 467–488.
[15] J. Bramble and X. Zhang, *Handbook of numerical analysis*, in *Handbook of Numerical Analysis*, vol. VII, P. Ciarlet and J.-L. Lions, eds., North–Holland, Amsterdam, 2000, pp. 173–415.
[16] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, New York, 1991.
[17] Z. Cai and G. Starke, *First-order system least squares for the stress-displacement formulation: linear elasticity*, SIAM J. Numer. Anal., 41 (2003), pp. 715–730.
[18] Z. Cai and G. Starke, *Least-squares methods for linear elasticity*, SIAM J. Numer. Anal., To appear.
[19] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North–Holland, Amsterdam, 1978.
[20] H. C. Elman and G. H. Golub, *Inexact and preconditioned Uzawa algorithms for saddle point problems*, SIAM J. Numer. Anal., 31 (1994), pp. 1645–1661.
[21] R. E. Ewing and J. Wang, *Analysis of the Schwarz algorithm for mixed finite element methods*, RAIRO Math. Model. Anal. Numer., 26 (1992), pp. 739–756.
[22] 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 (1988), pp. 123–141.
[23] P. Grisvard, *Singularities in Boundary Value Problems*, Research Notes in Appl. Math. 22, Springer-Verlag, New York, 1992.
[24] W. Hackbusch, *Multi-Grid Methods and Applications*, Springer-Verlag, Berlin, 1985.
[25] R. Hiptmair, *Multigrid method for H(div) in three dimensions*, Electron. Trans. Numer. Anal., 6 (1997), pp. 133–152.
[26] C. Johnson and B. Mercier, *Some equilbrium finite element methods for two-dimensional elasticity problems*, Numer. Math., 30 (1978), pp. 103–116.
[27] T. Rusten and R. Winther, *A preconditioned iterative method for saddlepoint problem*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 887–904.
[28] Y. Wang, *Overlapping Schwarz preconditioner for the mixed formulation of plane elasticity*, Applied. Numer. Math., 54 (2005), pp. 292–309.
[29] J. Xu, *Iterative methods by space decomposition and subspace correction*, SIAM Rev., 34 (1992), pp. 581–613.
[30] X. Zhang, *Multilevel Schwarz methods*, Numer. Math., 63 (1992), pp. 521–539.