MATHEMATICS OF COMPUTATION
http://dx.doi.org/10.1090/mcom/3285
Article electronically published on November 9, 2017

FAST AUXILIARY SPACE PRECONDITIONERS FOR LINEAR ELASTICITY IN MIXED FORM

LONG CHEN, JUN HU, AND XUEHAI HUANG

Abstract. A block-diagonal preconditioner with the minimal residual method and an approximate block-factorization preconditioner with the generalized minimal residual method are developed for Hu-Zhang mixed finite element methods for linear elasticity. They are based on a new stability result for the saddle point system in mesh-dependent norms. The mesh-dependent norm for the stress corresponds to the mass matrix which is easy to invert while for the displacement it is spectral equivalent to the Schur complement. A fast auxiliary space preconditioner based on the H^1-conforming linear element of the linear elasticity problem is then designed for solving the Schur complement. For both diagonal and triangular preconditioners, it is proved that the conditioning numbers of the preconditioned systems are bounded above by a constant independent of both the crucial Lamé constant and the mesh size. Numerical examples are presented to support theoretical results. As byproducts, a new stabilized low order mixed finite element method is proposed and analyzed and superconvergence results for the Hu-Zhang element are obtained.

1. Introduction

We consider fast solvers for the Hu-Zhang mixed finite element methods [46, 48, 49, 85, 86] for linear elasticity, namely fast solvers for inverting the following saddle point system,

$$\begin{pmatrix} M_h^\lambda & B_h^T \\ B_h & O \end{pmatrix}, \tag{1.1}$$

where M_h^λ is the mass matrix weighted by the compliance tensor and B_h is the discretization of the div operator. The subscript h is the mesh size of an underlying triangulation and the superscript λ is the Lamé number which could be very large for nearly incompressible material. We aim to develop preconditioners robust to both h and λ.

Received by the editor April 8, 2016, and, in revised form, February 12, 2017.

2010 *Mathematics Subject Classification.* Primary 65N55, 65F10, 65N22, 65N30.

The first author was supported by the National Science Foundation (NSF) DMS-1418934 and in part by the Sea Poly Project of Beijing Overseas Talents. This work was finished when the first author visited Peking University in the fall of 2015. He would like to thank Peking University for the support and hospitality, as well as for their exciting research atmosphere.

The second author was supported by the NSFC Projects 11625101, 91430213 and 11421101.

The third author is the corresponding author.

The third author was supported by the NSFC Projects 11771338, 11301396 and 11671304, Zhejiang Provincial Natural Science Foundation of China Projects LY17A010010, LY15A010015 and LY15A010016, and Wenzhou Science and Technology Plan Project G20160019.

In [46,48,49,85,86], a stability result is established in the $\boldsymbol{H}(\mathbf{div},\Omega;\mathbb{S})\times\boldsymbol{L}^2(\Omega;\mathbb{R}^n)$ norm whose matrix form is

$$\begin{pmatrix} M_h + B_h^T M_{u,h}^{-1} B_h & O \\ O & M_{u,h} \end{pmatrix},$$

where $M_{u,h}$ is the mass matrix for the displacement and M_h is the abbreviation of M_h^0. By the theory developed by Mardal and Winther [64], the following block-diagonal preconditioner leads to a parameter independent condition number of the preconditioned system

$$\begin{pmatrix} (M_h + B_h^T M_{u,h}^{-1} B_h)^{-1} & O \\ O & M_{u,h}^{-1} \end{pmatrix}. \tag{1.2}$$

To compute the first block of (1.2), however, a nontrivial solver should be designed to account for the discrete div operator.

Motivated by works [4,33,47], we shall establish another stability result of (1.1) in mesh-dependent norms $\|\cdot\|_{0,h}\times|\cdot|_{1,h}$ whose equivalent matrix form is

$$\begin{pmatrix} M_h & O \\ O & B_h M_h^{-1} B_h^T \end{pmatrix}.$$

Therefore we can use the block-diagonal preconditioner

$$\begin{pmatrix} M_h^{-1} & O \\ O & (B_h M_h^{-1} B_h^T)^{-1} \end{pmatrix} \tag{1.3}$$

together with the MINRES method to solve (1.1). It is worth mentioning that the stability in mesh-dependent norms established in [17,63] has been used to devise preconditioners for the mixed finite element methods for the Poisson problem in [43,69]. The mass matrix M_h can be further replaced by its diagonal matrix and thus a spectral equivalent approximation of M_h^{-1} is easy to construct. The difficulty is the inverse of the Schur complement.

We shall develop a fast auxiliary space preconditioner for the Schur complement. The auxiliary space preconditioner was initially designed by Xu [79] to avoid the difficulty in creating a sequence of nonnested grids or nonnested finite element spaces. As a two level method, the auxiliary space preconditioner involves smoothing on the fine level space which is usually the to-be-solved finite element space, and a coarse grid correction on an auxiliary space which is much more flexible to choose. It has been successfully applied to many finite element methods for partial differential equations [80], including conforming and nonconforming finite element methods for the second or fourth order problem [79,81], $H(\mathrm{curl})$ and $H(\mathrm{div})$ problems [44,53–55,74], DG type discretizations [29,32,35,61,82], and general symmetric positive definite problems [56], etc.

We use the H^1-conforming linear finite element discretization on the same mesh for the linear elasticity equation with parameter $\lambda = 0$ as the auxiliary problem to precondition the Schur complement. Since $\lambda = 0$, we can solve the auxiliary problem by geometric multigrid methods for structured meshes and algebraic multigrid methods in general. Using Korn's inequality, we can further adopt the H^1-conforming linear finite element discretization for the vector-type Poisson equation as the auxiliary problem.

Our stability result is robust with respect to both the parameters λ and h, therefore the condition number of the preconditioned system is uniformly bounded with respect to both the size of the problem and the parameter λ. The latter is difficult to construct for linear elasticity. Furthermore our results hold without the full regularity assumption.

To further improve the performance, we propose the following approximate block-factorization preconditioner

$$\begin{pmatrix} I & D_h^{-1}B_h^T \\ 0 & -I \end{pmatrix} \begin{pmatrix} D_h & 0 \\ B_h & \tilde{S}_h \end{pmatrix}^{-1}, \tag{1.4}$$

where D_h is the diagonal of M_h and $\tilde{S}_h = B_h D_h^{-1} B_h^T$ will be further preconditioned by the auxiliary space preconditioner we mentioned before. Numerical results in Section 6 show that the preconditioned GMRES converges around 40 steps to push the relative tolerance below 10^{-8}.

Results in this paper can also be applied to other $H(\mathrm{div})$ conforming and symmetric stress elements developed in [1, 5, 6, 8]. Indeed we present our results for both the original Hu-Zhang element $k \geq n+1$ and a new stabilized version for $1 \leq k \leq n$.

We now give a brief literature review on robust multigrid methods for the linear elasticity problem. Discretization of the linear elasticity equations can be roughly classified into three categories: the displacement primary formulation, the displacement-pressure mixed formulation, and the stress-displacement mixed formulation. Robust conforming and nonconforming multigrid methods for the primary formulation have been discussed in [60, 71, 77], and discontinuous Galerkin $H(\mathrm{div})$-conforming method in [45]. The W-cycle multigrid methods are the most studied multigrid methods for the displacement-pressure mixed formulation, which can be found in [25, 59] for conforming discretization and [19, 20] for nonconforming discretization. A V-cycle multigrid method for the finite difference discretization was developed in [84]. In [10], the Taylor-Hood element method was reduced to the pressure Schur complement equation, based on which an inner/outer iteration scheme was set up. So far the solvers for the stress-displacement mixed formulation are mainly concentrated on the block-diagonal preconditioned MINRES method; see [52, 66, 76]. In [52], the multigrid preconditioner was advanced for the PEERS element method with weakly symmetric stress. As for the Arnold-Winther element discretization, the overlapping Schwarz preconditioner was exploited in [76], and the variable V-cycle multigrid preconditioner was developed in [66]. Recently, an augmented Lagrangian Uzawa iteration was devised for the mixed finite element method with weakly imposed symmetry in [83], in which the saddle point problem was reduced to a nearly singular system and the HX preconditioner [44] was used for this nearly singular system. The majority of existing works deal with the discrete null space ker(div) by using either blockwise Gauss-Seidels smoother or overlapping Schwarz smoothers; only works in [25, 45, 52, 60, 66, 83] do not rely on the H^2 regularity assumption. As we mentioned early our approach does not require a prior knowledge of the discrete ker(div). We transfer this difficulty to solve the Schur complement problem but with $\lambda = 0$, which only involves standard Poisson-type solvers. So it is much easier to implement and analyze.

The rest of this article is organized as follows. In Section 2, we present the mixed finite element methods for linear elasticity. In Section 3, we establish the stability based on the mesh-dependent norms. Then we describe the block-diagonal and triangular preconditioners in Section 4 and construct an auxiliary space preconditioner in Section 5. In Section 6, we give some numerical experiments to demonstrate the efficiency and robustness of our preconditioners. Throughout this paper, we use "$\lesssim \cdots$" to mean that "$\leq C\cdots$", where C is a generic positive constant independent of h and the Lamé constant λ, which may take different values at different times, and $a \eqsim b$ means $a \lesssim b$ and $b \lesssim a$.

2. Mixed finite element methods

Assume that $\Omega \subset \mathbb{R}^n$ is a bounded polytope. Denote by $\mathbb{S}$ the space of all symmetric $n\times n$ tensors. Given a bounded domain $G \subset \mathbb{R}^n$ and a nonnegative integer m, let $H^m(G)$ be the usual Sobolev space of functions on G, and let $\boldsymbol{H}^m(G;\mathbb{X})$ be the usual Sobolev space of functions taking values in the finite-dimensional vector space $\mathbb{X}$ for $\mathbb{X}$ being $\mathbb{S}$ or $\mathbb{R}^n$. The corresponding norm and semi-norm are denoted, respectively, by $\|\cdot\|_{m,G}$ and $|\cdot|_{m,G}$. Let $(\cdot,\cdot)_G$ be the standard inner product on $L^2(G)$ or $\boldsymbol{L}^2(G;\mathbb{X})$. If G is Ω, we abbreviate $\|\cdot\|_{m,G}$, $|\cdot|_{m,G}$ and $(\cdot,\cdot)_G$ by $\|\cdot\|_m$, $|\cdot|_m$ and $(\cdot,\cdot)$, respectively. Let $\boldsymbol{H}_0^m(G;\mathbb{R}^n)$ be the closure of $\boldsymbol{C}_0^\infty(G;\mathbb{R}^n)$ with respect to the norm $\|\cdot\|_{m,G}$. Denote by $\boldsymbol{H}(\mathbf{div},G;\mathbb{S})$ the Sobolev space of square-integrable symmetric tensor fields with square-integrable divergence. For any $\boldsymbol{\tau}\in\boldsymbol{H}(\mathbf{div},\Omega;\mathbb{S})$, we define the following norm:

$$\|\boldsymbol{\tau}\|_{\boldsymbol{H}(\mathbf{div})} := \left(\|\boldsymbol{\tau}\|_0^2 + \|\mathbf{div}\boldsymbol{\tau}\|_0^2\right)^{1/2}.$$

The Hellinger-Reissner mixed formulation of the linear elasticity under the load $\boldsymbol{f}\in\boldsymbol{L}^2(\Omega;\mathbb{R}^n)$ is given as follows: Find $(\boldsymbol{\sigma},\boldsymbol{u})\in\boldsymbol{\Sigma}\times\boldsymbol{V} := \boldsymbol{H}(\mathbf{div},\Omega;\mathbb{S})\times\boldsymbol{L}^2(\Omega;\mathbb{R}^n)$ such that

$$a(\boldsymbol{\sigma},\boldsymbol{\tau}) + b(\boldsymbol{\tau},\boldsymbol{u}) = 0 \qquad \forall\,\boldsymbol{\tau}\in\boldsymbol{\Sigma}, \tag{2.1}$$

$$b(\boldsymbol{\sigma},\boldsymbol{v}) = -(\boldsymbol{f},\boldsymbol{v}) \qquad \forall\,\boldsymbol{v}\in\boldsymbol{V}, \tag{2.2}$$

where

$$a(\boldsymbol{\sigma},\boldsymbol{\tau}) := (\mathfrak{A}\boldsymbol{\sigma},\boldsymbol{\tau}), \quad b(\boldsymbol{\tau},\boldsymbol{v}) := (\mathbf{div}\boldsymbol{\tau},\boldsymbol{v})$$

with $\mathfrak{A}$ being the compliance tensor of fourth order defined by

$$\mathfrak{A}\boldsymbol{\sigma} := \frac{1}{2\mu}\left(\boldsymbol{\sigma} - \frac{\lambda}{n\lambda+2\mu}(\mathrm{tr}\boldsymbol{\sigma})\boldsymbol{I}\right).$$

Here $\boldsymbol{I}$ is the identity tensor, tr is the trace operator, and the positive constants λ and μ are the Lamé constants.

Suppose the domain Ω is subdivided by a family of shape regular simplicial grids $\mathcal{T}_h$ (cf. [26,34]) with $h := \max_{K\in\mathcal{T}_h} h_K$ and $h_K := \mathrm{diam}(K)$. Let $\mathcal{F}_h$ be the union of all $(n-1)$-dimensional faces of $\mathcal{T}_h$ and let $\mathcal{F}_h^i$ be the union of all $(n-1)$-dimensional interior faces. For any $F\in\mathcal{F}_h$, denote by h_F its diameter and fix a unit normal vector $\boldsymbol{\nu}_F$. Let $P_m(G)$ stand for the set of all polynomials in G with the total degree no more than m, and let $\boldsymbol{P}_m(G;\mathbb{X})$ denote the tensor or vector version of $P_m(G)$ for $\mathbb{X}$ being $\mathbb{S}$ or $\mathbb{R}^n$, respectively.

Consider two adjacent simplices K^+ and K^- sharing an interior face F. Denote by $\boldsymbol{\nu}^+$ and $\boldsymbol{\nu}^-$ the unit outward normals to the common face F of the simplices K^+ and K^-, respectively. For a vector-valued function $\boldsymbol{w}$, write $\boldsymbol{w}^+ := \boldsymbol{w}|_{K^+}$ and $\boldsymbol{w}^- := \boldsymbol{w}|_{K^-}$. Then define the jump of $\boldsymbol{w}$ as

$$[\boldsymbol{w}] := \begin{cases} \boldsymbol{w}^+(\boldsymbol{\nu}^+\cdot\boldsymbol{\nu}_F) + \boldsymbol{w}^-(\boldsymbol{\nu}^-\cdot\boldsymbol{\nu}_F), & \text{if } F\in\mathcal{F}_h^i, \\ \boldsymbol{w}, & \text{if } F\in\mathcal{F}_h\backslash\mathcal{F}_h^i. \end{cases}$$

For each $K \in \mathcal{T}_h$, define an $\boldsymbol{H}(\mathbf{div}, K; \mathbb{S})$ bubble function space of polynomials of degree k as

$$\boldsymbol{B}_{K,k} := \{\boldsymbol{\tau} \in \boldsymbol{P}_k(K;\mathbb{S}) : \boldsymbol{\tau}\boldsymbol{\nu}|_{\partial K} = \mathbf{0}\}.$$

It is easy to check that $\boldsymbol{B}_{K,1}$ is merely the zero space. Denote the vertices of simplex K by $\boldsymbol{x}_{K,0}, \ldots, \boldsymbol{x}_{K,n}$. If it does not cause confusion, we will abbreviate $\boldsymbol{x}_{K,i}$ as $\boldsymbol{x}_i$ for $i = 0, \ldots, n$. For any edge $\boldsymbol{x}_i\boldsymbol{x}_j (i \neq j)$ of element K, let $\boldsymbol{t}_{i,j}$ be the associated unit tangent vectors and let

$$\boldsymbol{T}_{i,j} := \boldsymbol{t}_{i,j}\boldsymbol{t}_{i,j}^T, \quad 0 \leq i < j \leq n.$$

It has been proved in [46] that the $(n+1)n/2$ symmetric tensors $\boldsymbol{T}_{i,j}$ form a basis of $\mathbb{S}$, and for $k \geq 2$,

$$\boldsymbol{B}_{K,k} = \sum_{0 \leq i < j \leq n} \lambda_i \lambda_j P_{i,j;k-2}(K) \boldsymbol{T}_{i,j},$$

where $P_{i,j;k-2}(K) := P_{k-2}(K)$ and λ_i is the associated barycentric coordinate corresponding to $\boldsymbol{x}_i$ for $0 \leq i < j \leq n$. Some global finite element spaces are given by

$$\begin{aligned}
\boldsymbol{B}_{k,h} &:= \{\boldsymbol{\tau} \in \boldsymbol{H}(\mathbf{div}, \Omega; \mathbb{S}) : \boldsymbol{\tau}|_K \in \boldsymbol{B}_{K,k} \quad \forall\, K \in \mathcal{T}_h\}, \\
\widetilde{\boldsymbol{\Sigma}}_{k,h} &:= \left\{\boldsymbol{\tau} \in \boldsymbol{H}^1(\Omega; \mathbb{S}) : \boldsymbol{\tau}|_K \in \boldsymbol{P}_k(K;\mathbb{S}) \quad \forall\, K \in \mathcal{T}_h\right\}, \\
\boldsymbol{\Sigma}_h &:= \widetilde{\boldsymbol{\Sigma}}_{k,h} + \boldsymbol{B}_{k,h}, \\
\boldsymbol{V}_h &:= \left\{\boldsymbol{v} \in \boldsymbol{L}^2(\Omega; \mathbb{R}^n) : \boldsymbol{v}|_K \in \boldsymbol{P}_{k-1}(K;\mathbb{R}^n) \quad \forall\, K \in \mathcal{T}_h\right\},
\end{aligned}$$

with integer $k \geq 1$. The local rigid motion space is defined as

$$\boldsymbol{R}(K) := \left\{\boldsymbol{v} \in \boldsymbol{H}^1(K;\mathbb{R}^n) : \boldsymbol{\varepsilon}(\boldsymbol{v}) = \mathbf{0}\right\}$$

with $\boldsymbol{\varepsilon}(\boldsymbol{v}) := \left(\boldsymbol{\nabla}\boldsymbol{v} + (\boldsymbol{\nabla}\boldsymbol{v})^T\right)/2$ being the linearized strain tensor.

With previous preparation, the mixed finite element method for linear elasticity proposed in [30,46,48,49] is defined as follows: Find $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h) \in \boldsymbol{\Sigma}_h \times \boldsymbol{V}_h$ such that

$$a(\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) + b(\boldsymbol{\tau}_h, \boldsymbol{u}_h) = 0 \qquad \forall\, \boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h, \tag{2.3}$$

$$b(\boldsymbol{\sigma}_h, \boldsymbol{v}_h) - c(\boldsymbol{u}_h, \boldsymbol{v}_h) = -(\boldsymbol{f}, \boldsymbol{v}_h) \qquad \forall\, \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{2.4}$$

where

$$c(\boldsymbol{u}_h, \boldsymbol{v}_h) := \eta \sum_{F \in \mathcal{F}_h} h_F^{-1} \int_F [\boldsymbol{u}_h] \cdot [\boldsymbol{v}_h] \, \mathrm{d}s,$$

$$\eta := \begin{cases} 0, & \text{if } k \geq n+1, \\ 1, & \text{if } 1 \leq k \leq n. \end{cases}$$

The bilinear form $c(\cdot,\cdot)$ involving the jump of displacement is introduced to stabilize the discretization which is only necessary for low order polynomials, i.e., $1 \leq k \leq n$. Note that the scaling h_F^{-1} is different than the one in [30].

Choosing appropriate bases of $\boldsymbol{\Sigma}_h$ and $\boldsymbol{V}_h$, we can write the matrix form of (2.3)-(2.4) as

$$\begin{pmatrix} M_h^\lambda & B_h^T \\ B_h & -C_h \end{pmatrix} \begin{pmatrix} \boldsymbol{\sigma}_h \\ \boldsymbol{u}_h \end{pmatrix} = \begin{pmatrix} 0 \\ \boldsymbol{f} \end{pmatrix}, \tag{2.5}$$

where M_h^λ is the mass matrix weighted by the compliance tensor, B_h is the discretization of the div operator, and C_h corresponds to the stabilization term. Here with a slight abuse of notation, we use the same notation $\boldsymbol{\sigma}_h, \boldsymbol{u}_h$, and $\boldsymbol{f}$ for the vector representations of corresponding functions.

Let

$$\hat{\boldsymbol{\Sigma}}_h := \{\boldsymbol{\tau} \in \boldsymbol{\Sigma}_h : \int_\Omega \operatorname{tr}\boldsymbol{\tau} \,\mathrm{d}x = 0\},$$

$$\mathbb{A}(\boldsymbol{\sigma}_h, \boldsymbol{u}_h; \boldsymbol{\tau}_h, \boldsymbol{v}_h) := a(\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) + b(\boldsymbol{\tau}_h, \boldsymbol{u}_h) + b(\boldsymbol{\sigma}_h, \boldsymbol{v}_h) - c(\boldsymbol{u}_h, \boldsymbol{v}_h).$$

For $k \geq n+1$, the following inf-sup condition is the immediate result of (3.4)-(3.5) in [46]:

$$\|\widetilde{\boldsymbol{\sigma}}_h\|_{\boldsymbol{H}(\mathbf{div})} + \|\widetilde{\boldsymbol{u}}_h\|_0 \lesssim \sup_{(\boldsymbol{\tau}_h, \boldsymbol{v}_h) \in \hat{\boldsymbol{\Sigma}}_h \times \boldsymbol{V}_h} \frac{\mathbb{A}(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{u}}_h; \boldsymbol{\tau}_h, \boldsymbol{v}_h)}{\|\boldsymbol{\tau}_h\|_{\boldsymbol{H}(\mathbf{div})} + \|\boldsymbol{v}_h\|_0}, \tag{2.6}$$

for any $(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{u}}_h) \in \hat{\boldsymbol{\Sigma}}_h \times \boldsymbol{V}_h$.

Thanks to the inf-sup condtion (2.6), the system (2.5) is stable in the space $\boldsymbol{\Sigma}_h \times \boldsymbol{V}_h$ equipped with the $\boldsymbol{H}(\mathbf{div}, \Omega; \mathbb{S}) \times \boldsymbol{L}^2(\Omega; \mathbb{R}^n)$-norm which leads to a block-diagonal preconditioner requiring a nontrivial solver for $(M_h + B_h^T M_{u,h}^{-1} B_h)^{-1}$. In the next section we shall establish another stability result of (2.5) in mesh-dependent norms which leads to a new block-diagonal preconditioner.

3. Stability based on mesh-dependent norms

To construct a new block-diagonal preconditioner, we will show that the bilinear form $\mathbb{A}(\cdot,\cdot;\cdot,\cdot)$ is stable on $\hat{\boldsymbol{\Sigma}}_h \times \boldsymbol{V}_h$ with mesh-dependent norms.

For each $K \in \mathcal{T}_h$, denote by $\boldsymbol{\nu}_i$ the unit outward normal vector of the ith face F_i of element K. For any $\boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h$ and $\boldsymbol{v}_h \in \boldsymbol{V}_h$, define

$$\|\boldsymbol{\tau}_h\|_{0,h}^2 := \|\boldsymbol{\tau}_h\|_0^2 + \sum_{F \in \mathcal{F}_h} h_F \|\boldsymbol{\tau}_h \boldsymbol{\nu}_F\|_{0,F}^2,$$

$$|\boldsymbol{v}_h|_{1,h}^2 := \|\boldsymbol{\varepsilon}_h(\boldsymbol{v}_h)\|_0^2 + \sum_{F \in \mathcal{F}_h} h_F^{-1} \|[\boldsymbol{v}_h]\|_{0,F}^2,$$

$$\|\boldsymbol{v}_h\|_c^2 := c(\boldsymbol{v}_h, \boldsymbol{v}_h).$$

Here $\boldsymbol{\varepsilon}_h$ is the elementwise symmetric gradient. We shall prove the stability of (2.5) in the mesh-dependent norms $\|\cdot\|_{0,h} \times |\cdot|_{1,h}$. The key is the following inf-sup condition: for $k \geq n+1$,

$$|\boldsymbol{v}_h|_{1,h} \lesssim \sup_{\boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h} \frac{b(\boldsymbol{\tau}_h, \boldsymbol{v}_h)}{\|\boldsymbol{\tau}_h\|_{0,h}} \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h. \tag{3.1}$$

For low order cases $1 \leq k \leq n$, in addition to a variant of the inf-sup condition, we also need a coercivity result in the null space of the div operator.

3.1. Properties of mesh-dependent norms.

We first present a different basis of the symmetric tensor space $\mathbb{S}$. Inside a simplex formed by vertices $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_n$, we label the face opposite to $\boldsymbol{x}_i$ as the ith face F_i. For the edge $\boldsymbol{x}_i\boldsymbol{x}_j$, $i \neq j$, define

$$\boldsymbol{N}_{i,j} := \frac{1}{2(\boldsymbol{\nu}_i^T \boldsymbol{t}_{i,j})(\boldsymbol{\nu}_j^T \boldsymbol{t}_{i,j})} (\boldsymbol{\nu}_i \boldsymbol{\nu}_j^T + \boldsymbol{\nu}_j \boldsymbol{\nu}_i^T), \quad 0 \leq i < j \leq n.$$

Here recall that $\boldsymbol{t}_{i,j}$ is a unit tangent vector of edge $\boldsymbol{x}_i\boldsymbol{x}_j$ and $\boldsymbol{\nu}_i$ is the unit outwards normal vector of face F_i. Due to the shape regularity of the triangulation, it holds that

$$\boldsymbol{\nu}_i^T \boldsymbol{t}_{i,j} \eqsim 1, \quad 0 \leq i < j \leq n.$$

By direct manipulation, we have the following results about $\boldsymbol{T}_{i,j}$ and $\boldsymbol{N}_{i,j}$:

$$\boldsymbol{T}_{i,j} : \boldsymbol{N}_{k,l} = \delta_{ik}\delta_{jl}, \quad 0 \leq i < j \leq n,\ 0 \leq k < l \leq n, \tag{3.2}$$

$$\boldsymbol{T}_{i,j} : \boldsymbol{T}_{i,j} = 1, \quad \boldsymbol{N}_{i,j} : \boldsymbol{N}_{i,j} \eqsim 1, \quad 0 \le i < j \le n. \tag{3.3}$$

Thus the $(n+1)n/2$ symmetric tensors $\{\boldsymbol{N}_{i,j}\}$ also form a basis of $\mathbb{S}$ which is the dual to $\{\boldsymbol{T}_{i,j}\}$.

Lemma 3.1. *For any $q_{ij} \in L^2(K)$, $0 \le i < j \le n$, let $\boldsymbol{\tau}_1 = \sum_{0\le i<j\le n} q_{ij}\boldsymbol{T}_{i,j}$ and let $\boldsymbol{\tau}_2 = \sum_{0\le i<j\le n} q_{ij}\boldsymbol{N}_{i,j}$, then it holds that*

$$\|\boldsymbol{\tau}_1\|_{0,K}^2 \eqsim \|\boldsymbol{\tau}_2\|_{0,K}^2 \eqsim \sum_{0\le i<j\le n} \|q_{ij}\|_{0,K}^2.$$

Proof. Using the Cauchy-Schwarz inequality and (3.3), we have

$$\|\boldsymbol{\tau}_1\|_{0,K}^2 \le \frac{(n+1)n}{2} \sum_{0\le i<j\le n} \|q_{ij}\boldsymbol{T}_{i,j}\|_{0,K}^2 = \frac{(n+1)n}{2} \sum_{0\le i<j\le n} \|q_{ij}\|_{0,K}^2,$$

$$\|\boldsymbol{\tau}_2\|_{0,K}^2 \le \frac{(n+1)n}{2} \sum_{0\le i<j\le n} \|q_{ij}\boldsymbol{N}_{i,j}\|_{0,K}^2 \lesssim \frac{(n+1)n}{2} \sum_{0\le i<j\le n} \|q_{ij}\|_{0,K}^2.$$

On the other side, it follows from Cauchy-Schwarz inequality and (3.2) that

$$\begin{aligned}
\sum_{0\le i<j\le n} \|q_{ij}\|_{0,K}^2 &= \sum_{0\le i<j\le n} \int_K q_{ij}^2 \,\mathrm{d}x = \sum_{0\le i<j\le n} \sum_{0\le k<l\le n} \int_K q_{ij} q_{kl} \delta_{ik}\delta_{jl} \,\mathrm{d}x \\
&= \sum_{0\le i<j\le n} \sum_{0\le k<l\le n} \int_K q_{ij}\boldsymbol{T}_{i,j} : q_{kl}\boldsymbol{N}_{k,l} \,\mathrm{d}x \\
&= \int_K \boldsymbol{\tau}_1 : \boldsymbol{\tau}_2 \,\mathrm{d}x \le \|\boldsymbol{\tau}_1\|_{0,K} \|\boldsymbol{\tau}_2\|_{0,K}.
\end{aligned}$$

Hence we conclude the result by combining the last three inequalities. □

We then map $\boldsymbol{\varepsilon}_h(\boldsymbol{V}_h)$ into the $\boldsymbol{H}(\mathbf{div}, K; \mathbb{S})$ bubble function space. For each element $K \in \mathcal{T}_h$, introduce a bijective connection operator $\boldsymbol{E}_K : \boldsymbol{P}_{k-2}(K;\mathbb{S}) \to \boldsymbol{B}_{K,k}$ with $k \ge 2$ as follows: For any $\boldsymbol{\tau} = \sum_{0\le i<j\le n} q_{ij}\boldsymbol{N}_{i,j}$ with $q_{ij} \in P_{k-2}(K)$, $0 \le i < j \le n$, define

$$\boldsymbol{E}_K \boldsymbol{\tau} := \sum_{0\le i<j\le n} \lambda_i\lambda_j q_{ij} \boldsymbol{T}_{i,j}.$$

Applying Lemma 3.1 and the scaling argument, we get for any $\boldsymbol{\tau} \in \boldsymbol{P}_{k-2}(K;\mathbb{S})$,

$$\|\boldsymbol{E}_K\boldsymbol{\tau}\|_{0,K}^2 \eqsim \sum_{0\le i<j\le n} \|\lambda_i\lambda_j q_{ij}\|_{0,K}^2 \eqsim \sum_{0\le i<j\le n} \|q_{ij}\|_{0,K}^2 \eqsim \|\boldsymbol{\tau}\|_{0,K}^2, \tag{3.4}$$

$$\int_K \boldsymbol{E}_K\boldsymbol{\tau} : \boldsymbol{\tau} \,\mathrm{d}x = \sum_{0\le i<j\le n} \int_K \lambda_i\lambda_j q_{ij}^2 \,\mathrm{d}x \eqsim \sum_{0\le i<j\le n} \|q_{ij}\|_{0,K}^2 \eqsim \|\boldsymbol{\tau}\|_{0,K}^2. \tag{3.5}$$

Denote by $\boldsymbol{E}$ the elementwise global version of $\boldsymbol{E}_K$, i.e., $\boldsymbol{E}|_K := \boldsymbol{E}_K$ for each $K \in \mathcal{T}_h$.

Next, we give an equivalent formulation of the mesh-dependent norm $|\cdot|_{1,h}$. For each $F \in \mathcal{F}_h$, denote by $\boldsymbol{\pi}_F$ the orthogonal projection operator from $\boldsymbol{L}^2(F;\mathbb{R}^n)$ onto $\boldsymbol{P}_1(F;\mathbb{R}^n)$. Define the broken $\boldsymbol{H}^1$ space as

$$\boldsymbol{H}^1(\mathcal{T}_h;\mathbb{R}^n) := \left\{\boldsymbol{v} \in \boldsymbol{L}^2(\Omega;\mathbb{R}^n) : \boldsymbol{v}|_K \in \boldsymbol{H}^1(K;\mathbb{R}^n) \quad \forall\, K \in \mathcal{T}_h\right\}.$$

The domain of the mesh-dependent norm $|\cdot|_{1,h}$ can be extended from $\boldsymbol{V}_h$ to $\boldsymbol{H}^1(\mathcal{T}_h;\mathbb{R}^n)$.

Lemma 3.2. *We have the norm equivalence:*

$$|\boldsymbol{v}|_{1,h}^2 \eqsim \|\boldsymbol{\varepsilon}_h(\boldsymbol{v})\|_0^2 + \sum_{F\in\mathcal{F}_h} h_F^{-1}\|\boldsymbol{\pi}_F[\boldsymbol{v}]\|_{0,F}^2 \quad \forall\ \boldsymbol{v}\in \boldsymbol{H}^1(\mathcal{T}_h;\mathbb{R}^n). \tag{3.6}$$

Proof. For any element $K\in\mathcal{T}_h$, let $\boldsymbol{\pi}_K$ be an interpolation operator from $\boldsymbol{H}_1(K;\mathbb{R}^n)$ onto $\boldsymbol{R}(K)$ defined by (3.1)-(3.2) in [22], and let $\boldsymbol{\pi}$ be the elementwise global version of $\boldsymbol{\pi}_K$, i.e., $\boldsymbol{\pi}|_K := \boldsymbol{\pi}_K$ for each $K\in\mathcal{T}_h$. It follows from (3.3)-(3.4) in [22] that for any $\boldsymbol{v}\in \boldsymbol{H}^1(\mathcal{T}_h;\mathbb{R}^n)$,

$$\begin{aligned}\sum_{F\in\mathcal{F}_h} h_F^{-1}\|[\boldsymbol{v}]-\boldsymbol{\pi}_F[\boldsymbol{v}]\|_{0,F}^2 &= \sum_{F\in\mathcal{F}_h} h_F^{-1}\|[\boldsymbol{v}-\boldsymbol{\pi}\boldsymbol{v}]-\boldsymbol{\pi}_F[\boldsymbol{v}-\boldsymbol{\pi}\boldsymbol{v}]\|_{0,F}^2\\ &\le \sum_{F\in\mathcal{F}_h} h_F^{-1}\|[\boldsymbol{v}-\boldsymbol{\pi}\boldsymbol{v}]\|_{0,F}^2 \lesssim \|\boldsymbol{\varepsilon}_h(\boldsymbol{v})\|_0^2.\end{aligned} \tag{3.7}$$

Then the equivalence (3.6) follows from the triangle inequality. □

We shall also use the following discrete Korn's inequality (cf. (1.22) in [22] and (34) in [7])

$$\|\boldsymbol{\nabla}_h\boldsymbol{v}\|_0^2 + \|\boldsymbol{v}\|_0^2 \lesssim \|\boldsymbol{\varepsilon}_h(\boldsymbol{v})\|_0^2 + \sum_{F\in\mathcal{F}_h} h_F^{-1}\|\boldsymbol{\pi}_F[\boldsymbol{v}]\|_{0,F}^2 \quad \forall\ \boldsymbol{v}\in \boldsymbol{H}^1(\mathcal{T}_h;\mathbb{R}^n). \tag{3.8}$$

Together with (3.6), we conclude $|\cdot|_{1,h}$ defines a norm on $\boldsymbol{V}_h$.

3.2. **inf-sup condition in mesh-dependent norms.** The inf-sup condition we need is actually for the subspace $\hat{\boldsymbol{\Sigma}}_h$ with vanished mean trace, cf., (3.9) below. It is obvious that the inf-sup condition (3.9) implies the inf-sup condition (3.1). On the other hand, if the inf-sup condition (3.1) is true, then (3.9) holds by taking $\hat{\boldsymbol{\tau}}_h = \boldsymbol{\tau}_h - (\frac{1}{n}\int_\Omega \mathrm{tr}\boldsymbol{\tau}_h\ \mathrm{d}x)\boldsymbol{\delta}$. Therefore the inf-sup conditions (3.1) and (3.9) are equivalent.

Lemma 3.3. *For $k\ge n+1$, we have the following inf-sup condition,*

$$|\boldsymbol{v}_h|_{1,h} \lesssim \sup_{\hat{\boldsymbol{\tau}}_h\in\hat{\boldsymbol{\Sigma}}_h} \frac{b(\hat{\boldsymbol{\tau}}_h,\boldsymbol{v}_h)}{\|\hat{\boldsymbol{\tau}}_h\|_{0,h}}, \tag{3.9}$$

for any $\boldsymbol{v}_h\in\boldsymbol{V}_h$.

Proof. Given a $\boldsymbol{v}_h\in\boldsymbol{V}_h$, we shall construct a $\hat{\boldsymbol{\tau}}_h\in\hat{\boldsymbol{\Sigma}}_h$ to verify (3.9).

We first control the norm $\|\boldsymbol{\varepsilon}_h(\boldsymbol{v}_h)\|_0$. For any $\boldsymbol{v}_h\in\boldsymbol{V}_h$, take $\boldsymbol{\tau}_1 = \boldsymbol{E}\boldsymbol{\varepsilon}_h(\boldsymbol{v}_h)$. It follows from (3.4) that

$$\|\boldsymbol{\tau}_1\|_0 \eqsim \|\boldsymbol{\varepsilon}_h(\boldsymbol{v}_h)\|_0. \tag{3.10}$$

According to integration by parts and (3.5), there exists a constant $C_1>0$ such that

$$b(\boldsymbol{\tau}_1,\boldsymbol{v}_h) = \int_\Omega \boldsymbol{\tau}_1 : \boldsymbol{\varepsilon}_h(\boldsymbol{v}_h)\ \mathrm{d}x \ge C_1\|\boldsymbol{\varepsilon}_h(\boldsymbol{v}_h)\|_0^2. \tag{3.11}$$

Next we control the jump term. Choose $\boldsymbol{\tau}_2\in\boldsymbol{\Sigma}_h$ such that all the degrees of freedom (cf. Lemma 2.1 in [30]) for $\boldsymbol{\tau}_2$ vanish except the following one:

$$\int_F (\boldsymbol{\tau}_2\boldsymbol{\nu}_F)\cdot\boldsymbol{w}\ \mathrm{d}s = h_F^{-1}\int_F [\boldsymbol{v}_h]\cdot\boldsymbol{w}\ \mathrm{d}s \quad \forall\ \boldsymbol{w}\in\boldsymbol{P}_1(F;\mathbb{R}^n) \text{ on each face } F.$$

Then we have

$$\int_K \boldsymbol{\tau}_2 : \boldsymbol{\varepsilon}_h(\boldsymbol{v}_h)\ \mathrm{d}x = 0, \quad \int_F (\boldsymbol{\tau}_2\boldsymbol{\nu}_F)\cdot \boldsymbol{\pi}_F[\boldsymbol{v}_h]\ \mathrm{d}s = h_F^{-1}\|\boldsymbol{\pi}_F[\boldsymbol{v}_h]\|_{0,F}^2,$$

$$\|\boldsymbol{\tau}_2\|_0^2 \lesssim \sum_{F\in\mathcal{T}_h} h_F^{-1}\|\boldsymbol{\pi}_F[\boldsymbol{v}_h]\|_{0,F}^2. \tag{3.12}$$

Thus by (3.7) and (3.12), there exists a constant $C_2 > 0$ such that

$$\begin{aligned} b(\boldsymbol{\tau}_2, \boldsymbol{v}_h) &= \sum_{F\in\mathcal{F}_h} \int_F (\boldsymbol{\tau}_2\boldsymbol{\nu}_F)\cdot[\boldsymbol{v}_h]\ \mathrm{d}s \\ &= \sum_{F\in\mathcal{F}_h} \int_F (\boldsymbol{\tau}_2\boldsymbol{\nu}_F)\cdot([\boldsymbol{v}_h] - \boldsymbol{\pi}_F[\boldsymbol{v}_h])\ \mathrm{d}s + \sum_{F\in\mathcal{F}_h} h_F^{-1}\|\boldsymbol{\pi}_F[\boldsymbol{v}_h]\|_{0,F}^2 \\ &\geq -C_2\|\boldsymbol{\varepsilon}_h(\boldsymbol{v}_h)\|_0^2 + \frac{1}{2}\sum_{F\in\mathcal{F}_h} h_F^{-1}\|\boldsymbol{\pi}_F[\boldsymbol{v}_h]\|_{0,F}^2. \end{aligned} \tag{3.13}$$

Now taking $\boldsymbol{\tau}_h = \boldsymbol{\tau}_1 + \frac{C_1}{2C_2}\boldsymbol{\tau}_2$, it holds from (3.11) and (3.13) that

$$\begin{aligned} b(\boldsymbol{\tau}_h, \boldsymbol{v}_h) &= b(\boldsymbol{\tau}_1, \boldsymbol{v}_h) + \frac{C_1}{2C_2} b(\boldsymbol{\tau}_2, \boldsymbol{v}_h) \\ &\geq \frac{C_1}{2}\|\boldsymbol{\varepsilon}_h(\boldsymbol{v}_h)\|_0^2 + \frac{C_1}{4C_2}\sum_{F\in\mathcal{F}_h} h_F^{-1}\|\boldsymbol{\pi}_F[\boldsymbol{v}_h]\|_{0,F}^2. \end{aligned}$$

Thanks to (3.6), we get

$$|\boldsymbol{v}_h|_{1,h}^2 \lesssim b(\boldsymbol{\tau}_h, \boldsymbol{v}_h).$$

On the other hand, it follows from the inverse inequality, (3.10) and (3.12) that

$$\|\boldsymbol{\tau}_h\|_{0,h} \lesssim \|\boldsymbol{\tau}_h\|_0 \lesssim |\boldsymbol{v}_h|_{1,h}.$$

Finally, the inf-sup condition (3.1) is the result of the last two inequalities and consequently (3.9) holds by taking $\hat{\boldsymbol{\tau}}_h = \boldsymbol{\tau}_h - (\frac{1}{n}\int_\Omega \mathrm{tr}\boldsymbol{\tau}_h\ \mathrm{d}x)\boldsymbol{\delta}$. □

3.3. **Coercivity in the null space of the div operator.** Besides the inf-sup condition, another issue of the linear elasticity in the mixed form is the coercivity of bilinear form $a(\cdot,\cdot)$. On the whole space: for all $\boldsymbol{\sigma}\in\boldsymbol{\Sigma}$,

$$a(\boldsymbol{\sigma},\boldsymbol{\sigma}) \geq \frac{1}{n\lambda + 2\mu}\|\boldsymbol{\sigma}\|_0^2. \tag{3.14}$$

The coercivity constant is, unfortunately, in the order of $\mathcal{O}(1/\lambda)$ as $\lambda \to +\infty$. Namely it is not robust to λ. To obtain a robust coercivity, we first recall the following inequality which implies the coercivity in the null space of the div operator.

Lemma 3.4 (Proposition 9.1.1 in [16]). *For $\boldsymbol{\tau}\in\boldsymbol{H}(\mathbf{div},\Omega;\mathbb{S})$ satisfying $\int_\Omega \mathrm{tr}\boldsymbol{\tau}\ \mathrm{d}x = 0$, we have*

$$\|\boldsymbol{\tau}\|_0 \lesssim \|\boldsymbol{\tau}\|_a + \|\operatorname{div}\boldsymbol{\tau}\|_{-1},$$

where $\|\boldsymbol{\tau}\|_a^2 := a(\boldsymbol{\tau},\boldsymbol{\tau})$ and $\|\operatorname{div}\boldsymbol{\tau}\|_{-1} := \sup_{\boldsymbol{v}\in\boldsymbol{H}_0^1(\Omega;\mathbb{R}^n)} b(\boldsymbol{\tau},\boldsymbol{v})/|\boldsymbol{v}|_1$.

We then move to the discrete case. Define discrete norms

$$\|\operatorname{div}\boldsymbol{\tau}\|_{-1,h} := \sup_{\boldsymbol{v}_h\in\boldsymbol{V}_h} \frac{b(\boldsymbol{\tau},\boldsymbol{v}_h)}{|\boldsymbol{v}_h|_{1,h}},$$

$$\|h\operatorname{div}\boldsymbol{\tau}\|^2 := \sum_{K\in\mathcal{T}_h} h_K^2\|\mathbf{div}\boldsymbol{\tau}\|_{0,K}^2.$$

Let $\boldsymbol{Q}_h^{k-1}$ be the L^2 orthogonal projection from $\boldsymbol{L}^2(\Omega;\mathbb{R}^n)$ onto $\boldsymbol{V}_h$, which will be abbreviated as $\boldsymbol{Q}_h$. It holds the following error estimate (cf. [26, 34]):

$$\|\boldsymbol{v}-\boldsymbol{Q}_h\boldsymbol{v}\|_{0,K}+h_K^{1/2}\|\boldsymbol{v}-\boldsymbol{Q}_h\boldsymbol{v}\|_{0,\partial K}\lesssim h_K^{\min\{k,m\}}|\boldsymbol{v}|_{m,K}\quad\forall\,\boldsymbol{v}\in\boldsymbol{H}^m(\Omega;\mathbb{R}^n),\tag{3.15}$$

with integer $m\geq 1$.

Lemma 3.5. *For any $\boldsymbol{\tau}\in\boldsymbol{H}(\mathbf{div},\Omega;\mathbb{S})$ satisfying $\int_\Omega \operatorname{tr}\boldsymbol{\tau}\,\mathrm{d}x=0$, we have*

$$\|\boldsymbol{\tau}\|_0\lesssim\|\boldsymbol{\tau}\|_a+\|h\operatorname{div}\boldsymbol{\tau}\|+\|\operatorname{div}\boldsymbol{\tau}\|_{-1,h}.$$

Proof. It is sufficient to prove the case $k=1$. Let $\boldsymbol{v}\in\boldsymbol{H}_0^1(\Omega;\mathbb{R}^n)$, then it follows from the Cauchy-Schwarz inequality and (3.15) that

$$\begin{aligned}b(\boldsymbol{\tau},\boldsymbol{v})&=b(\boldsymbol{\tau},\boldsymbol{v}-\boldsymbol{Q}_h\boldsymbol{v})+b(\boldsymbol{\tau},\boldsymbol{Q}_h\boldsymbol{v})\\&\lesssim\|h\operatorname{div}\boldsymbol{\tau}\||\boldsymbol{v}|_1+b(\boldsymbol{\tau},\boldsymbol{Q}_h\boldsymbol{v}).\end{aligned}$$

Again, by (3.15) and $[\boldsymbol{v}]=\mathbf{0}$ on $\mathcal{F}_h$ since $\boldsymbol{v}\in\boldsymbol{H}_0^1(\Omega;\mathbb{R}^n)$, it holds that

$$|\boldsymbol{Q}_h\boldsymbol{v}|_{1,h}^2=\sum_{F\in\mathcal{F}_h}h_F^{-1}\|[\boldsymbol{Q}_h\boldsymbol{v}]\|_{0,F}^2=\sum_{F\in\mathcal{F}_h}h_F^{-1}\|[\boldsymbol{Q}_h\boldsymbol{v}-\boldsymbol{v}]\|_{0,F}^2\lesssim|\boldsymbol{v}|_1^2.\tag{3.16}$$

Hence we get from the last two inequalities that

$$\|\operatorname{div}\boldsymbol{\tau}\|_{-1}=\sup_{\boldsymbol{v}\in\boldsymbol{H}_0^1(\Omega;\mathbb{R}^n)}\frac{b(\boldsymbol{\tau},\boldsymbol{v})}{|\boldsymbol{v}|_1}\lesssim\|h\operatorname{div}\boldsymbol{\tau}\|+\sup_{\boldsymbol{v}_h\in\boldsymbol{V}_h}\frac{b(\boldsymbol{\tau},\boldsymbol{v}_h)}{|\boldsymbol{v}_h|_{1,h}}.$$

Therefore we can end the proof by using Lemma 3.4. ∎

3.4. **Stability in the mesh-dependent norms.** We now present stability in mesh-dependent norms. For $k\geq n+1$, since there is no stabilization term and $\operatorname{div}\boldsymbol{\Sigma}_h\subset\boldsymbol{V}_h$, then $\ker(\operatorname{div})\cap\boldsymbol{\Sigma}_h\subset\ker(\operatorname{div})\cap\boldsymbol{\Sigma}$. The stability follows from Lemma 3.4 and inf-sup condition (3.9).

Theorem 3.6. *For $k\geq n+1$, it follows that for any $(\widetilde{\boldsymbol{\sigma}}_h,\widetilde{\boldsymbol{u}}_h)\in\hat{\boldsymbol{\Sigma}}_h\times\boldsymbol{V}_h$,*

$$\|\widetilde{\boldsymbol{\sigma}}_h\|_{0,h}+|\widetilde{\boldsymbol{u}}_h|_{1,h}\lesssim\sup_{(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\in\hat{\boldsymbol{\Sigma}}_h\times\boldsymbol{V}_h}\frac{\mathbb{A}(\widetilde{\boldsymbol{\sigma}}_h,\widetilde{\boldsymbol{u}}_h;\boldsymbol{\tau}_h,\boldsymbol{v}_h)}{\|\boldsymbol{\tau}_h\|_{0,h}+|\boldsymbol{v}_h|_{1,h}}.\tag{3.17}$$

Corollary 3.7. *Let $k\geq n+1$. Assume that $\boldsymbol{\sigma}\in\boldsymbol{H}^{k+1}(\Omega;\mathbb{S})$ and $\boldsymbol{u}\in\boldsymbol{H}^k(\Omega;\mathbb{R}^n)$, then*

$$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{0,h}+|\boldsymbol{Q}_h\boldsymbol{u}-\boldsymbol{u}_h|_{1,h}\lesssim h^{k+1}\|\boldsymbol{\sigma}\|_{k+1},\tag{3.18}$$

$$|\boldsymbol{u}-\boldsymbol{u}_h|_{1,h}\lesssim h^{k-1}\left(\|\boldsymbol{\sigma}\|_{k+1}+\|\boldsymbol{u}\|_k\right).\tag{3.19}$$

Moreover, when Ω is convex, we have

$$\|\boldsymbol{Q}_h\boldsymbol{u}-\boldsymbol{u}_h\|_0\lesssim h^{k+2}\|\boldsymbol{\sigma}\|_{k+1}.\tag{3.20}$$

Proof. Subtracting (2.3)-(2.4) from (2.1)-(2.2), we get the error equation

$$a(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{\tau}_h)+b(\boldsymbol{\tau}_h,\boldsymbol{u}-\boldsymbol{u}_h)=0\qquad\forall\,\boldsymbol{\tau}_h\in\boldsymbol{\Sigma}_h,\tag{3.21}$$

$$b(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{v}_h)=0\qquad\forall\,\boldsymbol{v}_h\in\boldsymbol{V}_h.\tag{3.22}$$

Let $\boldsymbol{I}_h^{HZ}$ be the standard interpolation from $\boldsymbol{H}^1(\Omega;\mathbb{S})$ to $\boldsymbol{\Sigma}_h$ defined in [46, Remark 3.1], and it holds that

$$\mathbf{div}(\boldsymbol{I}_h^{HZ}\boldsymbol{\sigma})=\boldsymbol{Q}_h(\mathbf{div}\boldsymbol{\sigma}).\tag{3.23}$$

Thus we have from (3.22) that

$$b(\boldsymbol{I}_h^{HZ}\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{v}_h) = b(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{v}_h) = 0.$$

By the definition of $\boldsymbol{Q}_h$ and (3.21),

$$b(\boldsymbol{\tau}_h, \boldsymbol{Q}_h\boldsymbol{u} - \boldsymbol{u}_h) = b(\boldsymbol{\tau}_h, \boldsymbol{u} - \boldsymbol{u}_h) = -a(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{\tau}_h).$$

Combining the last two equalities, it holds that

$$\begin{aligned}
&\mathbb{A}(\boldsymbol{I}_h^{HZ}\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{Q}_h\boldsymbol{u} - \boldsymbol{u}_h; \boldsymbol{\tau}_h, \boldsymbol{v}_h)\\
&\quad = a(\boldsymbol{I}_h^{HZ}\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) + b(\boldsymbol{\tau}_h, \boldsymbol{Q}_h\boldsymbol{u} - \boldsymbol{u}_h) + b(\boldsymbol{I}_h^{HZ}\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{v}_h)\\
&\quad = a(\boldsymbol{I}_h^{HZ}\boldsymbol{\sigma} - \boldsymbol{\sigma}, \boldsymbol{\tau}_h),
\end{aligned}$$

which together with (3.17) implies

$$\|\boldsymbol{I}_h^{HZ}\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,h} + |\boldsymbol{Q}_h\boldsymbol{u} - \boldsymbol{u}_h|_{1,h} \lesssim \|\boldsymbol{I}_h^{HZ}\boldsymbol{\sigma} - \boldsymbol{\sigma}\|_{0,h}.$$

Therefore we will achieve (3.18)-(3.19) by using the last inequality, and the error estimate of $\boldsymbol{I}_h^{HZ}$ and $\boldsymbol{Q}_h$. The error estimate (3.20) can be derived by using the duality argument as in [39, 73]. □

Remark 3.8. The optimal convergence rate of $\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,h}$ has been proved in [46, Remarks 3.1-3.2] and [49, Remarks 3.6], but the second order higher superconvergent rates of $|\boldsymbol{Q}_h\boldsymbol{u} - \boldsymbol{u}_h|_{1,h}$ and $\|\boldsymbol{Q}_h\boldsymbol{u} - \boldsymbol{u}_h\|_0$ are new which can be used to reconstruct a better approximation of displacement. The convergence rate of $|\boldsymbol{u} - \boldsymbol{u}_h|_{1,h}$ is also optimal.

Due to the stabilization term (for the inf-sup condition), our proof of the stability is more complicated for the lower order case $1 \le k \le n$, which is similar to the one in [30, Lemma 3.2].

Theorem 3.9. *For $1 \le k \le n$, it holds for any $(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{u}}_h) \in \hat{\boldsymbol{\Sigma}}_h \times \boldsymbol{V}_h$ that*

$$\|\widetilde{\boldsymbol{\sigma}}_h\|_{0,h} + |\widetilde{\boldsymbol{u}}_h|_{1,h} \lesssim \sup_{(\boldsymbol{\tau}_h, \boldsymbol{v}_h) \in \hat{\boldsymbol{\Sigma}}_h \times \boldsymbol{V}_h} \frac{\mathbb{A}(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{u}}_h; \boldsymbol{\tau}_h, \boldsymbol{v}_h)}{\|\boldsymbol{\tau}_h\|_{0,h} + |\boldsymbol{v}_h|_{1,h}}. \tag{3.24}$$

Proof. As demonstrated in Lemma 3.3, it is equivalent to prove that

$$\|\widetilde{\boldsymbol{\sigma}}_h\|_{0,h} + |\widetilde{\boldsymbol{u}}_h|_{1,h} \lesssim \sup_{(\boldsymbol{\tau}_h, \boldsymbol{v}_h) \in \boldsymbol{\Sigma}_h \times \boldsymbol{V}_h} \frac{\mathbb{A}(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{u}}_h; \boldsymbol{\tau}_h, \boldsymbol{v}_h)}{\|\boldsymbol{\tau}_h\|_{0,h} + |\boldsymbol{v}_h|_{1,h}} := \beta. \tag{3.25}$$

The notation β is introduced just for ease of presentation. Let $\boldsymbol{\tau}_1 = \boldsymbol{E}\varepsilon_h(\widetilde{\boldsymbol{u}}_h)$ for $k \ge 2$ and let $\boldsymbol{\tau}_1 = \boldsymbol{0}$ for $k = 1$, then it holds from Cauchy-Schwarz inequality that

$$\mathbb{A}(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{u}}_h; \boldsymbol{\tau}_1, 0) = a(\widetilde{\boldsymbol{\sigma}}_h, \boldsymbol{\tau}_1) + b(\boldsymbol{\tau}_1, \widetilde{\boldsymbol{u}}_h) \ge -\|\widetilde{\boldsymbol{\sigma}}_h\|_a \|\boldsymbol{\tau}_1\|_a + b(\boldsymbol{\tau}_1, \widetilde{\boldsymbol{u}}_h).$$

Using (3.10)-(3.11), there exists a constant $C_3 > 0$ such that

$$\begin{aligned}
\mathbb{A}(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{u}}_h; \boldsymbol{\tau}_1, 0) &\ge C_1\|\varepsilon_h(\widetilde{\boldsymbol{u}}_h)\|_0^2 - C_3\|\widetilde{\boldsymbol{\sigma}}_h\|_a\|\varepsilon_h(\widetilde{\boldsymbol{u}}_h)\|_0\\
&\ge \frac{C_1}{2}\|\varepsilon_h(\widetilde{\boldsymbol{u}}_h)\|_0^2 - \frac{C_3^2}{2C_1}\|\widetilde{\boldsymbol{\sigma}}_h\|_a^2.
\end{aligned} \tag{3.26}$$

Let $\boldsymbol{v}_1 \in \boldsymbol{V}_h$ such that $\boldsymbol{v}_1|_K = h_K^2 \mathbf{div}\widetilde{\boldsymbol{\sigma}}_h$ for each $K \in \mathcal{T}_h$. Applying inverse inequality, we have

$$|\boldsymbol{v}_1|_{1,h} \lesssim \|h\,\mathrm{div}\,\widetilde{\boldsymbol{\sigma}}_h\| \lesssim \|\widetilde{\boldsymbol{\sigma}}_h\|_0. \tag{3.27}$$

Thus there exists a constant $C_4 > 0$ such that

$$\begin{aligned}\mathbb{A}(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{u}}_h; 0, \boldsymbol{v}_1) &= b(\widetilde{\boldsymbol{\sigma}}_h, \boldsymbol{v}_1) - c(\widetilde{\boldsymbol{u}}_h, \boldsymbol{v}_1) \geq \|h \operatorname{div} \widetilde{\boldsymbol{\sigma}}_h\|^2 - \|\widetilde{\boldsymbol{u}}_h\|_c \|\boldsymbol{v}_1\|_c \\ &\geq \|h \operatorname{div} \widetilde{\boldsymbol{\sigma}}_h\|^2 - C_4 \|\widetilde{\boldsymbol{u}}_h\|_c \|h \operatorname{div} \widetilde{\boldsymbol{\sigma}}_h\| \\ &\geq \frac{1}{2}\|h \operatorname{div} \widetilde{\boldsymbol{\sigma}}_h\|^2 - \frac{C_4^2}{2}\|\widetilde{\boldsymbol{u}}_h\|_c^2. \end{aligned} \tag{3.28}$$

Now taking $\boldsymbol{\tau}_h = \widetilde{\boldsymbol{\sigma}}_h + \frac{C_1}{C_3^2}\boldsymbol{\tau}_1$ and $\boldsymbol{v}_h = -\widetilde{\boldsymbol{u}}_h + \frac{1}{C_4^2}\boldsymbol{v}_1$, we have from (3.26) and (3.28)

$$\begin{aligned}&\mathbb{A}(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{u}}_h; \boldsymbol{\tau}_h, \boldsymbol{v}_h) \\ &\quad = \mathbb{A}(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{u}}_h; \widetilde{\boldsymbol{\sigma}}_h, -\widetilde{\boldsymbol{u}}_h) + \frac{C_1}{C_3^2}\mathbb{A}(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{u}}_h; \boldsymbol{\tau}_1, 0) + \frac{1}{C_4^2}\mathbb{A}(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{u}}_h; 0, \boldsymbol{v}_1) \\ &\quad \geq \frac{1}{2}\|\widetilde{\boldsymbol{\sigma}}_h\|_a^2 + \frac{1}{2}\|\widetilde{\boldsymbol{u}}_h\|_c^2 + \frac{1}{2C_4^2}\|h \operatorname{div} \widetilde{\boldsymbol{\sigma}}_h\|^2 + \frac{C_1^2}{2C_3^2}\|\boldsymbol{\varepsilon}_h(\widetilde{\boldsymbol{u}}_h)\|_0^2,\end{aligned}$$

which together with (3.10) and (3.27) indicates

$$\|\widetilde{\boldsymbol{\sigma}}_h\|_a^2 + \|h \operatorname{div} \widetilde{\boldsymbol{\sigma}}_h\|^2 + |\widetilde{\boldsymbol{u}}_h|_{1,h}^2 \lesssim \mathbb{A}(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{u}}_h; \boldsymbol{\tau}_h, \boldsymbol{v}_h) \lesssim \beta(\|\widetilde{\boldsymbol{\sigma}}_h\|_0 + |\widetilde{\boldsymbol{u}}_h|_{1,h}).$$

According to Lemma 3.5 and the definition of β, it holds that

$$\begin{aligned}\|\widetilde{\boldsymbol{\sigma}}_h\|_0 &\lesssim \|\widetilde{\boldsymbol{\sigma}}_h\|_a + \|h \operatorname{div} \widetilde{\boldsymbol{\sigma}}_h\| + \sup_{\boldsymbol{v}_h \in \boldsymbol{V}_h} \frac{b(\widetilde{\boldsymbol{\sigma}}_h, \boldsymbol{v}_h)}{|\boldsymbol{v}_h|_{1,h}} \\ &= \|\widetilde{\boldsymbol{\sigma}}_h\|_a + \|h \operatorname{div} \widetilde{\boldsymbol{\sigma}}_h\| + \sup_{\boldsymbol{v}_h \in \boldsymbol{V}_h} \frac{\mathbb{A}(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{u}}_h; \boldsymbol{0}, \boldsymbol{v}_h) + c(\widetilde{\boldsymbol{u}}_h, \boldsymbol{v}_h)}{|\boldsymbol{v}_h|_{1,h}} \\ &\lesssim \|\widetilde{\boldsymbol{\sigma}}_h\|_a + \|h \operatorname{div} \widetilde{\boldsymbol{\sigma}}_h\| + |\widetilde{\boldsymbol{u}}_h|_{1,h} + \beta.\end{aligned}$$

Thus we obtain from the last two inequalities that

$$\begin{aligned}\|\widetilde{\boldsymbol{\sigma}}_h\|_0^2 + |\widetilde{\boldsymbol{u}}_h|_{1,h}^2 &\lesssim \|\widetilde{\boldsymbol{\sigma}}_h\|_a^2 + \|h \operatorname{div} \widetilde{\boldsymbol{\sigma}}_h\|^2 + |\widetilde{\boldsymbol{u}}_h|_{1,h}^2 + \beta^2 \\ &\lesssim \beta(\|\widetilde{\boldsymbol{\sigma}}_h\|_0 + |\widetilde{\boldsymbol{u}}_h|_{1,h}) + \beta^2,\end{aligned}$$

which implies inf-sup condition (3.25). □

Corollary 3.10. *Let $1 \leq k \leq n$. Assume that $\boldsymbol{\sigma} \in \boldsymbol{H}^{k+1}(\Omega; \mathbb{S})$ and $\boldsymbol{u} \in \boldsymbol{H}^k(\Omega; \mathbb{R}^n)$, then*

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,h} + |\boldsymbol{u} - \boldsymbol{u}_h|_{1,h} \lesssim h^{k-1}\left(\|\boldsymbol{\sigma}\|_{k+1} + \|\boldsymbol{u}\|_k\right).$$

The convergence rate of $|\boldsymbol{u} - \boldsymbol{u}_h|_{1,h}$ is optimal. But the L^2-type error of $\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,h}$ is of two order less.

Remark 3.11. Using the stability in mesh-dependent norms established in [17, 63], the MINRES method with additive Schwarz preconditioner was developed for the mixed finite element methods of the Poisson problem in [69], and the CG method with the auxiliary space preconditioner for the corresponding Schur complement problem was designed in [43]. Similar stability in the mesh-dependent norm for the mixed finite macroelement methods of the linear elasticity can be found in [72], hence the fast auxiliary space preconditioner constructed in this paper can be easily extended to these mixed methods.

3.5. **Postprocessing.** Based on the superconvergent results of the displacement in (3.18) and (3.20), we will construct a superconvergent postprocessed displacement from $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h)$ for the higher order case $k \geq n+1$ in this subsection.

To this end, let

$$\boldsymbol{V}_h^* := \left\{ \boldsymbol{v} \in \boldsymbol{L}^2(\Omega;\mathbb{R}^n) : \boldsymbol{v}|_K \in \boldsymbol{P}_{k+1}(K;\mathbb{R}^n) \quad \forall\, K \in \mathcal{T}_h \right\}.$$

Then a postprocessed displacement can be defined as follows: Find $\boldsymbol{u}_h^* \in \boldsymbol{V}_h^*$ such that

$$\boldsymbol{Q}_h \boldsymbol{u}_h^* = \boldsymbol{u}_h, \tag{3.29}$$

$$(\boldsymbol{\varepsilon}(\boldsymbol{u}_h^*), \boldsymbol{\varepsilon}(\boldsymbol{v}))_K = (\mathfrak{A}\boldsymbol{\sigma}_h, \boldsymbol{\varepsilon}(\boldsymbol{v}))_K \quad \forall\ \boldsymbol{v} \in (\boldsymbol{I} - \boldsymbol{Q}_h)\boldsymbol{V}_h^*|_K, \tag{3.30}$$

for any $K \in \mathcal{T}_h$. The postprocessing (3.29)-(3.30) can be recast as the following local mixed method: Find $\boldsymbol{u}_h^* \in \boldsymbol{V}_h^*$ and $\boldsymbol{\phi}_h \in \boldsymbol{V}_h$ satisfying

$$\begin{aligned}(\boldsymbol{\varepsilon}(\boldsymbol{u}_h^*), \boldsymbol{\varepsilon}(\boldsymbol{v}))_K + (\boldsymbol{v}, \boldsymbol{\phi}_h)_K &= (\mathfrak{A}\boldsymbol{\sigma}_h, \boldsymbol{\varepsilon}(\boldsymbol{v}))_K \quad && \forall\, \boldsymbol{v} \in \boldsymbol{V}_h^*|_K,\\ (\boldsymbol{u}_h^*, \boldsymbol{\psi})_K &= (\boldsymbol{u}_h, \boldsymbol{\psi})_K && \forall\, \boldsymbol{\psi} \in \boldsymbol{V}_h|_K.\end{aligned}$$

To derive the error estimate for the postprocessed displacement $\boldsymbol{u}_h^*$, we will merge the mixed finite element method (2.3)-(2.4) and the postprocessing (3.29)-(3.30) into one method as in [63]. To be specific, find $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h^*) \in \boldsymbol{\Sigma}_h \times \boldsymbol{V}_h^*$ such that

$$\mathbb{A}_h(\boldsymbol{\sigma}_h, \boldsymbol{u}_h^*; \boldsymbol{\tau}_h, \boldsymbol{v}_h^*) = -(\boldsymbol{Q}_h \boldsymbol{f}, \boldsymbol{v}_h^*) \quad \forall\ (\boldsymbol{\tau}_h, \boldsymbol{v}_h^*) \in \boldsymbol{\Sigma}_h \times \boldsymbol{V}_h^*, \tag{3.31}$$

where

$$\mathbb{A}_h(\boldsymbol{\sigma}_h, \boldsymbol{u}_h^*; \boldsymbol{\tau}_h, \boldsymbol{v}_h^*) := \mathbb{A}(\boldsymbol{\sigma}_h, \boldsymbol{u}_h^*; \boldsymbol{\tau}_h, \boldsymbol{v}_h^*) + (\varepsilon_h(\boldsymbol{u}_h^*) - \mathfrak{A}\boldsymbol{\sigma}_h, \varepsilon_h(\boldsymbol{v}_h^* - \boldsymbol{Q}_h \boldsymbol{v}_h^*)).$$

Lemma 3.12. *The mixed finite element method* (2.3)-(2.4) *and the problem* (3.31) *are equivalent in the following sense: If* $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h^*) \in \boldsymbol{\Sigma}_h \times \boldsymbol{V}_h^*$ *is the solution of the problem* (3.31) *and* $\boldsymbol{u}_h = \boldsymbol{Q}_h \boldsymbol{u}_h^*$*, then* $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h) \in \boldsymbol{\Sigma}_h \times \boldsymbol{V}_h$ *solves the mixed finite element method* (2.3)-(2.4)*; Conversely, if* $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h) \in \boldsymbol{\Sigma}_h \times \boldsymbol{V}_h$ *is the solution of the mixed finite element method* (2.3)-(2.4) *and* $\boldsymbol{u}_h^* \in \boldsymbol{V}_h^*$ *is the postprocessed displacement defined by* (3.29)-(3.30)*, then* $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h^*) \in \boldsymbol{\Sigma}_h \times \boldsymbol{V}_h^*$ *solves the problem* (3.31)*.*

Proof. Taking any $(\boldsymbol{\tau}_h, \boldsymbol{v}_h) \in \boldsymbol{\Sigma}_h \times \boldsymbol{V}_h$, and noting the fact that $\boldsymbol{v}_h = \boldsymbol{Q}_h \boldsymbol{v}_h$ and $\mathbf{div}\boldsymbol{\Sigma}_h \subset \boldsymbol{V}_h$, we have

$$\begin{aligned}\mathbb{A}_h(\boldsymbol{\sigma}_h, \boldsymbol{u}_h^*; \boldsymbol{\tau}_h, \boldsymbol{v}_h) &= \mathbb{A}(\boldsymbol{\sigma}_h, \boldsymbol{u}_h^*; \boldsymbol{\tau}_h, \boldsymbol{v}_h) = a(\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) + b(\boldsymbol{\tau}_h, \boldsymbol{u}_h^*) + b(\boldsymbol{\sigma}_h, \boldsymbol{v}_h)\\ &= a(\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) + b(\boldsymbol{\tau}_h, \boldsymbol{Q}_h \boldsymbol{u}_h^*) + b(\boldsymbol{\sigma}_h, \boldsymbol{v}_h) = \mathbb{A}(\boldsymbol{\sigma}_h, \boldsymbol{Q}_h \boldsymbol{u}_h^*; \boldsymbol{\tau}_h, \boldsymbol{v}_h).\end{aligned} \tag{3.32}$$

Hence we can see from (3.32) that $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h)$ solves the mixed finite element method (2.3)-(2.4) if $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h^*)$ is the solution of problem (3.31).

Conversely, since $\mathbf{div}\boldsymbol{\Sigma}_h \subset \boldsymbol{V}_h$ and $(\boldsymbol{I} - \boldsymbol{Q}_h)^2 = \boldsymbol{I} - \boldsymbol{Q}_h$, it follows from (3.32) and (3.29) that

$$\begin{aligned}\mathbb{A}_h(\boldsymbol{\sigma}_h, \boldsymbol{u}_h^*; \boldsymbol{\tau}_h, \boldsymbol{v}_h^*) &= \mathbb{A}_h(\boldsymbol{\sigma}_h, \boldsymbol{u}_h^*; \boldsymbol{\tau}_h, \boldsymbol{Q}_h \boldsymbol{v}_h^*) + \mathbb{A}_h(\boldsymbol{\sigma}_h, \boldsymbol{u}_h^*; \boldsymbol{0}, \boldsymbol{v}_h^* - \boldsymbol{Q}_h \boldsymbol{v}_h^*)\\ &= \mathbb{A}(\boldsymbol{\sigma}_h, \boldsymbol{u}_h; \boldsymbol{\tau}_h, \boldsymbol{Q}_h \boldsymbol{v}_h^*) + (\varepsilon_h(\boldsymbol{u}_h^*) - \mathfrak{A}\boldsymbol{\sigma}_h, \varepsilon_h(\boldsymbol{v}_h^* - \boldsymbol{Q}_h \boldsymbol{v}_h^*))\\ &= -(\boldsymbol{f}, \boldsymbol{Q}_h \boldsymbol{v}_h^*) + (\varepsilon_h(\boldsymbol{u}_h^*) - \mathfrak{A}\boldsymbol{\sigma}_h, \varepsilon_h(\boldsymbol{v}_h^* - \boldsymbol{Q}_h \boldsymbol{v}_h^*)),\end{aligned} \tag{3.33}$$

which together with (3.30) means that $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h^*)$ solves problem (3.31). □

Lemma 3.13. *For any $\boldsymbol{v} \in \boldsymbol{H}^1(\mathcal{T}_h; \mathbb{R}^n)$, it holds that*

$$|\boldsymbol{v} - \boldsymbol{Q}_h \boldsymbol{v}|_{1,h} \eqsim \|\boldsymbol{\varepsilon}_h(\boldsymbol{v} - \boldsymbol{Q}_h \boldsymbol{v})\|_0. \tag{3.34}$$

Proof. It is sufficient to prove that

$$\sum_{F\in\mathcal{F}_h} h_F^{-1}\|[\boldsymbol{v} - \boldsymbol{Q}_h \boldsymbol{v}]\|_{0,F}^2 \lesssim \|\boldsymbol{\varepsilon}_h(\boldsymbol{v} - \boldsymbol{Q}_h \boldsymbol{v})\|_0^2. \tag{3.35}$$

Let $\boldsymbol{\pi}$ be defined as in Lemma 3.2 and $\boldsymbol{w} = \boldsymbol{v} - \boldsymbol{Q}_h \boldsymbol{v}$. It follows from (3.3) in [22]

$$\begin{aligned}\sum_{F\in\mathcal{F}_h} h_F^{-1}\|[\boldsymbol{w} - \boldsymbol{Q}_h \boldsymbol{w}]\|_{0,F}^2 &= \sum_{F\in\mathcal{F}_h} h_F^{-1}\|[(\boldsymbol{w} - \boldsymbol{\pi}\boldsymbol{w}) - \boldsymbol{Q}_h(\boldsymbol{w} - \boldsymbol{\pi}\boldsymbol{w})]\|_{0,F}^2 \\ &\lesssim \sum_{K\in\mathcal{T}_h} |\boldsymbol{w} - \boldsymbol{\pi}\boldsymbol{w}|_{1,K}^2 \lesssim \|\boldsymbol{\varepsilon}_h(\boldsymbol{w})\|_0^2.\end{aligned} \tag{3.36}$$

On the other hand,

$$\sum_{F\in\mathcal{F}_h} h_F^{-1}\|[\boldsymbol{v} - \boldsymbol{Q}_h \boldsymbol{v}]\|_{0,F}^2 = \sum_{F\in\mathcal{F}_h} h_F^{-1}\|[\boldsymbol{w} - \boldsymbol{Q}_h \boldsymbol{w}]\|_{0,F}^2.$$

Therefore (3.35) follows from (3.36). $\square$

Theorem 3.14. *For any $(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{u}}_h^*) \in \hat{\boldsymbol{\Sigma}}_h \times \boldsymbol{V}_h^*$, it follows that*

$$\|\widetilde{\boldsymbol{\sigma}}_h\|_{0,h} + |\widetilde{\boldsymbol{u}}_h^*|_{1,h} \lesssim \sup_{(\boldsymbol{\tau}_h, \boldsymbol{v}_h^*)\in\hat{\boldsymbol{\Sigma}}_h\times\boldsymbol{V}_h^*} \frac{\mathbb{A}_h(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{u}}_h^*; \boldsymbol{\tau}_h, \boldsymbol{v}_h^*)}{\|\boldsymbol{\tau}_h\|_{0,h} + |\boldsymbol{v}_h^*|_{1,h}}. \tag{3.37}$$

Proof. For any $\boldsymbol{v}_h \in \boldsymbol{V}_h$, we have from (3.32) that

$$\mathbb{A}_h(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{u}}_h^*; \boldsymbol{\tau}_h, \boldsymbol{v}_h) = \mathbb{A}(\widetilde{\boldsymbol{\sigma}}_h, \boldsymbol{Q}_h\widetilde{\boldsymbol{u}}_h^*; \boldsymbol{\tau}_h, \boldsymbol{v}_h).$$

Since $(\widetilde{\boldsymbol{\sigma}}_h, \boldsymbol{Q}_h\widetilde{\boldsymbol{u}}_h^*) \in \hat{\boldsymbol{\Sigma}}_h \times \boldsymbol{V}_h$, it holds from (3.17) that

$$\begin{aligned}\|\widetilde{\boldsymbol{\sigma}}_h\|_{0,h} + |\boldsymbol{Q}_h\widetilde{\boldsymbol{u}}_h^*|_{1,h} &\lesssim \sup_{(\boldsymbol{\tau}_h, \boldsymbol{v}_h)\in\hat{\boldsymbol{\Sigma}}_h\times\boldsymbol{V}_h} \frac{\mathbb{A}(\widetilde{\boldsymbol{\sigma}}_h, \boldsymbol{Q}_h\widetilde{\boldsymbol{u}}_h^*; \boldsymbol{\tau}_h, \boldsymbol{v}_h)}{\|\boldsymbol{\tau}_h\|_{0,h} + |\boldsymbol{v}_h|_{1,h}} \\ &= \sup_{(\boldsymbol{\tau}_h, \boldsymbol{v}_h)\in\hat{\boldsymbol{\Sigma}}_h\times\boldsymbol{V}_h} \frac{\mathbb{A}_h(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{u}}_h^*; \boldsymbol{\tau}_h, \boldsymbol{v}_h)}{\|\boldsymbol{\tau}_h\|_{0,h} + |\boldsymbol{v}_h|_{1,h}} \\ &\le \sup_{(\boldsymbol{\tau}_h, \boldsymbol{v}_h^*)\in\hat{\boldsymbol{\Sigma}}_h\times\boldsymbol{V}_h^*} \frac{\mathbb{A}_h(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{u}}_h^*; \boldsymbol{\tau}_h, \boldsymbol{v}_h^*)}{\|\boldsymbol{\tau}_h\|_{0,h} + |\boldsymbol{v}_h^*|_{1,h}}.\end{aligned} \tag{3.38}$$

If $\boldsymbol{\varepsilon}_h(\widetilde{\boldsymbol{u}}_h^* - \boldsymbol{Q}_h\widetilde{\boldsymbol{u}}_h^*) = \boldsymbol{0}$, then (3.37) is the immediate result of the triangle inequality, (3.38) and (3.34). Next we only need to focus on $\boldsymbol{\varepsilon}_h(\widetilde{\boldsymbol{u}}_h^* - \boldsymbol{Q}_h\widetilde{\boldsymbol{u}}_h^*) \neq \boldsymbol{0}$. Similarly, as in (3.33), we get

$$\mathbb{A}_h(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{u}}_h^*; \boldsymbol{0}, \widetilde{\boldsymbol{u}}_h^* - \boldsymbol{Q}_h\widetilde{\boldsymbol{u}}_h^*) = (\boldsymbol{\varepsilon}_h(\widetilde{\boldsymbol{u}}_h^*) - \mathfrak{A}\widetilde{\boldsymbol{\sigma}}_h, \boldsymbol{\varepsilon}_h(\widetilde{\boldsymbol{u}}_h^* - \boldsymbol{Q}_h\widetilde{\boldsymbol{u}}_h^*)).$$

Then we rewrite it as

$$\begin{aligned}\|\boldsymbol{\varepsilon}_h(\widetilde{\boldsymbol{u}}_h^* - \boldsymbol{Q}_h\widetilde{\boldsymbol{u}}_h^*)\|_0^2 =&(\mathfrak{A}\widetilde{\boldsymbol{\sigma}}_h - \boldsymbol{\varepsilon}_h(\boldsymbol{Q}_h\widetilde{\boldsymbol{u}}_h^*), \boldsymbol{\varepsilon}_h(\widetilde{\boldsymbol{u}}_h^* - \boldsymbol{Q}_h\widetilde{\boldsymbol{u}}_h^*)) \\ &+ \mathbb{A}_h(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{u}}_h^*; \boldsymbol{0}, \widetilde{\boldsymbol{u}}_h^* - \boldsymbol{Q}_h\widetilde{\boldsymbol{u}}_h^*).\end{aligned} \tag{3.39}$$

By the triangle inequality and (3.38), it holds that

$$\begin{aligned}\|\mathfrak{A}\widetilde{\boldsymbol{\sigma}}_h - \boldsymbol{\varepsilon}_h(\boldsymbol{Q}_h\widetilde{\boldsymbol{u}}_h^*)\|_0 \le&\|\mathfrak{A}\widetilde{\boldsymbol{\sigma}}_h\|_0 + \|\boldsymbol{\varepsilon}_h(\boldsymbol{Q}_h\widetilde{\boldsymbol{u}}_h^*)\|_0 \lesssim \|\widetilde{\boldsymbol{\sigma}}_h\|_0 + |\boldsymbol{Q}_h\widetilde{\boldsymbol{u}}_h^*|_{1,h} \\ &\lesssim \sup_{(\boldsymbol{\tau}_h, \boldsymbol{v}_h^*)\in\hat{\boldsymbol{\Sigma}}_h\times\boldsymbol{V}_h^*} \frac{\mathbb{A}_h(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{u}}_h^*; \boldsymbol{\tau}_h, \boldsymbol{v}_h^*)}{\|\boldsymbol{\tau}_h\|_{0,h} + |\boldsymbol{v}_h^*|_{1,h}}.\end{aligned}$$

Due to (3.34), we have

$$\begin{aligned}
&\mathbb{A}_h(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{u}}_h^*; \mathbf{0}, \widetilde{\boldsymbol{u}}_h^* - \boldsymbol{Q}_h\widetilde{\boldsymbol{u}}_h^*) \\
&\quad\le \|\boldsymbol{\varepsilon}_h(\widetilde{\boldsymbol{u}}_h^* - \boldsymbol{Q}_h\widetilde{\boldsymbol{u}}_h^*)\|_0 \sup_{\boldsymbol{v}_h^*\in\boldsymbol{V}_h^*,\, \boldsymbol{\varepsilon}_h(\boldsymbol{v}_h^*-\boldsymbol{Q}_h\boldsymbol{v}_h^*)\neq\mathbf{0}} \frac{\mathbb{A}_h(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{u}}_h^*; \mathbf{0}, \boldsymbol{v}_h^* - \boldsymbol{Q}_h\boldsymbol{v}_h^*)}{\|\boldsymbol{\varepsilon}_h(\boldsymbol{v}_h^* - \boldsymbol{Q}_h\boldsymbol{v}_h^*)\|_0} \\
&\quad\lesssim \|\boldsymbol{\varepsilon}_h(\widetilde{\boldsymbol{u}}_h^* - \boldsymbol{Q}_h\widetilde{\boldsymbol{u}}_h^*)\|_0 \sup_{\boldsymbol{v}_h^*\in\boldsymbol{V}_h^*,\, \boldsymbol{\varepsilon}_h(\boldsymbol{v}_h^*-\boldsymbol{Q}_h\boldsymbol{v}_h^*)\neq\mathbf{0}} \frac{\mathbb{A}_h(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{u}}_h^*; \mathbf{0}, \boldsymbol{v}_h^* - \boldsymbol{Q}_h\boldsymbol{v}_h^*)}{|\boldsymbol{v}_h^* - \boldsymbol{Q}_h\boldsymbol{v}_h^*|_{1,h}} \\
&\quad\lesssim \|\boldsymbol{\varepsilon}_h(\widetilde{\boldsymbol{u}}_h^* - \boldsymbol{Q}_h\widetilde{\boldsymbol{u}}_h^*)\|_0 \sup_{(\boldsymbol{\tau}_h,\boldsymbol{v}_h^*)\in\hat{\boldsymbol{\Sigma}}_h\times\boldsymbol{V}_h^*} \frac{\mathbb{A}_h(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{u}}_h^*; \boldsymbol{\tau}_h, \boldsymbol{v}_h^*)}{\|\boldsymbol{\tau}_h\|_{0,h} + |\boldsymbol{v}_h^*|_{1,h}}.
\end{aligned}$$

Using the last two inequalities and the Cauchy-Schwarz inequality, we get from (3.39) that

$$\|\boldsymbol{\varepsilon}_h(\widetilde{\boldsymbol{u}}_h^* - \boldsymbol{Q}_h\widetilde{\boldsymbol{u}}_h^*)\|_0 \lesssim \sup_{(\boldsymbol{\tau}_h,\boldsymbol{v}_h^*)\in\hat{\boldsymbol{\Sigma}}_h\times\boldsymbol{V}_h^*} \frac{\mathbb{A}_h(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{u}}_h^*; \boldsymbol{\tau}_h, \boldsymbol{v}_h^*)}{\|\boldsymbol{\tau}_h\|_{0,h} + |\boldsymbol{v}_h^*|_{1,h}},$$

which together with (3.34) implies

$$|\widetilde{\boldsymbol{u}}_h^* - \boldsymbol{Q}_h\widetilde{\boldsymbol{u}}_h^*|_{1,h} \lesssim \sup_{(\boldsymbol{\tau}_h,\boldsymbol{v}_h^*)\in\hat{\boldsymbol{\Sigma}}_h\times\boldsymbol{V}_h^*} \frac{\mathbb{A}_h(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{u}}_h^*; \boldsymbol{\tau}_h, \boldsymbol{v}_h^*)}{\|\boldsymbol{\tau}_h\|_{0,h} + |\boldsymbol{v}_h^*|_{1,h}}. \tag{3.40}$$

Finally, we can finish the proof by combining (3.38) and (3.40). □

Theorem 3.15. *Assume that $\boldsymbol{\sigma} \in \boldsymbol{H}^{k+1}(\Omega;\mathbb{S})$ and $\boldsymbol{u} \in \boldsymbol{H}^{k+2}(\Omega;\mathbb{R}^n)$, then*

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,h} + |\boldsymbol{u} - \boldsymbol{u}_h^*|_{1,h} \lesssim h^{k+1}\left(\|\boldsymbol{\sigma}\|_{k+1} + \|\boldsymbol{u}\|_{k+2}\right). \tag{3.41}$$

Moreover, when Ω is convex, we have

$$\|\boldsymbol{u} - \boldsymbol{u}_h^*\|_0 \lesssim h^{k+2}\left(\|\boldsymbol{\sigma}\|_{k+1} + \|\boldsymbol{u}\|_{k+2}\right). \tag{3.42}$$

Proof. By direct computation, we have

$$\mathbb{A}_h(\boldsymbol{\sigma}, \boldsymbol{u}; \boldsymbol{\tau}_h, \boldsymbol{v}_h^*) = -(\boldsymbol{f}, \boldsymbol{v}_h^*) \quad \forall\ (\boldsymbol{\tau}_h, \boldsymbol{v}_h^*) \in \boldsymbol{\Sigma}_h \times \boldsymbol{V}_h^*.$$

Combining with (3.31), we get the error equation

$$\mathbb{A}_h(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{u} - \boldsymbol{u}_h^*; \boldsymbol{\tau}_h, \boldsymbol{v}_h^*) = (\boldsymbol{Q}_h\boldsymbol{f} - \boldsymbol{f}, \boldsymbol{v}_h^*) \quad \forall\ (\boldsymbol{\tau}_h, \boldsymbol{v}_h^*) \in \boldsymbol{\Sigma}_h \times \boldsymbol{V}_h^*. \tag{3.43}$$

Let $\boldsymbol{Q}_h^*$ be the L^2 orthogonal projection from $\boldsymbol{L}^2(\Omega;\mathbb{R}^n)$ onto $\boldsymbol{V}_h^*$. It holds from (3.23) that

$$\begin{aligned}
&\mathbb{A}_h(\boldsymbol{\sigma} - \boldsymbol{I}_h^{HZ}\boldsymbol{\sigma}, \boldsymbol{u} - \boldsymbol{Q}_h^*\boldsymbol{u}; \boldsymbol{\tau}_h, \boldsymbol{v}_h^*) \\
&\quad = a(\boldsymbol{\sigma} - \boldsymbol{I}_h^{HZ}\boldsymbol{\sigma}, \boldsymbol{\tau}_h) + b(\boldsymbol{\sigma} - \boldsymbol{I}_h^{HZ}\boldsymbol{\sigma}, \boldsymbol{v}_h^*) \\
&\qquad + (\boldsymbol{\varepsilon}_h(\boldsymbol{u} - \boldsymbol{Q}_h^*\boldsymbol{u}) - \mathfrak{A}(\boldsymbol{\sigma} - \boldsymbol{I}_h^{HZ}\boldsymbol{\sigma}), \boldsymbol{\varepsilon}_h(\boldsymbol{v}_h^* - \boldsymbol{Q}_h\boldsymbol{v}_h^*)) \\
&\quad = a(\boldsymbol{\sigma} - \boldsymbol{I}_h^{HZ}\boldsymbol{\sigma}, \boldsymbol{\tau}_h) + (\boldsymbol{Q}_h\boldsymbol{f} - \boldsymbol{f}, \boldsymbol{v}_h^*) \\
&\qquad + (\boldsymbol{\varepsilon}_h(\boldsymbol{u} - \boldsymbol{Q}_h^*\boldsymbol{u}) - \mathfrak{A}(\boldsymbol{\sigma} - \boldsymbol{I}_h^{HZ}\boldsymbol{\sigma}), \boldsymbol{\varepsilon}_h(\boldsymbol{v}_h^* - \boldsymbol{Q}_h\boldsymbol{v}_h^*)).
\end{aligned}$$

Then we obtain from (3.43), Cauchy-Schwarz inequality and the error estimates of $\boldsymbol{I}_h^{HZ}, \boldsymbol{Q}_h^*$ and $\boldsymbol{Q}_h$ that

$$\begin{aligned}
&\mathbb{A}_h(\boldsymbol{I}_h^{HZ}\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{Q}_h^*\boldsymbol{u} - \boldsymbol{u}_h^*; \boldsymbol{\tau}_h, \boldsymbol{v}_h^*) \\
&\quad = a(\boldsymbol{I}_h^{HZ}\boldsymbol{\sigma} - \boldsymbol{\sigma}, \boldsymbol{\tau}_h) - (\boldsymbol{\varepsilon}_h(\boldsymbol{u} - \boldsymbol{Q}_h^*\boldsymbol{u}) - \mathfrak{A}(\boldsymbol{\sigma} - \boldsymbol{I}_h^{HZ}\boldsymbol{\sigma}), \boldsymbol{\varepsilon}_h(\boldsymbol{v}_h^* - \boldsymbol{Q}_h\boldsymbol{v}_h^*)) \\
&\quad \lesssim h^{k+1}\left(\|\boldsymbol{\sigma}\|_{k+1} + \|\boldsymbol{u}\|_{k+2}\right)(\|\boldsymbol{\tau}_h\|_0 + \|\boldsymbol{\varepsilon}_h(\boldsymbol{v}_h^*)\|_0).
\end{aligned}$$

Here we indeed have used the following estimate derived from the inverse inequality and (3.3)-(3.4) in [22]

$$\|\boldsymbol{\varepsilon}_h(\boldsymbol{v}_h^* - \boldsymbol{Q}_h\boldsymbol{v}_h^*)\|_{0,K} \lesssim h_K^{-1}\|\boldsymbol{v}_h^* - \boldsymbol{Q}_h\boldsymbol{v}_h^*\|_{0,K} \lesssim h_K^{-1}\|\boldsymbol{v}_h^* - \boldsymbol{\pi}_K\boldsymbol{v}_h^*\|_{0,K} \lesssim \|\boldsymbol{\varepsilon}_h(\boldsymbol{v}_h^*)\|_{0,K}$$

for any $K \in \mathcal{T}_h$. Applying the inf-sup condition (3.37), it follows that

$$\|\boldsymbol{I}_h^{HZ}\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,h} + |\boldsymbol{Q}_h^*\boldsymbol{u} - \boldsymbol{u}_h^*|_{1,h} \lesssim h^{k+1}\left(\|\boldsymbol{\sigma}\|_{k+1} + \|\boldsymbol{u}\|_{k+2}\right).$$

Hence we will achieve (3.41) by using the triangle inequality, and the error estimates of $\boldsymbol{I}_h^{HZ}$ and $\boldsymbol{Q}_h^*$.

When Ω is convex, we have from the triangle inequality, the error estimate of $\boldsymbol{Q}_h$ and (3.29) that

$$\begin{aligned}\|\boldsymbol{u} - \boldsymbol{u}_h^*\|_0 \le& \|(\boldsymbol{I} - \boldsymbol{Q}_h)(\boldsymbol{u} - \boldsymbol{u}_h^*)\|_0 + \|\boldsymbol{Q}_h\boldsymbol{u} - \boldsymbol{Q}_h\boldsymbol{u}_h^*\|_0\\ \lesssim& h|\boldsymbol{u} - \boldsymbol{u}_h^*|_{1,h} + \|\boldsymbol{Q}_h\boldsymbol{u} - \boldsymbol{u}_h\|_0.\end{aligned}$$

Finally (3.42) is achieved by using (3.41) and (3.20). □

Remark 3.16. The convergence rate of $|\boldsymbol{u} - \boldsymbol{u}_h^*|_{1,h}$ is of second order higher than that of $|\boldsymbol{u} - \boldsymbol{u}_h|_{1,h}$. This superconvergent postprocessed displacement $\boldsymbol{u}_h^*$ has been used in [31] to construct an a posteriori error estimator involving the displacement.

4. Block-diagonal and block-factorization preconditioners

Direct use of the mesh-dependent norm $\|\cdot\|_{0,h}\times|\cdot|_{1,h}$ would require the additional assembling of the jump term. In this section, we first derive equivalent matrix forms for these mesh-dependent norms and then construct block-diagonal and an approximate block-factorization preconditioners.

4.1. **Equivalent matrix forms of the mesh-dependent norms.**

By the trace theorem and the inverse inequality, it is easy to see that

$$\|\boldsymbol{\tau}_h\|_{0,h} \eqsim \|\boldsymbol{\tau}_h\|_0 \quad \forall\ \boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h, \tag{4.1}$$

which implies that we can use the weighted mass matrix M_h^λ with $\lambda = 0$, i.e., M_h.

For each $\boldsymbol{v}_h \in \boldsymbol{V}_h$, denote by $\underline{\boldsymbol{v_h}}$ the vector representation of $\boldsymbol{v}_h$ based on the basis of $\boldsymbol{V}_h$ used to form the mass matrix $M_{u,h}$ (cf. [78, Subsection 4.4]). For the mesh-dependent norm $|\cdot|_{1,h}$ of the displacement, we can use the Schur complement of the $(1,1)$ block, i.e., $S_h := B_hM_h^{-1}B_h^T + C_h$. It is easy to see S_h is symmetric and positive definite (SPD) and induces a norm $\|\cdot\|_{S_h}$ on $\boldsymbol{V}_h$, i.e.,

$$\|\boldsymbol{v}_h\|_{S_h}^2 := \underline{\boldsymbol{v_h}}^T S_h \underline{\boldsymbol{v_h}} \quad \forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h.$$

Lemma 4.1. *We have the norm equivalence:*

$$|\boldsymbol{v}_h|_{1,h} \eqsim \|\boldsymbol{v}_h\|_{S_h} \quad \forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h.$$

Proof. We focus on the case $k \ge n+1$ first. The low order case $1 \le k \le n$ can be proved similarly by adding the stabilization term.

The inf-sup condition (3.1) implies B_h^T is injective and thus S_h is SPD and defines an inner product on $\boldsymbol{V}_h$. The identity

$$(\underline{\boldsymbol{v_h}}^T S_h \underline{\boldsymbol{v_h}})^{1/2} = \|M_h^{-1/2}B_h^T\underline{\boldsymbol{v_h}}\| = \sup_{\boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h} \frac{b(\boldsymbol{\tau}_h, \boldsymbol{v}_h)}{\|\boldsymbol{\tau}_h\|_0} \quad \forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h \tag{4.2}$$

follows from the Riesz representation. Here $\|\cdot\|$ denotes the Euclidean norm of a vector. The inequality $|\boldsymbol{v}_h|_{1,h} \lesssim \|\boldsymbol{v}_h\|_{S_h}$ is a combination of (3.1), (4.1), and (4.2).

From an integration by parts, we can easily get $b(\boldsymbol{\tau}_h, \boldsymbol{v}_h) \lesssim \|\boldsymbol{\tau}_h\|_{0,h} |\boldsymbol{v}_h|_{1,h}$. Then the inequality $\|\boldsymbol{v}_h\|_{S_h} \lesssim |\boldsymbol{v}_h|_{1,h}$ follows from (4.1) and (4.2). □

We define the operator $\mathcal{P}_h : \boldsymbol{\Sigma}_h' \times \boldsymbol{V}_h' \to \boldsymbol{\Sigma}_h \times \boldsymbol{V}_h$ with the matrix representation

$$\mathcal{P}_h := \begin{pmatrix} M_h^{-1} & 0 \\ 0 & S_h^{-1} \end{pmatrix}, \tag{4.3}$$

and denote

$$\mathcal{L}_h^\lambda := \begin{pmatrix} M_h^\lambda & B_h^T \\ B_h & -C_h \end{pmatrix}.$$

By the stability results (3.17) and (3.24), we have the following result [15, 64, 67].

Theorem 4.2. *The $\mathcal{P}_h$ is a uniform preconditioner for $\mathcal{L}_h^\lambda$, i.e., the corresponding operator norms*

$$\|\mathcal{P}_h \mathcal{L}_h^\lambda\|_{\boldsymbol{\Sigma}_h \times \boldsymbol{V}_h \to \boldsymbol{\Sigma}_h \times \boldsymbol{V}_h}, \|(\mathcal{P}_h \mathcal{L}_h^\lambda)^{-1}\|_{\boldsymbol{\Sigma}_h \times \boldsymbol{V}_h \to \boldsymbol{\Sigma}_h \times \boldsymbol{V}_h}$$

are bounded and independent of parameters h and λ.

The inverse of the mass matrix M_h^{-1} can be further replaced by the inverse of the diagonal matrix or symmetric Gauss-Seidel iteration and thus the computation of M_h^{-1} is not a problem. The difficulty is the inverse of the Schur complement which will be further preconditioned by an auxiliary space preconditioner in the next section.

4.2. **Approximate block-factorization preconditioner.** Let $\tilde{M}_h$ be a symmetric and positive definite matrix which is spectrally equivalent to M_h. We can make use of the block decomposition

$$\begin{pmatrix} \tilde{M}_h & B_h^T \\ B_h & -C_h \end{pmatrix} \begin{pmatrix} I & \tilde{M}_h^{-1} B_h^T \\ 0 & -I \end{pmatrix} = \begin{pmatrix} \tilde{M}_h & 0 \\ B_h & \tilde{S}_h \end{pmatrix} \tag{4.4}$$

to obtain an approximate block-factorization preconditioner, where

$$\tilde{S}_h := B_h \tilde{M}_h^{-1} B_h^T + C_h.$$

It is apparent that $\tilde{S}_h$ is spectrally equivalent to S_h.

We define the operator $\mathcal{G}_h : \boldsymbol{\Sigma}_h' \times \boldsymbol{V}_h' \to \boldsymbol{\Sigma}_h \times \boldsymbol{V}_h$ as

$$\mathcal{G}_h := \begin{pmatrix} I & \tilde{M}_h^{-1} B_h^T \\ 0 & -I \end{pmatrix} \begin{pmatrix} \tilde{M}_h & 0 \\ B_h & \tilde{S}_h \end{pmatrix}^{-1}. \tag{4.5}$$

If we denote

$$\tilde{\mathcal{L}}_h := \begin{pmatrix} \tilde{M}_h & B_h^T \\ B_h & -C_h \end{pmatrix},$$

it is trivial to verify that $\mathcal{G}_h = \tilde{\mathcal{L}}_h^{-1}$. Sine $\tilde{M}_h$ is spectrally equivalent to M_h, so $\tilde{\mathcal{L}}_h$ is also stable in the mesh-dependent norm. We thus obtain the following result [9, 15, 64, 67]. Detailed eigenvalue analysis of the preconditioned system can be found in [13].

Theorem 4.3. *The $\mathcal{G}_h$ is a uniform preconditioner for $\mathcal{L}_h^\lambda$, i.e., the corresponding operator norms*

$$\|\mathcal{G}_h \mathcal{L}_h^\lambda\|_{\boldsymbol{\Sigma}_h \times \boldsymbol{V}_h \to \boldsymbol{\Sigma}_h \times \boldsymbol{V}_h}, \|(\mathcal{G}_h \mathcal{L}_h^\lambda)^{-1}\|_{\boldsymbol{\Sigma}_h \times \boldsymbol{V}_h \to \boldsymbol{\Sigma}_h \times \boldsymbol{V}_h}$$

are bounded and independent of parameters h and λ.

Lemma 4.4. *All the eigenvalues of $\mathcal{G}_h \mathcal{L}_h^\lambda$ are positive.*

Proof. This claim has been proved in [37, 51]. Here we reprove it by a direct computation. Suppose ζ is an eigenvalue of $\mathcal{G}_h\mathcal{L}_h^\lambda$, then there exists nonzero complex vector (y, z) such that

$$\begin{pmatrix} M_h^\lambda & B_h^T \\ B_h & -C_h \end{pmatrix}\begin{pmatrix} y \\ z \end{pmatrix} = \zeta \begin{pmatrix} \tilde{M}_h & B_h^T \\ B_h & -C_h \end{pmatrix}\begin{pmatrix} y \\ z \end{pmatrix}. \tag{4.6}$$

Multiplying $(0, \bar{z}^T)$ on the left of (4.6), where $\bar{z}$ is the complex conjugate of z, we get

$$\bar{z}^T B_h y - \bar{z}^T C_h z = \zeta(\bar{z}^T B_h y - \bar{z}^T C_h z).$$

If $\zeta = 1$, the proof is finished. Then we only consider $\zeta \neq 1$, which indicates

$$\bar{z}^T B_h y = \bar{z}^T C_h z.$$

Thus

$$\bar{y}^T B_h^T z = \overline{\bar{z}^T B_h y} = \overline{\bar{z}^T C_h z} = \bar{z}^T C_h z.$$

Multiplying $(\bar{y}^T, 0)$ on the left of (4.6), we have

$$\bar{y}^T M_h^\lambda y + \bar{y}^T B_h^T z = \zeta(\bar{y}^T \tilde{M}_h y + \bar{y}^T B_h^T z).$$

Then we get from the last two equalities that

$$\bar{y}^T M_h^\lambda y + \bar{z}^T C_h z = \zeta(\bar{y}^T \tilde{M}_h y + \bar{z}^T C_h z). \tag{4.7}$$

Next we show that $\bar{y}^T \tilde{M}_h y + \bar{z}^T C_h z \neq 0$. Otherwise, we obtain $y = 0$ and $C_h z = 0$. By (4.6) and the fact that $\zeta \neq 1$, we get $B_h^T z = 0$. Hence it holds that

$$\begin{pmatrix} M_h^\lambda & B_h^T \\ B_h & -C_h \end{pmatrix}\begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

which indicates the vector $(y, z) = 0$. Thus $\bar{y}^T \tilde{M}_h y + \bar{z}^T C_h z \neq 0$. Therefore we get from (4.7) that ζ is positive. $\square$

Combining Theorem 4.3 and Lemma 4.4, we have the following result.

Theorem 4.5. *It holds for each eigenvalue ζ of $\mathcal{G}_h\mathcal{L}_h^\lambda$ that*

$$\zeta \eqsim 1.$$

In both diagonal- and block-factorization preconditioners, to be practical, we do not compute S_h^{-1} or $\tilde{S}_h^{-1}$. Instead it can be solved by many existing fast solvers, such as multigrid methods [11, 12, 23, 24, 28, 42], domain decomposition methods [2, 3, 14, 41], and other two-level or multilevel preconditioning techniques [11, 12, 29, 36, 57, 58, 82]. We shall apply the fast auxiliary space preconditioner to be developed in the next section.

At the end of this section, we show the convergence of the preconditioned GMRES method. For ease of presentation, for any two symmetric matrices H_1 and H_2, let $H_1 \leq H_2$ mean $H_2 - H_1$ is a semi-positive definite matrix.

Lemma 4.6. *For any $\boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h$ satisfying $\int_\Omega \mathrm{tr}\boldsymbol{\tau}_h \,\mathrm{d}x = 0$, we have*

$$\|\boldsymbol{\tau}_h\|_0 \lesssim \|\boldsymbol{\tau}_h\|_a + \|\operatorname{div}\boldsymbol{\tau}_h\|_{-1,h}.$$

Proof. First we can see that

$$\|h \operatorname{div} \boldsymbol{\tau}_h\| = \sup_{\boldsymbol{v}_h \in \boldsymbol{V}_h} \frac{\sum\limits_{K \in \mathcal{T}_h} (h_K \operatorname{div} \boldsymbol{\tau}_h, \boldsymbol{v}_h)_K}{\|\boldsymbol{v}_h\|_0} = \sup_{\boldsymbol{v}_h \in \boldsymbol{V}_h} \frac{(\operatorname{div} \boldsymbol{\tau}_h, \boldsymbol{v}_h)}{\left(\sum\limits_{K \in \mathcal{T}_h} \|h_K^{-1} \boldsymbol{v}_h\|_{0,K}^2\right)^{1/2}}.$$

By the inverse inequality, it follows that

$$|\boldsymbol{v}_h|_{1,h}^2 \lesssim \sum_{K \in \mathcal{T}_h} \|h_K^{-1} \boldsymbol{v}_h\|_{0,K}^2.$$

Hence

$$\|h \operatorname{div} \boldsymbol{\tau}_h\| \lesssim \sup_{\boldsymbol{v}_h \in \boldsymbol{V}_h} \frac{(\operatorname{div} \boldsymbol{\tau}_h, \boldsymbol{v}_h)}{|\boldsymbol{v}_h|_{1,h}} = \|\operatorname{div} \boldsymbol{\tau}_h\|_{-1,h},$$

which together with Lemma 3.5 will end the proof. $\square$

By Lemma 4.1 and Lemma 4.6, there exists a constant $C_1 > 0$ such that

$$C_1 M_h \le M_h^\lambda + B_h^T S_h^{-1} B_h. \tag{4.8}$$

Let S_{Fasp} be a symmetric and positive definite matrix which is spectrally equivalent to $\tilde{S}_h$, and

$$\hat{\mathcal{G}}_h := \begin{pmatrix} I & \tilde{M}_h^{-1} B_h^T \\ 0 & -I \end{pmatrix} \begin{pmatrix} \tilde{M}_h & 0 \\ B_h & S_{Fasp} \end{pmatrix}^{-1}.$$

Assume there exist positive constants $\alpha \le \alpha' < 1$ and $\beta' \le \beta < 4$ such that

$$(1-\alpha')\tilde{M}_h \le M_h \le (1-\alpha)\tilde{M}_h, \quad \beta' S_{Fasp} \le \tilde{S}_h \le \beta S_{Fasp}. \tag{4.9}$$

Assumption (4.9) is easily satisfied by scaling $\tilde{M}_h$ and S_{Fasp} with constants independent of h. Indeed, for any given matrix $\hat{M}_h$ being spectrally equivalent to M_h and matrix $\hat{S}_h$ being spectrally equivalent to $\tilde{S}_h$, take $\tilde{M}_h = \omega_1 \hat{M}_h$ and $S_{Fasp} = \omega_2 \hat{S}_h$ with constants ω_1 and ω_2 being independent of h, then (4.9) can be verified only if constants ω_1 and ω_2 are chosen large enough.

Obviously

$$\alpha \tilde{M}_h \le \tilde{M}_h - M_h^\lambda \le \tilde{M}_h, \tag{4.10}$$

$$(1-\alpha') S_h \le \tilde{S}_h \le S_h. \tag{4.11}$$

Let $E_M := I - \tilde{M}_h^{-1} M_h^\lambda$, then

$$\alpha M_h^\lambda \le E_M^T M_h^\lambda \le M_h^\lambda, \quad E_M^T M_h^\lambda E_M \le M_h^\lambda, \tag{4.12}$$

$$\alpha^2 \tilde{M}_h \le E_M^T \tilde{M}_h E_M \le \tilde{M}_h, \quad \alpha^2 \tilde{M}_h^{-1} \le E_M \tilde{M}_h^{-1} E_M^T \le \tilde{M}_h^{-1}. \tag{4.13}$$

By direct computation, we get

$$\hat{\mathcal{G}}_h \mathcal{L}_h^\lambda = \mathcal{D}_h^{-1} \begin{pmatrix} E_M^T M_h^\lambda + E_M^T B_h^T S_{Fasp}^{-1} B_h E_M & E_M^T B_h^T (I - S_{Fasp}^{-1} \tilde{S}_h) \\ -B_h E_M & \tilde{S}_h \end{pmatrix}.$$

Let

$$\mathcal{D}_h := \begin{pmatrix} \tilde{M}_h - M_h^\lambda & 0 \\ 0 & S_{Fasp} \end{pmatrix}.$$

Lemma 4.7 (Field-of-value equivalence)**.** *Assume* (4.9) *holds, then we have*

$$\frac{(\hat{\mathcal{G}}_h\mathcal{L}_h^\lambda x,x)_{\mathcal{D}_h}}{(x,x)_{\mathcal{D}_h}}\geq\gamma\quad\text{and}\quad\frac{\|\hat{\mathcal{G}}_h\mathcal{L}_h^\lambda x\|_{\mathcal{D}_h}^2}{\|x\|_{\mathcal{D}_h}^2}\leq\Gamma$$

for all nonzero vectors $x^T:=(y^T,z^T)$*, where*

$$\gamma:=\min\left\{C_1\alpha^3(1-\alpha'),\ C_1\alpha^2\beta'(1-\alpha')(1-\frac{\sqrt{\beta}}{2}),\ \beta'(1-\frac{\sqrt{\beta}}{2})\right\},$$

$$\Gamma:=\max\left\{\frac{3}{\alpha^2}((1-\alpha)^2+\beta^2)+\frac{2\beta}{\alpha},\ \frac{3\beta}{\alpha}(1-\beta)^2+2\beta^2,\ \frac{3\beta}{\alpha}(1-\beta')^2+2\beta^2\right\}.$$

Here C_1 is the constant from (4.8).

Proof. It follows from the Cauchy-Schwarz inequality and (4.9) that

$$\begin{aligned}(B_hE_My,S_{Fasp}^{-1}\tilde{S}_hz)\leq&\|B_hE_My\|_{S_{Fasp}^{-1}}\|\tilde{S}_hz\|_{S_{Fasp}^{-1}}\leq\sqrt{\beta}\|B_hE_My\|_{S_{Fasp}^{-1}}\|z\|_{\tilde{S}_h}\\ \leq&\frac{\sqrt{\beta}}{2}(\|B_hE_My\|_{S_{Fasp}^{-1}}^2+\|z\|_{\tilde{S}_h}^2).\end{aligned}$$

Due to (4.12), it holds that

$$\begin{aligned}(\hat{\mathcal{G}}_h\mathcal{L}_h^\lambda x,x)_{\mathcal{D}_h}=&(E_M^TM_h^\lambda y,y)+\|B_hE_My\|_{S_{Fasp}^{-1}}^2+\|z\|_{\tilde{S}_h}^2-(B_hE_My,S_{Fasp}^{-1}\tilde{S}_hz)\\ \geq&\alpha(M_h^\lambda y,y)+(1-\frac{\sqrt{\beta}}{2})(\|B_hE_My\|_{S_{Fasp}^{-1}}^2+\|z\|_{\tilde{S}_h}^2).\end{aligned}\tag{4.14}$$

By (4.9) and (4.8),

$$C_1(1-\alpha')\|E_My\|_{\tilde{M}_h}^2\leq C_1\|E_My\|_{M_h}^2\leq(M_h^\lambda E_My,E_My)+\|B_hE_My\|_{S_h^{-1}}^2.$$

Combining with (4.11)-(4.13) and (4.9), we get

$$\begin{aligned}\|y\|_{\tilde{M}_h}^2\leq&\frac{1}{\alpha^2}\|E_My\|_{\tilde{M}_h}^2\leq\frac{1}{C_1\alpha^2(1-\alpha')}\left((M_h^\lambda E_My,E_My)+\|B_hE_My\|_{S_h^{-1}}^2\right)\\ \leq&\frac{1}{C_1\alpha^2(1-\alpha')}\left((M_h^\lambda y,y)+\|B_hE_My\|_{\tilde{S}_h^{-1}}^2\right)\\ \leq&\frac{1}{C_1\alpha^2(1-\alpha')}(M_h^\lambda y,y)+\frac{1}{C_1\alpha^2(1-\alpha')\beta'}\|B_hE_My\|_{S_{Fasp}^{-1}}^2.\end{aligned}$$

According to (4.9)-(4.10) and (4.14), we obtain

$$\begin{aligned}\gamma(x,x)_{\mathcal{D}_h}\leq&\gamma\|y\|_{\tilde{M}_h}^2+\gamma\|z\|_{S_{Fasp}}^2\leq\gamma\|y\|_{\tilde{M}_h}^2+\frac{\gamma}{\beta'}\|z\|_{\tilde{S}_h}^2\\ \leq&\frac{\gamma}{C_1\alpha^2(1-\alpha')}(M_h^\lambda y,y)+\frac{\gamma}{C_1\alpha^2(1-\alpha')\beta'}\|B_hE_My\|_{S_{Fasp}^{-1}}^2+\frac{\gamma}{\beta'}\|z\|_{\tilde{S}_h}^2\\ \leq&(\hat{\mathcal{G}}_h\mathcal{L}_h^\lambda x,x)_{\mathcal{D}_h}.\end{aligned}$$

On the other hand, by (4.12) and (4.9),

$$\begin{aligned}\|M_h^\lambda y\|_{\tilde{M}_h^{-1}}^2=&(M_h^\lambda y,y)-(E_M^TM_h^\lambda y,y)\leq(1-\alpha)(M_h^\lambda y,y)\\ \leq&(1-\alpha)\|y\|_{M_h}^2\leq(1-\alpha)^2\|y\|_{\tilde{M}_h}^2.\end{aligned}$$

Since $B_h\tilde{M}_h^{-1}B_h^T\tilde{S}_h^{-1} = I - C_h\tilde{S}_h^{-1}$ and $\tilde{M}_h^{-1}B_h^T\tilde{S}_h^{-1}B_h$ have the same nonzero eigenvalues, the eigenvalues of the generalized eigenvalue problem $\tilde{B}_h^T\tilde{S}_h^{-1}B_hy = \lambda M_hy$ are not bigger than one. We get from (4.9) and (4.13) that

$$\begin{aligned}\|B_h^TS_{Fasp}^{-1}B_hE_My\|^2_{\tilde{M}_h^{-1}} =& y^TE_M^TB_h^TS_{Fasp}^{-1}B_h\tilde{M}_h^{-1}B_h^TS_{Fasp}^{-1}B_hE_My\\ \le& y^TE_M^TB_h^TS_{Fasp}^{-1}\tilde{S}_hS_{Fasp}^{-1}B_hE_My\\ \le& \beta y^TE_M^TB_h^TS_{Fasp}^{-1}B_hE_My\\ \le& \beta^2y^TE_M^TB_h^T\tilde{S}_h^{-1}B_hE_My\\ \le& \beta^2y^TE_M^T\tilde{M}_hE_My \le \beta^2\|y\|^2_{\tilde{M}_h}.\end{aligned}$$

Similarly, we obtain from (4.9) and (4.13) that

$$\begin{aligned}\|B_h^T(I - S_{Fasp}^{-1}\tilde{S}_h)z\|^2_{\tilde{M}_h^{-1}} =& z^T(I - \tilde{S}_hS_{Fasp}^{-1})B_h\tilde{M}_h^{-1}B_h^T(I - S_{Fasp}^{-1}\tilde{S}_h)z\\ \le& z^T(I - \tilde{S}_hS_{Fasp}^{-1})\tilde{S}_h(I - S_{Fasp}^{-1}\tilde{S}_h)z\\ \le& \max\{(1-\beta)^2, (1-\beta')^2\}\|z\|^2_{\tilde{S}_h},\end{aligned}$$

$$\begin{aligned}\|-B_hE_My + \tilde{S}_hz\|^2_{\tilde{S}_h^{-1}} \le& 2\|-B_hE_My\|^2_{\tilde{S}_h^{-1}} + 2\|\tilde{S}_hz\|^2_{\tilde{S}_h^{-1}}\\ =& 2y^TE_M^TB_h^T\tilde{S}_h^{-1}B_hE_My + 2\|z\|^2_{\tilde{S}_h}\\ \le& 2y^TE_M^T\tilde{M}_hE_My + 2\|z\|^2_{\tilde{S}_h} \le 2\|y\|^2_{\tilde{M}_h} + 2\|z\|^2_{\tilde{S}_h}.\end{aligned}$$

Combining the last four inequalities, we achieve from (4.9)-(4.10) and (4.13) that

$$\begin{aligned}\|\hat{\mathcal{G}}_h\mathcal{L}_h^\lambda x\|^2_{\mathcal{D}_h} \le& \|M_h^\lambda y + B_h^TS_{Fasp}^{-1}B_hE_My + B_h^T(I - S_{Fasp}^{-1}\tilde{S}_h)z\|^2_{E_M(\tilde{M}_h - M_h^\lambda)^{-1}E_M^T}\\ &+ \|-B_hE_My + \tilde{S}_hz\|^2_{S_{Fasp}^{-1}}\\ \le& \frac{1}{\alpha}\|M_h^\lambda y + B_h^TS_{Fasp}^{-1}B_hE_My + B_h^T(I - S_{Fasp}^{-1}\tilde{S}_h)z\|^2_{\tilde{M}_h^{-1}}\\ &+ \beta\|-B_hE_My + \tilde{S}_hz\|^2_{\tilde{S}_h^{-1}}\\ \le& \frac{3}{\alpha}\|M_h^\lambda y\|^2_{\tilde{M}_h^{-1}} + \frac{3}{\alpha}\|B_h^TS_{Fasp}^{-1}B_hE_My\|^2_{\tilde{M}_h^{-1}}\\ &+ \frac{3}{\alpha}\|B_h^T(I - S_{Fasp}^{-1}\tilde{S}_h)z\|^2_{\tilde{M}_h^{-1}} + \beta\|-B_hE_My + \tilde{S}_hz\|^2_{\tilde{S}_h^{-1}}\\ \le& (\frac{3}{\alpha}((1-\alpha)^2 + \beta^2) + 2\beta)\|y\|^2_{\tilde{M}_h}\\ &+ (\frac{3}{\alpha}\max\{(1-\beta)^2, (1-\beta')^2\} + 2\beta)\|z\|^2_{\tilde{S}_h} \le \Gamma\|x\|^2_{\mathcal{D}_h},\end{aligned}$$

as required. □

By the convergence theories for the GMRES method developed in [40, 70] and [62, Algorithm 2.2], from Lemma 4.7 we conclude the uniform convergence of the GMRES method in the $\|\cdot\|_{\mathcal{D}_h}$-norm as follows.

Theorem 4.8. *Assume the SPD matrices $\tilde{M}_h$ and S_{Fasp} satisfy* (4.9)*, then the GMRES method with the approximate block-factorization preconditioner $\hat{\mathcal{G}}_h$ for linear system* (2.5) *converges uniformly in the $\|\cdot\|_{\mathcal{D}_h}$-norm with respect to the parameters h and λ.*

5. Auxiliary space preconditioner

In this section we first review the framework on the auxiliary space preconditioners developed by Xu [79] and then construct one for the linear elasticity problem in mixed forms. We use H^1-conforming linear element and primary formulation of linear elasticity with $\lambda = 0$ as the auxiliary space preconditioner and verify all assumptions needed in the framework.

5.1. Framework.

Let

$$\boldsymbol{\mathcal{V}}_h := \left\{\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega;\mathbb{R}^n) : \boldsymbol{v}|_K \in \boldsymbol{P}_1(K;\mathbb{R}^n) \quad \forall\, K \in \mathcal{T}_h\right\}.$$

Then $\boldsymbol{\mathcal{V}}_h \subset \boldsymbol{V}_h$ for $k \geq 2$, and

$$|\boldsymbol{v}_h|_{1,h} = \|\boldsymbol{\varepsilon}(\boldsymbol{v}_h)\|_0 \eqsim |\boldsymbol{v}_h|_1 \quad \forall\, \boldsymbol{v}_h \in \boldsymbol{\mathcal{V}}_h. \tag{5.1}$$

The conforming linear finite element method for the linear elasticity with $\lambda = 0$ is defined as follows: Find $\boldsymbol{u}_h \in \boldsymbol{\mathcal{V}}_h$ such that

$$2\mu(\boldsymbol{\varepsilon}(\boldsymbol{u}_h), \boldsymbol{\varepsilon}(\boldsymbol{v}_h)) = (\boldsymbol{f}, \boldsymbol{v}_h) \qquad \forall\, \boldsymbol{v}_h \in \boldsymbol{\mathcal{V}}_h. \tag{5.2}$$

Denote $\mathcal{A} : \boldsymbol{\mathcal{V}}_h \to \boldsymbol{\mathcal{V}}_h$ by

$$(\mathcal{A}\boldsymbol{w}_h, \boldsymbol{v}_h) := 2\mu(\boldsymbol{\varepsilon}(\boldsymbol{w}_h), \boldsymbol{\varepsilon}(\boldsymbol{v}_h)) \quad \forall\, \boldsymbol{w}_h, \boldsymbol{v}_h \in \boldsymbol{\mathcal{V}}_h.$$

It is apparent that the operator $\mathcal{A}$ is SPD.

In what follows we assume $\mathcal{T}_h$ is quasi-uniform. Define operator $M : \boldsymbol{\Sigma}_h \to \boldsymbol{\Sigma}_h$ by

$$(M\boldsymbol{\varsigma}_h, \boldsymbol{\tau}_h) = \frac{1}{2\mu}(\boldsymbol{\varsigma}_h, \boldsymbol{\tau}_h) \quad \forall\, \boldsymbol{\varsigma}_h, \boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h,$$

operator $B : \boldsymbol{\Sigma}_h \to \boldsymbol{V}_h$ by

$$(B\boldsymbol{\tau}_h, \boldsymbol{v}_h) = b(\boldsymbol{\tau}_h, \boldsymbol{v}_h) \quad \forall\, \boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h, \boldsymbol{v}_h \in \boldsymbol{v}_h,$$

and operator $C : \boldsymbol{V}_h \to \boldsymbol{V}_h$ by

$$(C\boldsymbol{w}_h, \boldsymbol{v}_h) = c(\boldsymbol{u}_h, \boldsymbol{v}_h) \quad \forall\, \boldsymbol{w}_h, \boldsymbol{v}_h \in \boldsymbol{V}_h.$$

Then the Schur complement operator $S = BM^{-1}B^T + C$. Based on the norm equivalence (5.1), we can easily derive the estimate of spectral radius and condition number of the Schur complement operator S that

$$\rho_S = \lambda_{\max}(S) \eqsim h^{-2}, \quad \kappa(S) = \frac{\lambda_{\max}(S)}{\lambda_{\min}(S)} \eqsim h^{-2}. \tag{5.3}$$

The relation between S and S_h is given by

$$S_h = M_{u,h}\underline{S}$$

with $\underline{S}$ being the matrix representation of S.

We introduce the auxiliary space preconditioner for the Schur complement. The idea is to construct a multigrid method using $\boldsymbol{V}_h$ as the "fine" space and $\boldsymbol{\mathcal{V}}_h$ as the "coarse" space. Denote $\mathcal{B} : \boldsymbol{\mathcal{V}}_h \to \boldsymbol{\mathcal{V}}_h$ to be such a "coarse" solver. It can be either an exact solver or an approximate solver that satisfies certain conditions, which will be given later. Next, on the fine space, we need a smoother $R : \boldsymbol{V}_h \to \boldsymbol{V}_h$, which is symmetric and positive definite. For example, R can be a Jacobi or symmetric Gauss-Seidel smoother. Finally, to connect the "coarse" space with the "fine" space, we need a "prolongation" operator $\Pi : \boldsymbol{\mathcal{V}}_h \to \boldsymbol{V}_h$. A "restriction" operator $\Pi^t : \boldsymbol{V}_h \to \boldsymbol{\mathcal{V}}_h$ is consequently defined by

$$(\Pi^t\boldsymbol{v}, \boldsymbol{w}) = (\boldsymbol{v}, \Pi\boldsymbol{w}) \quad \text{for } \boldsymbol{v} \in \boldsymbol{V}_h \text{ and } \boldsymbol{w} \in \boldsymbol{\mathcal{V}}_h.$$

It is also well-known that the matrix representation of the restriction operator Π^t is just the transpose of the matrix representation of the prolongation operator Π. Then, the auxiliary space preconditioner $X : \boldsymbol{V}_h \to \boldsymbol{V}_h$, following the definition in [79], is given by

$$\text{(5.4)} \qquad \text{Additive} \qquad X = R + \Pi\mathcal{B}\Pi^t,$$

$$\text{(5.5)} \qquad \text{Multiplicative} \qquad I - XS = (I - R^tS)(I - \Pi\mathcal{B}\Pi^tS)(I - RS).$$

According to [79], the following theorem holds.

Theorem 5.1 (Xu [79])**.** *Assume that for all* $\boldsymbol{v} \in \boldsymbol{V}_h$, $\boldsymbol{w} \in \boldsymbol{\mathcal{V}}_h$,

$$(S\boldsymbol{v}, \boldsymbol{v}) \lesssim (R^{-1}\boldsymbol{v}, \boldsymbol{v}) \lesssim \rho_S(\boldsymbol{v}, \boldsymbol{v}), \tag{5.6}$$

$$(\mathcal{A}\boldsymbol{w}, \boldsymbol{w}) \lesssim (\mathcal{B}\mathcal{A}\boldsymbol{w}, \mathcal{A}\boldsymbol{w}) \lesssim (\mathcal{A}\boldsymbol{w}, \boldsymbol{w}), \tag{5.7}$$

$$|\Pi\boldsymbol{w}|_{1,h} \lesssim |\boldsymbol{w}|_1 \qquad \textit{(stability of } \Pi\textit{)}, \tag{5.8}$$

and furthermore, assume that there exists a linear operator $P : \boldsymbol{V}_h \to \boldsymbol{\mathcal{V}}_h$ *such that*

$$|P\boldsymbol{v}|_1 \lesssim |\boldsymbol{v}|_{1,h} \qquad \textit{(stability of } P\textit{)}, \tag{5.9}$$

$$\|\boldsymbol{v} - \Pi P\boldsymbol{v}\|_0^2 \lesssim \rho_S^{-1}|\boldsymbol{v}|_{1,h}^2 \qquad \textit{(approximability)}. \tag{5.10}$$

Then the preconditioner X *defined in* (5.4) *or* (5.5) *satisfies*

$$\kappa(XS) \lesssim 1.$$

5.2. **Construction.** Now we construct an auxiliary space preconditioner which satisfies all conditions in Theorem 5.1, namely, inequalities (5.6)-(5.10). It is straightforward to pick $\mathcal{B}$ that satisfies condition (5.7). For example, $\mathcal{B}$ can be either the direct solver, for which $\mathcal{B} \sim \mathcal{A}^{-1}$, or one step of classical multigrid iteration which satisfies condition (5.7).

The smoother R is also easy to define. A Jacobi or a symmetric Gauss-Seidel smoother [18] will satisfy the condition (5.6). Operator Π is the natural inclusion for $k \geq 2$ and the L^2 projection $\boldsymbol{Q}_h$ for $k = 1$, i.e., taking the average of nodal values inside each simplex. Then the condition (5.8) follows from (5.1) and (3.16) immediately.

The technical part is to define an operator $P : \boldsymbol{V}_h \to \boldsymbol{\mathcal{V}}_h$ that satisfy the conditions (5.9)-(5.10). Note that the operator P is needed only in the theoretical analysis. In the implementation, one needs $\mathcal{B}$, R, and Π only.

Construction of P is equivalent to specify function values at each vertex. For an interior vertex $\boldsymbol{x}_i$ of $\mathcal{T}_h$, denoted by Ω_i the vertex patch of $\boldsymbol{x}_i$, we will simply choose $(P\boldsymbol{v})(\boldsymbol{x}_i) := |\Omega_i|^{-1}\int_{\Omega_i}\boldsymbol{v}\,\mathrm{d}x$, i.e., the average of a discontinuous polynomial $\boldsymbol{v}$ in the vertex patch. For boundary vertex $\boldsymbol{x}_i \in \partial\Omega$, we set $(P\boldsymbol{v})(\boldsymbol{x}_i) := 0$.

For any $K \in \mathcal{T}_h$, let

$$\boldsymbol{Q}_K^0\boldsymbol{v} := (\boldsymbol{Q}_h^0\boldsymbol{v})|_K = \frac{1}{|K|}\int_K \boldsymbol{v}\,\mathrm{d}x \quad \boldsymbol{v} \in \boldsymbol{L}^2(\Omega;\mathbb{R}^n).$$

Define

$$\mathcal{T}_{h,i} := \{K \in \mathcal{T}_h : K \subset \Omega_i\}, \quad \mathcal{F}_{h,i} := \{F \in \mathcal{F}_h : \boldsymbol{x}_i \in F\}.$$

Obviously for interior nodes we have (cf. [21])

$$(P\boldsymbol{v})(\boldsymbol{x}_i) = \sum_{K \in \mathcal{T}_{h,i}} \frac{|K|}{|\Omega_i|}\boldsymbol{Q}_K^0\boldsymbol{v}. \tag{5.11}$$

The error estimate of the operator P can be derived by the standard argument used in [22, 27, 50, 75]. For completeness, we show it in detail as follows.

Lemma 5.2. *The operator P satisfies*

$$\|\boldsymbol{v} - P\boldsymbol{v}\|_0 + h|P\boldsymbol{v}|_1 \lesssim h|\boldsymbol{v}|_{1,h} \quad \forall\ \boldsymbol{v} \in \boldsymbol{V}_h.$$

Proof. According to (5.11), it holds for each interior node $\boldsymbol{x}_i$ that

$$|\boldsymbol{Q}_K^0\boldsymbol{v} - (P\boldsymbol{v})(\boldsymbol{x}_i)|^2 \lesssim \sum_{K'\in\mathcal{T}_{h,i}} |\boldsymbol{Q}_K^0\boldsymbol{v} - \boldsymbol{Q}_{K'}^0\boldsymbol{v}|^2 \lesssim \sum_{F\in\mathcal{F}_{h,i}} |[\boldsymbol{Q}_h^0\boldsymbol{v}]|^2.$$

For each boundary node $\boldsymbol{x}_i$, we obtain by a similar technique and the definition of jump on the boundary,

$$|\boldsymbol{Q}_K^0\boldsymbol{v} - (P\boldsymbol{v})(\boldsymbol{x}_i)|^2 = |\boldsymbol{Q}_K^0\boldsymbol{v}|^2 \lesssim \sum_{F\in\mathcal{F}_{h,i}} |[\boldsymbol{Q}_h^0\boldsymbol{v}]|^2.$$

Then using the scaling argument, we have

$$\begin{aligned}\sum_{K\in\mathcal{T}_h} h_K^{-2}\|\boldsymbol{Q}_K^0\boldsymbol{v} - P\boldsymbol{v}\|_{0,K}^2 &= \sum_{K\in\mathcal{T}_h}\sum_{i=0}^{n} h_K^{n-2}|\boldsymbol{Q}_K^0\boldsymbol{v} - (P\boldsymbol{v})(\boldsymbol{x}_{K,i})|^2 \\ &\lesssim \sum_{F\in\mathcal{F}_h} h_F^{-1}\|[\boldsymbol{Q}_h^0\boldsymbol{v}]\|_{0,F}^2.\end{aligned}$$

From the L^2 error estimate (3.15), the discrete Korn's inequality (3.8), and the norm equivalence (3.6), we get

$$\begin{aligned}\sum_{K\in\mathcal{T}_h} h_K^{-2}\|\boldsymbol{v} - P\boldsymbol{v}\|_{0,K}^2 &\lesssim \sum_{K\in\mathcal{T}_h} h_K^{-2}\|\boldsymbol{v} - \boldsymbol{Q}_K^0\boldsymbol{v}\|_{0,K}^2 + \sum_{K\in\mathcal{T}_h} h_K^{-2}\|\boldsymbol{Q}_K^0\boldsymbol{v} - P\boldsymbol{v}\|_{0,K}^2 \\ &\lesssim |\boldsymbol{v}|_{1,h}^2 + \sum_{F\in\mathcal{F}_h} h_F^{-1}\|[\boldsymbol{Q}_h^0\boldsymbol{v}]\|_{0,F}^2 \\ &\lesssim |\boldsymbol{v}|_{1,h}^2 + \sum_{F\in\mathcal{F}_h} h_F^{-1}\|[\boldsymbol{Q}_h^0\boldsymbol{v} - \boldsymbol{v}]\|_{0,F}^2 \lesssim |\boldsymbol{v}|_{1,h}^2. \end{aligned} \tag{5.12}$$

It follows from (5.12) and (3.15) that

$$\begin{aligned}|P\boldsymbol{v}|_1^2 &= \sum_{K\in\mathcal{T}_h} |P\boldsymbol{v} - \boldsymbol{Q}_K^0\boldsymbol{v}|_{1,K}^2 \lesssim \sum_{K\in\mathcal{T}_h} h_K^{-2}\|P\boldsymbol{v} - \boldsymbol{Q}_K^0\boldsymbol{v}\|_{0,K}^2 \\ &\lesssim \sum_{K\in\mathcal{T}_h} h_K^{-2}\|\boldsymbol{v} - P\boldsymbol{v}\|_{0,K}^2 + \sum_{K\in\mathcal{T}_h} h_K^{-2}\|\boldsymbol{v} - \boldsymbol{Q}_K^0\boldsymbol{v}\|_{0,K}^2 \lesssim |\boldsymbol{v}|_{1,h}^2.\end{aligned}$$

Therefore we can finish the proof by combining the last two inequalities. ◻

Lemma 5.3. *For any $\boldsymbol{v} \in \boldsymbol{V}_h$, it holds that*

$$\|\boldsymbol{v} - \Pi P\boldsymbol{v}\|_0^2 \lesssim \rho_S^{-1}|\boldsymbol{v}|_{1,h}^2. \tag{5.13}$$

Proof. For $k \geq 2$, the inequality (5.13) is the result of Lemma 5.2 and (5.3). For $k = 1$, we obtain from the triangle inequality, (3.15), Lemma 5.2, and (5.3) that

$$\begin{aligned}\|\boldsymbol{v} - \Pi P\boldsymbol{v}\|_0^2 =& \|\boldsymbol{v} - \boldsymbol{Q}_h P\boldsymbol{v}\|_0^2 \lesssim \|\boldsymbol{v} - P\boldsymbol{v}\|_0^2 + \|P\boldsymbol{v} - \boldsymbol{Q}_h P\boldsymbol{v}\|_0^2 \\ &\lesssim \|\boldsymbol{v} - P\boldsymbol{v}\|_0^2 + h^2|P\boldsymbol{v}|_1^2 \lesssim h^2|\boldsymbol{v}|_{1,h}^2 \lesssim \rho_S^{-1}|\boldsymbol{v}|_{1,h}^2,\end{aligned}$$

as required. ◻

Combining Lemma 4.1, Theorem 5.1, and Lemmas 5.2 and 5.3, we have the following estimate of the condition number of XS.

Theorem 5.4. *Let R be a Jacobi or a symmetric Gauss-Seidel smoother, let $\mathcal{B}$ be one step of classical multigrid iteration, and let Π be $\boldsymbol{Q}_h$. Then the preconditioner X defined in* (5.4) *or* (5.5) *satisfies*

$$\kappa(XS) \lesssim 1.$$

6. Numerical results

In this section, we will report some numerical results to testify the efficiency and robustness of the auxiliary space preconditioners developed in Sections 4 and 5 for the mixed finite element method (2.3)-(2.4). Let $\Omega = (-1,1)^2$, $\mu = 0.5$ and the load $\boldsymbol{f} = \mathbf{1}$. We use the uniform triangulation $\mathcal{T}_h$ of Ω except in the last example. The stopping criteria of our iterative methods is that the relative residual is less than 10^{-8}, and the initial guess is zero. We run the code on the laptop with Intel Core i5-5300U CPU (2.30 GHz) and 4GB RAM.

6.1. **Block-diagonal preconditioner.** First we use the minimal residual (MINRES) method with the block-diagonal preconditioner

$$\begin{pmatrix} D_h^{-1} & 0 \\ 0 & (B_h D_h^{-1} B_h^T + C_h)^{-1} \end{pmatrix}$$

to solve the matrix equation (2.5) for the mixed finite element method (2.3)-(2.4), where D_h is the diagonal matrix of M_h. To precondition the Schur complement $B_h D_h^{-1} B_h^T + C_h$, we apply the multiplicative auxiliary space preconditioner (5.5), in which we employ three steps of the Gauss-Seidel smoother for R and one step of the V-cycle multigrid method with one pre-smoothing and one post-smoothing for $\mathcal{B}$.

The iteration numbers and CPU time for the block-diagonal preconditioned MINRES method are shown in Tables 1-3 for $k = 1, 2, 3$, from which we can see that the iteration steps are uniform with respect to the mesh-size h and the Lamé constant λ. But it depends on k as the traditional multigrid methods applied to the Poisson equation.

Remark 6.1. The iteration steps can be further reduced by introducing additional scaling in D_h and S_h, i.e.,

$$\begin{pmatrix} \omega_1 D_h^{-1} & 0 \\ 0 & \omega_2(\omega_1 B_h D_h^{-1} B_h^T + C_h)^{-1} \end{pmatrix},$$

and optimizing ω_1 and ω_2.

6.2. **Approximate block-factorization preconditioner.** Next we examine the generalized minimal residual (GMRES) method with the approximate block-factorization preconditioner

$$\begin{pmatrix} D_h & B_h^T \\ B_h & -C_h \end{pmatrix}^{-1} = \begin{pmatrix} I & D_h^{-1} B_h^T \\ 0 & -I \end{pmatrix} \begin{pmatrix} D_h & 0 \\ B_h & B_h D_h^{-1} B_h^T + C_h \end{pmatrix}^{-1}.$$

Set restart=20 in the GMRES method. We still exploit the same multiplicative auxiliary space preconditioner as in the block-diagonal preconditioner to solve the Schur complement.

The iteration numbers and CPU time for the approximate block-factorization preconditioned GMRES method are shown in Tables 4-7 for $k = 1, 2, 3, 4$. Again the iteration steps are uniform with respect to the mesh-size h and the Lamé constant λ.

TABLE 1. The iteration steps and CPU time (in seconds) of the block-diagonal preconditioned MINRES method for $k = 1$

#dofs	$\lambda = 0$		$\lambda = 10$		$\lambda = 100$		$\lambda = 1000$		$\lambda = +\infty$	
	steps	time	steps	time	steps	time	steps	time	steps	time
1891	43	0.08	65	0.11	74	0.13	74	0.13	74	0.13
7363	46	0.34	75	0.58	84	0.64	86	0.64	86	0.64
29059	47	1.39	78	2.31	91	2.65	92	2.67	92	2.67
115459	47	5.24	81	9.05	95	10.5	96	10.6	96	10.6
460291	47	21.9	81	37.4	97	44.7	98	45.2	98	45.2
1838083	46	88.0	81	151.3	97	183.6	100	189.3	100	189.3

TABLE 2. The iteration steps and CPU time (in seconds) of the block-diagonal preconditioned MINRES method for $k = 2$

#dofs	$\lambda = 0$		$\lambda = 10$		$\lambda = 100$		$\lambda = 1000$		$\lambda = +\infty$	
	steps	time	steps	time	steps	time	steps	time	steps	time
1811	57	0.16	85	0.23	93	0.25	94	0.25	94	0.25
7075	58	0.66	91	1.03	98	1.09	100	1.12	100	1.12
27971	58	2.64	93	4.13	102	4.57	102	4.57	102	4.57
111235	58	10.2	95	16.5	103	17.8	104	18.0	104	18.0
443651	57	41.3	96	68.4	104	73.9	104	73.9	106	75.3
1772035	57	170.7	97	285.5	104	307.3	106	313.2	106	313.2

TABLE 3. The iteration steps and CPU time (in seconds) of the block-diagonal preconditioned MINRES method for $k = 3$

#dofs	$\lambda = 0$		$\lambda = 10$		$\lambda = 100$		$\lambda = 1000$		$\lambda = +\infty$	
	steps	time	steps	time	steps	time	steps	time	steps	time
971	56	0.09	89	0.17	91	0.19	91	0.19	91	0.19
3763	58	0.58	88	0.81	94	0.87	94	0.87	94	0.87
14819	58	2.31	90	3.56	96	3.76	96	3.76	96	3.76
58819	58	8.78	90	13.8	96	14.4	96	14.4	97	14.6
234371	57	34.6	90	55.0	96	58.4	98	59.6	98	59.6
935683	57	141.4	90	221.7	96	236.5	98	241.4	98	241.4

The performance of the approximate block-factorization preconditioned GMRES method is better than the block-diagonal preconditioned MINRES method. The iteration steps and CPU time are almost halved compared with the block-diagonal preconditioner. By Tables 4-7, the approximate block-factorization preconditioner is not robust with respect to k. Indeed, the auxiliary linear conforming element space preconditioner is not robust even for the Lagrange element method for the Poisson equation with respect to k.

At last, we testify the approximate block-factorization preconditioner on unstructured meshes. Let the Lamé coefficient

$$\lambda = \begin{cases} 1, & x^2 + y^2 < 0.25, \\ 10000, & x^2 + y^2 \geq 0.25. \end{cases}$$

TABLE 4. The iteration steps and CPU time (in seconds) of the approximate block-factorization preconditioned GMRES method for $k = 1$

#dofs	$\lambda = 0$		$\lambda = 10$		$\lambda = 100$		$\lambda = 1000$		$\lambda = +\infty$	
	steps	time	steps	time	steps	time	steps	time	steps	time
1891	20	0.05	34	0.06	38	0.06	39	0.08	39	0.08
7363	22	0.16	39	0.30	46	0.33	47	0.36	47	0.36
29059	24	0.73	45	1.33	50	1.45	51	1.54	51	1.54
115459	24	2.87	47	5.63	54	6.52	55	6.71	55	6.71
460291	25	12.9	50	25.6	57	28.9	59	30.4	59	30.4
1838083	26	57.3	52	112.2	63	140.9	64	142.5	64	142.5

TABLE 5. The iteration steps and CPU time (in seconds) of the approximate block-factorization preconditioned GMRES method for $k = 2$

#dofs	$\lambda = 0$		$\lambda = 10$		$\lambda = 100$		$\lambda = 1000$		$\lambda = +\infty$	
	steps	time	steps	time	steps	time	steps	time	steps	time
1811	18	0.05	29	0.07	31	0.08	31	0.08	32	0.08
7075	20	0.23	32	0.36	34	0.39	35	0.41	35	0.41
27971	22	1.03	35	1.62	37	1.70	38	1.73	38	1.73
111235	23	4.20	37	6.80	40	7.38	41	7.60	41	7.60
443651	24	18.3	39	29.4	44	34.7	44	34.7	44	34.7
1772035	24	75.6	40	125.8	45	145.2	46	148.4	46	148.4

TABLE 6. The iteration steps and CPU time (in seconds) of the approximate block-factorization preconditioned GMRES method for $k = 3$

#dofs	$\lambda = 0$		$\lambda = 10$		$\lambda = 100$		$\lambda = 1000$		$\lambda = +\infty$	
	steps	time	steps	time	steps	time	steps	time	steps	time
971	20	0.03	27	0.05	28	0.06	28	0.06	28	0.06
3763	21	0.19	29	0.26	30	0.27	30	0.27	30	0.27
14819	22	0.86	30	1.19	32	1.25	32	1.25	32	1.25
58819	23	3.62	31	4.82	33	5.21	33	5.21	33	5.21
234371	24	15.5	32	20.6	34	21.7	35	22.3	35	22.3
935683	24	62.9	33	86.4	35	91.4	36	94.1	36	94.1

To generate unstructured meshes, we use the following a posteriori error estimator recently advanced in [31]

$$\sum_{K\in\mathcal{T}_h} h_K^4\|\mathrm{rot}\ \mathbf{rot}(\mathfrak{A}\boldsymbol{\sigma}_h)\|_{0,K}^2 + \sum_{K\in\mathcal{T}_h} h_K\|[\boldsymbol{t}^T((\mathfrak{A}\boldsymbol{\sigma}_h)\boldsymbol{t})]\|_{0,\partial K}^2$$
$$+ \sum_{K\in\mathcal{T}_h} h_K^3\|[\partial_t(\boldsymbol{\nu}^T((\mathfrak{A}\boldsymbol{\sigma}_h)\boldsymbol{t})) - \boldsymbol{t}^T\mathbf{rot}(\mathfrak{A}\boldsymbol{\sigma}_h)]\|_{0,\partial K}^2$$

TABLE 7. The iteration steps and CPU time (in seconds) of the approximate block-factorization preconditioned GMRES method for $k = 4$

#dofs	$\lambda = 0$		$\lambda = 10$		$\lambda = 100$		$\lambda = 1000$		$\lambda = +\infty$	
	steps	time	steps	time	steps	time	steps	time	steps	time
427	26	0.02	34	0.03	32	0.03	32	0.03	32	0.03
1627	28	0.14	35	0.17	36	0.19	36	0.19	36	0.19
6355	28	0.67	36	0.86	37	0.89	38	0.92	38	0.92
25123	29	2.90	37	3.70	38	3.74	38	3.74	38	3.74
99907	30	12.2	38	15.5	40	16.3	40	16.3	40	16.3
398467	31	50.9	40	65.8	42	69.1	43	70.8	43	70.8

to design the adaptive algorithm. The Dörfler marking strategy [38] with bulk parameter $\theta = 0.1$ and the newest vertex bisection [65] are employed in the adaptive algorithm. For consideration of unstructured meshes, we replace the auxiliary discrete problem (5.2) with the following vectorial Poisson equation discretized by the conforming linear element: Find $\boldsymbol{u}_h \in \boldsymbol{\mathcal{V}}_h$ such that

$$2\mu(\boldsymbol{\nabla} \boldsymbol{u}_h, \boldsymbol{\nabla} \boldsymbol{v}_h) = (\boldsymbol{f}, \boldsymbol{v}_h) \qquad \forall\, \boldsymbol{v}_h \in \boldsymbol{\mathcal{V}}_h.$$

This discrete problem will be solved efficiently by algebraic multigrid method [68]. The unstructured mesh with #dofs = 1669303 for $k = 3$ generated by the adaptive algorithm is drawn in Figure 1, which fairly exhibits the singularity of the solution around $x^2 + y^2 = 0.25$. Partial numerical results on the iteration steps of the approximate block-factorization preconditioned GMRES method for the adaptive algorithm are listed in Table 8. The approximate block-factorization preconditioner is still robust with respect to the #dofs and the Lamé coefficient λ even on adaptive meshes.

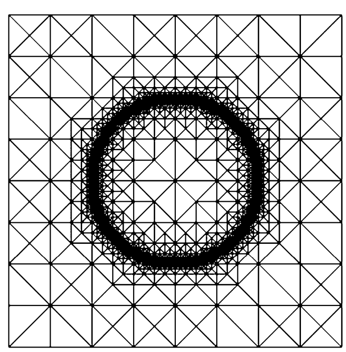

FIGURE 1. Adaptive mesh with #dofs = 1669303 for $k = 3$

TABLE 8. The iteration steps of the approximate block-factorization preconditioned GMRES method for the adaptive algorithm

$k=2$		$k=3$		$k=4$	
#dofs	steps	#dofs	steps	#dofs	steps
6226	34	5531	33	5992	39
35190	35	36026	35	37129	40
175532	40	158348	35	177197	40
386054	42	309797	35	347044	40
550932	43	501416	36	613309	40
778059	44	636848	35	790819	40
1102619	46	805682	36	897810	40
1394239	46	1028794	36	1036132	41
1756771	49	1311058	36	1180274	41
2198618	43	1669303	37	1319178	40

References

[1] S. Adams and B. Cockburn, *A mixed finite element method for elasticity in three dimensions*, J. Sci. Comput. **25** (2005), no. 3, 515–521, DOI 10.1007/s10915-004-4807-3. MR2221175

[2] P. F. Antonietti and B. Ayuso, *Schwarz domain decomposition preconditioners for discontinuous Galerkin approximations of elliptic problems: non-overlapping case*, M2AN Math. Model. Numer. Anal. **41** (2007), no. 1, 21–54, DOI 10.1051/m2an:2007006. MR2323689

[3] P. F. Antonietti and B. Ayuso, *Multiplicative Schwarz methods for discontinuous Galerkin approximations of elliptic problems*, M2AN Math. Model. Numer. Anal. **42** (2008), no. 3, 443–469, DOI 10.1051/m2an:2008012. MR2423794

[4] D. N. Arnold, *An interior penalty finite element method with discontinuous elements*, SIAM J. Numer. Anal. **19** (1982), no. 4, 742–760, DOI 10.1137/0719052. MR664882

[5] D. N. Arnold and G. Awanou, *Rectangular mixed finite elements for elasticity*, Math. Models Methods Appl. Sci. **15** (2005), no. 9, 1417–1429, DOI 10.1142/S0218202505000741. MR2166210

[6] D. N. Arnold, G. Awanou, and R. Winther, *Finite elements for symmetric tensors in three dimensions*, Math. Comp. **77** (2008), no. 263, 1229–1251, DOI 10.1090/S0025-5718-08-02071-1. MR2398766

[7] D. N. Arnold, F. Brezzi, and L. D. Marini, *A family of discontinuous Galerkin finite elements for the Reissner-Mindlin plate*, J. Sci. Comput. **22/23** (2005), 25–45, DOI 10.1007/s10915-004-4134-8. MR2142189

[8] D. N. Arnold and R. Winther, *Mixed finite elements for elasticity*, Numer. Math. **92** (2002), no. 3, 401–419, DOI 10.1007/s002110100348. MR1930384

[9] O. Axelsson and M. Neytcheva, *Preconditioning methods for linear systems arising in constrained optimization problems*, Numer. Linear Algebra Appl. **10** (2003), no. 1-2, 3–31, DOI 10.1002/nla.310. MR1964284

[10] O. Axelsson and A. Padiy, *On a robust and scalable linear elasticity solver based on a saddle point formulation*, Internat. J. Numer. Methods Engrg. **44** (1999), no. 6, 801–818, DOI 10.1002/(SICI)1097-0207(19990228)44:6⟨801::AID-NME525⟩3.0.CO;2-Y. MR1672010

[11] B. Ayuso de Dios, M. Holst, Y. Zhu, and L. Zikatanov, *Multilevel preconditioners for discontinuous, Galerkin approximations of elliptic problems, with jump coefficients*, Math. Comp. **83** (2014), no. 287, 1083–1120, DOI 10.1090/S0025-5718-2013-02760-3. MR3167451

[12] B. Ayuso de Dios and L. Zikatanov, *Uniformly convergent iterative methods for discontinuous Galerkin discretizations*, J. Sci. Comput. **40** (2009), no. 1-3, 4–36, DOI 10.1007/s10915-009-9293-1. MR2511726

[13] R. E. Bank, B. D. Welfert, and H. Yserentant, *A class of iterative methods for solving saddle point problems*, Numer. Math. **56** (1990), no. 7, 645–666, DOI 10.1007/BF01405194. MR1031439

[14] A. T. Barker, S. C. Brenner, and L.-Y. Sung, *Overlapping Schwarz domain decomposition preconditioners for the local discontinuous Galerkin method for elliptic problems*, J. Numer. Math. **19** (2011), no. 3, 165–187, DOI 10.1515/JNUM.2011.008. MR2854135

[15] M. Benzi, G. H. Golub, and J. Liesen, *Numerical solution of saddle point problems*, Acta Numer. **14** (2005), 1–137, DOI 10.1017/S0962492904000212. MR2168342

[16] D. Boffi, F. Brezzi, and M. Fortin, *Mixed finite element methods and applications*, Springer Series in Computational Mathematics, vol. 44, Springer, Heidelberg, 2013. MR3097958

[17] D. Braess and R. Verfürth, *A posteriori error estimators for the Raviart-Thomas element*, SIAM J. Numer. Anal. **33** (1996), no. 6, 2431–2444, DOI 10.1137/S0036142994264079. MR1427472

[18] J. H. Bramble and J. E. Pasciak, *The analysis of smoothers for multigrid algorithms*, Math. Comp. **58** (1992), no. 198, 467–488, DOI 10.2307/2153197. MR1122058

[19] S. C. Brenner, *A nonconforming mixed multigrid method for the pure displacement problem in planar linear elasticity*, SIAM J. Numer. Anal. **30** (1993), no. 1, 116–135, DOI 10.1137/0730006. MR1202659

[20] S. C. Brenner, *A nonconforming mixed multigrid method for the pure traction problem in planar linear elasticity*, Math. Comp. **63** (1994), no. 208, 435–460, S1–S5, DOI 10.2307/2153278. MR1257574

[21] S. C. Brenner, *Two-level additive Schwarz preconditioners for nonconforming finite elements*, Domain decomposition methods in scientific and engineering computing (University Park, PA, 1993), Contemp. Math., vol. 180, Amer. Math. Soc., Providence, RI, 1994, pp. 9–14, DOI 10.1090/conm/180/01951. MR1312372

[22] S. C. Brenner, *Korn's inequalities for piecewise H^1 vector fields*, Math. Comp. **73** (2004), no. 247, 1067–1087, DOI 10.1090/S0025-5718-03-01579-5. MR2047078

[23] S. C. Brenner, J. Cui, T. Gudi, and L.-Y. Sung, *Multigrid algorithms for symmetric discontinuous Galerkin methods on graded meshes*, Numer. Math. **119** (2011), no. 1, 21–47, DOI 10.1007/s00211-011-0379-y. MR2824854

[24] S. C. Brenner, J. Cui, and L.-Y. Sung, *Multigrid methods for the symmetric interior penalty method on graded meshes*, Numer. Linear Algebra Appl. **16** (2009), no. 6, 481–501, DOI 10.1002/nla.630. MR2522959

[25] S. C. Brenner, H. Li, and L.-Y. Sung, *Multigrid methods for saddle point problems: Stokes and Lamé systems*, Numer. Math. **128** (2014), no. 2, 193–216, DOI 10.1007/s00211-014-0607-3. MR3259489

[26] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, 3rd ed., Texts in Applied Mathematics, vol. 15, Springer, New York, 2008. MR2373954

[27] S. C. Brenner, K. Wang, and J. Zhao, *Poincaré-Friedrichs inequalities for piecewise H^2 functions*, Numer. Funct. Anal. Optim. **25** (2004), no. 5-6, 463–478, DOI 10.1081/NFA-200042165. MR2106270

[28] S. C. Brenner and J. Zhao, *Convergence of multigrid algorithms for interior penalty methods*, Appl. Numer. Anal. Comput. Math. **2** (2005), no. 1, 3–18, DOI 10.1002/anac.200410019. MR2157481

[29] K. Brix, M. Campos Pinto, and W. Dahmen, *A multilevel preconditioner for the interior penalty discontinuous Galerkin method*, SIAM J. Numer. Anal. **46** (2008), no. 5, 2742–2768, DOI 10.1137/07069691X. MR2421055

[30] L. Chen, J. Hu, and X. Huang, *Stabilized mixed finite element methods for linear elasticity on simplicial grids in $\mathbb{R}^n$*, Comput. Methods Appl. Math. **17** (2017), no. 1, 17–31, DOI 10.1515/cmam-2016-0035. MR3589339

[31] L. Chen, J. Hu, X. Huang, and H. Man, *Residual-based a posteriori error estimates for symmetric conforming mixed finite elements for linear elasticity problems*, arXiv:1705.04106, 2016. To appear in Sci. China Math.

[32] L. Chen, J. Wang, Y. Wang, and X. Ye, *An auxiliary space multigrid preconditioner for the weak Galerkin method*, Comput. Math. Appl. **70** (2015), no. 4, 330–344, DOI 10.1016/j.camwa.2015.04.016. MR3372035

[33] L. Chen, Y. Wu, L. Zhong, and J. Zhou, *Multigrid preconditioners for mixed finite element methods of vector Laplacian*, arXiv:1601.04095, 2016.

[34] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland Publishing Co., Amsterdam-New York-Oxford, 1978. Studies in Mathematics and its Applications, Vol. 4. MR0520174

[35] B. Cockburn, O. Dubois, J. Gopalakrishnan, and S. Tan, *Multigrid for an HDG method*, IMA J. Numer. Anal. **34** (2014), no. 4, 1386–1425, DOI 10.1093/imanum/drt024. MR3269430

[36] V. A. Dobrev, R. D. Lazarov, P. S. Vassilevski, and L. T. Zikatanov, *Two-level preconditioning of discontinuous Galerkin approximations of second-order elliptic equations*, Numer. Linear Algebra Appl. **13** (2006), no. 9, 753–770, DOI 10.1002/nla.504. MR2269798

[37] H. S. Dollar, *Constraint-style preconditioners for regularized saddle point problems*, SIAM J. Matrix Anal. Appl. **29** (2007), no. 2, 672–684, DOI 10.1137/050626168. MR2318371

[38] W. Dörfler, *A convergent adaptive algorithm for Poisson's equation*, SIAM J. Numer. Anal. **33** (1996), no. 3, 1106–1124, DOI 10.1137/0733054. MR1393904

[39] J. Douglas Jr. and J. E. Roberts, *Global estimates for mixed methods for second order elliptic equations*, Math. Comp. **44** (1985), no. 169, 39–52, DOI 10.2307/2007791. MR771029

[40] H. C. Elman. *Iterative Methods for Large, Sparse, Nonsymmetric Systems of Linear Equations*. PhD thesis, Yale University, New Haven, CT, 1982.

[41] X. Feng and O. A. Karakashian, *Two-level additive Schwarz methods for a discontinuous Galerkin approximation of second order elliptic problems*, SIAM J. Numer. Anal. **39** (2001), no. 4, 1343–1365, DOI 10.1137/S0036142900378480. MR1870847

[42] J. Gopalakrishnan and G. Kanschat, *A multilevel discontinuous Galerkin method*, Numer. Math. **95** (2003), no. 3, 527–550, DOI 10.1007/s002110200392. MR2012931

[43] A. Hannukainen, *Continuous preconditioners for the mixed Poisson problem*, BIT **52** (2012), no. 1, 65–83, DOI 10.1007/s10543-011-0346-0. MR2891654

[44] R. Hiptmair and J. Xu, *Nodal auxiliary space preconditioning in* $\mathbf{H}(\mathbf{curl})$ *and* $\mathbf{H}(\mathrm{div})$ *spaces*, SIAM J. Numer. Anal. **45** (2007), no. 6, 2483–2509, DOI 10.1137/060660588. MR2361899

[45] Q. Hong, J. Kraus, J. Xu, and L. Zikatanov, *A robust multigrid method for discontinuous Galerkin discretizations of Stokes and linear elasticity equations*, Numer. Math. **132** (2016), no. 1, 23–49, DOI 10.1007/s00211-015-0712-y. MR3439214

[46] J. Hu, *Finite element approximations of symmetric tensors on simplicial grids in* $\mathbb{R}^n$*: the higher order case*, J. Comput. Math. **33** (2015), no. 3, 283–296, DOI 10.4208/jcm.1412-m2014-0071. MR3352360

[47] J. Hu and G. Yu, *A unified analysis of quasi-optimal convergence for adaptive mixed finite element methods*, arXiv:1601.00132, 2016. To appear in SIAM J. Numer. Anal.

[48] J. Hu and S. Zhang, *A family of conforming mixed finite elements for linear elasticity on triangular grids, arXiv:1406.7457,* 2015.

[49] J. Hu and S. Zhang, *A family of symmetric mixed finite elements for linear elasticity on tetrahedral grids*, Sci. China Math. **58** (2015), no. 2, 297–307, DOI 10.1007/s11425-014-4953-5. MR3301063

[50] J. Huang and X. Huang, *Local and parallel algorithms for fourth order problems discretized by the Morley-Wang-Xu element method*, Numer. Math. **119** (2011), no. 4, 667–697, DOI 10.1007/s00211-011-0396-x. MR2854124

[51] C. Keller, N. I. M. Gould, and A. J. Wathen, *Constraint preconditioning for indefinite linear systems*, SIAM J. Matrix Anal. Appl. **21** (2000), no. 4, 1300–1317, DOI 10.1137/S0895479899351805. MR1780274

[52] A. Klawonn and G. Starke, *A preconditioner for the equations of linear elasticity discretized by the PEERS element*, Numer. Linear Algebra Appl. **11** (2004), no. 5-6, 493–510, DOI 10.1002/nla.357. MR2067817

[53] T. V. Kolev and P. S. Vassilevski, *Auxiliary space AMG for H(curl) problems*, Domain Decomposition Methods in Science and Engineering XVII, Lect. Notes Comput. Sci. Eng., vol. 60, Springer, Berlin, 2008, pp. 147–154, DOI 10.1007/978-3-540-75199-1_13. MR2436078

[54] T. V. Kolev and P. S. Vassilevski, *Parallel auxiliary space AMG for* $H(\mathrm{curl})$ *problems*, J. Comput. Math. **27** (2009), no. 5, 604–623, DOI 10.4208/jcm.2009.27.5.013. MR2536904

[55] J. Kraus, R. Lazarov, M. Lymbery, S. Margenov, and L. Zikatanov, *Preconditioning heterogeneous* $H(\mathrm{div})$ *problems by additive Schur complement approximation and applications*, SIAM J. Sci. Comput. **38** (2016), no. 2, A875–A898, DOI 10.1137/140974092. MR3474851

[56] J. Kraus, M. Lymbery, and S. Margenov, *Auxiliary space multigrid method based on additive Schur complement approximation*, Numer. Linear Algebra Appl. **22** (2015), no. 6, 965–986, DOI 10.1002/nla.1959. MR3426324

[57] J. K. Kraus and S. K. Tomar, *A multilevel method for discontinuous Galerkin approximation of three-dimensional anisotropic elliptic problems*, Numer. Linear Algebra Appl. **15** (2008), no. 5, 417–438, DOI 10.1002/nla.544. MR2423513

[58] J. K. Kraus and S. K. Tomar, *Multilevel preconditioning of two-dimensional elliptic problems discretized by a class of discontinuous Galerkin methods*, SIAM J. Sci. Comput. **30** (2008), no. 2, 684–706, DOI 10.1137/060667372. MR2385881

[59] C.-O. Lee, *Multigrid methods for the pure traction problem of linear elasticity: mixed formulation*, SIAM J. Numer. Anal. **35** (1998), no. 1, 121–145, DOI 10.1137/S0036142995282832. MR1618440

[60] Y.-J. Lee, J. Wu, and J. Chen, *Robust multigrid method for the planar linear elasticity problems*, Numer. Math. **113** (2009), no. 3, 473–496, DOI 10.1007/s00211-009-0232-8. MR2534133

[61] B. Li and X. Xie, *A two-level algorithm for the weak Galerkin discretization of diffusion problems*, J. Comput. Appl. Math. **287** (2015), 179–195, DOI 10.1016/j.cam.2015.03.043. MR3342024

[62] D. Loghin and A. J. Wathen, *Analysis of preconditioners for saddle-point problems*, SIAM J. Sci. Comput. **25** (2004), no. 6, 2029–2049, DOI 10.1137/S1064827502418203. MR2086829

[63] C. Lovadina and R. Stenberg, *Energy norm a posteriori error estimates for mixed finite element methods*, Math. Comp. **75** (2006), no. 256, 1659–1674, DOI 10.1090/S0025-5718-06-01872-2. MR2240629

[64] K.-A. Mardal and R. Winther, *Preconditioning discretizations of systems of partial differential equations*, Numer. Linear Algebra Appl. **18** (2011), no. 1, 1–40, DOI 10.1002/nla.716. MR2769031

[65] W. F. Mitchell, *A comparison of adaptive refinement techniques for elliptic problems*, ACM Trans. Math. Software **15** (1989), no. 4, 326–347 (1990), DOI 10.1145/76909.76912. MR1062496

[66] J. E. Pasciak and Y. Wang, *A multigrid preconditioner for the mixed formulation of linear plane elasticity*, SIAM J. Numer. Anal. **44** (2006), no. 2, 478–493, DOI 10.1137/040617820. MR2218956

[67] I. Perugia and V. Simoncini, *Block-diagonal and indefinite symmetric preconditioners for mixed finite element formulations*, Numer. Linear Algebra Appl. **7** (2000), no. 7-8, 585–616, DOI 10.1002/1099-1506(200010/12)7:7/8⟨585::AID-NLA214⟩3.3.CO;2-6. Preconditioning Techniques for Large Sparse Matrix Problems in Industrial Applications (Minneapolis, MN, 1999). MR1802361

[68] J. W. Ruge and K. Stüben, *Algebraic multigrid*, Multigrid Methods, Frontiers Appl. Math., vol. 3, SIAM, Philadelphia, PA, 1987, pp. 73–130. MR972756

[69] T. Rusten, P. S. Vassilevski, and R. Winther, *Interior penalty preconditioners for mixed finite element approximations of elliptic problems*, Math. Comp. **65** (1996), no. 214, 447–466, DOI 10.1090/S0025-5718-96-00720-X. MR1333325

[70] Y. Saad, *Iterative Methods for Sparse Linear Systems*, 2nd ed., Society for Industrial and Applied Mathematics, Philadelphia, PA, 2003. MR1990645

[71] J. Schöberl, *Multigrid methods for a parameter dependent problem in primal variables*, Numer. Math. **84** (1999), no. 1, 97–119, DOI 10.1007/s002110050465. MR1724348

[72] R. Stenberg, *On the construction of optimal mixed finite element methods for the linear elasticity problem*, Numer. Math. **48** (1986), no. 4, 447–462, DOI 10.1007/BF01389651. MR834332

[73] R. Stenberg, *Postprocessing schemes for some mixed finite elements*, RAIRO Modél. Math. Anal. Numér. **25** (1991), no. 1, 151–167, DOI 10.1051/m2an/1991250101511. MR1086845

[74] R. S. Tuminaro, J. Xu, and Y. Zhu, *Auxiliary space preconditioners for mixed finite element methods*, Domain Decomposition Methods in Science and Engineering XVIII, Lect. Notes Comput. Sci. Eng., vol. 70, Springer, Berlin, 2009, pp. 99–109, DOI 10.1007/978-3-642-02677-5_9. MR2743962

[75] M. Wang, *On the necessity and sufficiency of the patch test for convergence of nonconforming finite elements*, SIAM J. Numer. Anal. **39** (2001), no. 2, 363–384, DOI 10.1137/S003614299936473X. MR1860273

[76] Y. Wang, *Overlapping Schwarz preconditioner for the mixed formulation of plane elasticity*, Appl. Numer. Math. **54** (2005), no. 2, 292–309, DOI 10.1016/j.apnum.2004.09.032. MR2148047

[77] C. Wieners, *Robust multigrid methods for nearly incompressible elasticity*, Computing **64** (2000), no. 4, 289–306, DOI 10.1007/s006070070026. International GAMM-Workshop on Multigrid Methods (Bonn, 1998). MR1782022

[78] J. Xu, *Iterative methods by space decomposition and subspace correction*, SIAM Rev. **34** (1992), no. 4, 581–613, DOI 10.1137/1034116. MR1193013

[79] J. Xu, *The auxiliary space method and optimal multigrid preconditioning techniques for unstructured grids*, Computing **56** (1996), no. 3, 215–235, DOI 10.1007/BF02238513. International GAMM-Workshop on Multi-level Methods (Meisdorf, 1994). MR1393008

[80] J. Xu, *Fast Poisson-based solvers for linear and nonlinear PDEs*, Proceedings of the International Congress of Mathematicians. Volume IV, Hindustan Book Agency, New Delhi, 2010, pp. 2886–2912. MR2827997

[81] S. Zhang and J. Xu, *Optimal solvers for fourth-order PDEs discretized on unstructured grids*, SIAM J. Numer. Anal. **52** (2014), no. 1, 282–307, DOI 10.1137/120878148. MR3162408

[82] L. Zhong, E. T. Chung, and C. Liu, *Fast solvers for the symmetric IPDG discretization of second order elliptic problems*, Int. J. Numer. Anal. Model. **12** (2015), no. 3, 455–475. MR3352442

[83] Y. Zhu and Y.-J. Lee, *Auxiliary space preconditioner for linear elasticity equation with weakly imposed symmetry problems.* In: Conference on the Mathematics of Finite Elements and Applications, 2016.

[84] Y. Zhu, E. Sifakis, J. Teran, and A. Brandt, *An efficient multigrid method for the simulation of high-resolution elastic solids*, ACM Trans. Graph. (TOG), 29(2), MAR 2010.

[85] J. Hu, *A new family of efficient conforming mixed finite elements on both rectangular and cuboid meshes for linear elasticity in the symmetric formulation*, SIAM J. Numer. Anal., 53(2015), no. 3, 1438-1463, DOI 10.1137/130945272. MR3357634

[86] J. Hu and S. Zhang, *Finite element approximations of symmetric tensors on simplicial grids in $\mathbb{R}^n$: the lower order case*, Math. Models Methods Appl. Sci., 26(2016), no. 9, 1649-1669, DOI 10.1142/S0218202516500408. MR3529252

Department of Mathematics, University of California at Irvine, Irvine, California 92697 – and – Beijing Institute for Scientific and Engineering Computing, Beijing University of Technology, Beijing, 100124, People's Republic of China

E-mail address: `chenlong@math.uci.edu`

LMAM and School of Mathematical Sciences, Peking University, Beijing 100871, People's Republic of China

E-mail address: `hujun@math.pku.edu.cn`

College of Mathematics and Information Science, Wenzhou University, Wenzhou 325035, People's Republic of China

E-mail address: `xuehaihuang@wzu.edu.cn`