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,

\[
\begin{pmatrix}
M^\lambda_h & B^T_h \\
B_h & O
\end{pmatrix},
\]

where $M^\lambda_h$ 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 $\lambda$ is the Lamé number which could be very large for nearly incompressible material. We aim to develop preconditioners robust to both $h$ and $\lambda$. 

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 11771318, 11301396 and 11671304, Zhejiang Provincial Natural Science Foundation of China Projects LY17A010010, LY15A010015 and LY15A010016, and Wenzhou Science and Technology Plan Project G20160019.

©2017 American Mathematical Society
In [46][48][49][85][86], a stability result is established in the $H(\text{div}, \Omega; \mathbb{S}) \times L^2(\Omega; \mathbb{R}^n)$ norm whose matrix form is

$$(M_h + B_h^T M_{u,h}^{-1} B_h O M_{u,h}^{-1})^{-1}.$$  

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

$$(M_h + B_h^T M_{u,h}^{-1} B_h O M_{u,h}^{-1})^{-1}.$$  

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

$$(M_h O B_h M_h^{-1} B_h^T).$$  

Therefore we can use the block-diagonal preconditioner

$$(M_h^{-1} O (B_h M_h^{-1} B_h^T)^{-1})$$  

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], $H(\text{curl})$ and $H(\text{div})$ 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 $\lambda$ 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 $\lambda$. 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

\[
\begin{pmatrix}
I & D_h^{-1}B_h^T \\
0 & -I
\end{pmatrix}
\begin{pmatrix}
D_h & 0 \\
B_h & \tilde{S}_h
\end{pmatrix}^{-1},
\]

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 $H(\text{div})$ 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 $H(\text{div})$-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 [52,60,76]. 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], which in 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 $\ker(\text{div})$ by using either blockwise Gauss-Seidel smoothers 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 earlier our approach does not require a prior knowledge of the discrete $\ker(\text{div})$. 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 “\(\lesssim\) · · ·” to mean that “\(\lesssim C\) · · ·”, where \(C\) is a generic positive constant independent of \(h\) and the Lamé constant \(\lambda\), which may take different values at different times, and \(a \approx b\) 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 \(\mathcal{S}\) 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 \(H^m(G; X)\) be the usual Sobolev space of functions taking values in the finite-dimensional vector space \(X\) for \(X\) 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 \(L^2(G; X)\). If \(G\) is \(\Omega\), we abbreviate \(\| \cdot \|_{m, \Omega}\), \(| \cdot |_{m, \Omega}\) and \((\cdot, \cdot)_G\) by \(\| \cdot \|_m\), \(| \cdot |_m\) and \((\cdot, \cdot)\), respectively. Let \(H^m_0(\Omega; \mathbb{R}^n)\) be the closure of \(C^\infty_0(\Omega; \mathbb{R}^n)\) with respect to the norm \(\| \cdot \|_{m, \Omega}\). Denote by \(H(\text{div}; G; \mathcal{S})\) the Sobolev space of square-integrable symmetric tensor fields with square-integrable divergence. For any \(\tau \in H(\text{div}; \Omega; \mathcal{S})\), we define the following norm:

\[
\| \tau \|_{H(\text{div})} := \left( \| \tau \|_0^2 + \| \text{div} \tau \|_0^2 \right)^{1/2}.
\]

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

\[
\begin{align*}
\mathcal{A}\sigma + b(\tau, u) &= 0 \quad \forall \tau \in \Sigma, \\
\text{div}(\sigma) &= -f \quad \forall v \in V,
\end{align*}
\]

where

\[
\mathcal{A}\sigma := \frac{1}{2\mu} \left( \sigma - \frac{\lambda}{n\lambda + 2\mu}(\text{tr} \sigma)I \right),
\]

\[
b(\tau, v) := (\text{div} \tau, v)
\]

with \(\mathcal{A}\) being the compliance tensor of fourth order defined by

\(\mathcal{I}\) is the identity tensor, \(\text{tr}\) is the trace operator, and the positive constants \(\lambda\) and \(\mu\) are the Lamé constants.

Suppose the domain \(\Omega\) is subdivided by a family of shape regular simplicial grids \(T_h\) (cf. [26][34]) with \(h := \max_{K \in T_h} h_K\) and \(h_K := \text{diam}(K)\). Let \(\mathcal{F}_h\) be the union of all \((n-1)\)-dimensional faces of \(T_h\) 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 \(\nu_F\). Let \(P_m(G)\) stand for the set of all polynomials in \(G\) with the total degree no more than \(m\), and let \(P_m(G; X)\) denote the tensor or vector version of \(P_m(G)\) for \(X\) being \(\mathbb{S}\) or \(\mathbb{R}^n\), respectively.

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

\[
[w] := \begin{cases} w^+(\nu^+ \cdot \nu_F) + w^-(\nu^- \cdot \nu_F), & \text{if } F \in \mathcal{F}_h^i, \\
w, & \text{if } F \in \mathcal{F}_h \setminus \mathcal{F}_h^i. \end{cases}
\]
For each $K \in \mathcal{T}_h$, define an $H(\text{div}, K; \mathbb{S})$ bubble function space of polynomials of degree $k$ as
\[ B_{K,k} := \{ \tau \in P_k(K; \mathbb{S}) : \tau|_K = 0 \}. \]
It is easy to check that $B_{K,1}$ is merely the zero space. Denote the vertices of simplex $K$ by $x_{K,0}, \ldots, x_{K,n}$. If it does not cause confusion, we will abbreviate $x_{K,i}$ as $x_i$ for $i = 0, \ldots, n$. For any edge $[x_i, x_j]$ of element $K$, let $t_{i,j}$ be the associated unit tangent vectors and let
\[ T_{i,j} := t_{i,j}T_{i,j}^T, \quad 0 \leq i < j \leq n. \]
It has been proved in [16] that the $(n+1)n/2$ symmetric tensors $T_{i,j}$ form a basis of $\mathbb{S}$, and for $k \geq 2$,
\[ B_{K,k} = \sum_{0 \leq i < j \leq n} \lambda_i \lambda_j P_{i+j-2}(K) T_{i,j}, \]
where $P_{i+j-2}(K)$ and $\lambda_i$ is the associated barycentric coordinate corresponding to $x_i$ for $0 \leq i < j \leq n$. Some global finite element spaces are given by
\[
B_{k,h} := \{ \tau \in H(\text{div}, \Omega; \mathbb{S}) : \tau|_K \in B_{K,k} \quad \forall K \in \mathcal{T}_h \}, \\
\Sigma_{k,h} := \{ \tau \in H^1(\Omega; \mathbb{S}) : \tau|_K \in P_k(K; \mathbb{S}) \quad \forall K \in \mathcal{T}_h \}, \\
\Sigma_h := \Sigma_{k,h} + B_{k,h}, \\
V_h := \{ v \in L^2(\Omega; \mathbb{R}^n) : v|_K \in P_{k-1}(K; \mathbb{R}^n) \quad \forall K \in \mathcal{T}_h \},
\]
with integer $k \geq 1$. The local rigid motion space is defined as
\[ R(K) := \{ v \in H^1(K; \mathbb{R}^n) : \varepsilon(v) = 0 \}
\]
with $\varepsilon(v) := (\nabla v + (\nabla v)^T)/2$ being the linearized strain tensor.

With previous preparation, the mixed finite element method for linear elasticity proposed in [30, 40, 48, 49] is defined as follows: Find $(\sigma_h, u_h) \in \Sigma_h \times V_h$ such that
\begin{align}
\tag{2.3}
a(\sigma_h, \tau_h) + b(\tau_h, u_h) &= 0 \quad \forall \tau_h \in \Sigma_h, \\
\tag{2.4}
b(\sigma_h, v_h) - c(u_h, v_h) &= -(f, v_h) \quad \forall v_h \in V_h,
\end{align}
where
\[ c(u_h, v_h) := \eta \sum_{F \in \mathcal{F}_h} h_F^{-1} \int_F [u_h] \cdot [v_h] \, ds,
\]
\[ \eta := \begin{cases} 0, & \text{if } k \geq n + 1, \\ 1, & \text{if } 1 \leq k \leq n. \end{cases} \]
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 $\Sigma_h$ and $V_h$, we can write the matrix form of (2.3)-(2.4) as
\[ (M_h^\lambda \quad B_h^T \quad -C_h) \begin{pmatrix} \sigma_h \\ u_h \end{pmatrix} = \begin{pmatrix} 0 \\ f \end{pmatrix}, \]
where $M_h^\lambda$ 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 $\sigma_h, u_h$, and $f$ for the vector representations of corresponding functions.
Let
\[ \tilde{\Sigma}_h := \{ \tau \in \Sigma_h : \int_\Omega \text{tr}\tau \, dx = 0 \}, \]
\[ A_h(\sigma_h, u_h; \tau_h, v_h) := a(\sigma_h, \tau_h) + b(\tau_h, u_h) + b(\sigma_h, v_h) - c(u_h, v_h). \]
For \( k \geq n + 1 \), the following inf-sup condition is the immediate result of (3.4)-(3.5) in [10]:
\[ (2.6) \quad \| \tilde{\sigma}_h \|_{H(\text{div})} + \| u_h \|_0 \lesssim \sup_{(\tau_h, v_h) \in \Sigma_h \times V_h} \frac{A_h(\tilde{\sigma}_h, \tilde{u}_h; \tau_h, v_h)}{\| \tau_h \|_{H(\text{div})} + \| v_h \|_0}, \]
for any \((\tilde{\sigma}_h, \tilde{u}_h) \in \Sigma_h \times V_h\).

Thanks to the inf-sup condition (2.6), the system (2.5) is stable in the space \( \Sigma_h \times V_h \) equipped with the \( H(\text{div}, \Omega; \mathbb{S}) \times L^2(\Omega; \mathbb{R}^n) \)-norm which leads to a block-diagonal preconditioner requiring a nontrivial solver for \((M_h + B_h^T M_u^{-1} B_h)^{-1}\). In the next section we shall establish another stability result of (2.6) 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 \( A(\cdot, \cdot, \cdot) \) is stable on \( \Sigma_h \times V_h \) with mesh-dependent norms.

For each \( K \in \mathcal{T}_h \), denote by \( \nu_i \) the unit outward normal vector of the \( i \)-th face \( F_i \) of element \( K \). For any \( \tau_h \in \Sigma_h \) and \( v_h \in V_h \), define
\[ \| \tau_h \|_{0,h}^2 := \| \tau_h \|_0^2 + \sum_{F \in \mathcal{F}_h} h_F \| \tau_h \nu_F \|_{0,F}^2, \]
\[ \| v_h \|_{1,h}^2 := \| \varepsilon_h(v_h) \|_0^2 + \sum_{F \in \mathcal{F}_h} h_F^{-1} \| [v_h] \|_{0,F}^2, \]
\[ \| v_h \|_{0,h}^2 := c(v_h, v_h). \]
Here \( \varepsilon_h \) is the elementwise symmetric gradient. We shall prove the stability of (2.6) in the mesh-dependent norms \( \| \cdot \|_{0,h} \times \| \cdot \|_{1,h} \). The key is the following inf-sup condition: for \( k \geq n + 1 \),
\[ (3.1) \quad \| v_h \|_{1,h} \lesssim \sup_{\tau_h \in \Sigma_h} \frac{b(\tau_h, v_h)}{\| \tau_h \|_{0,h}} \quad \forall v_h \in V_h. \]
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 \( x_0, \ldots, x_n \), we label the face opposite to \( x_i \) as the \( i \)-th face \( F_i \). For the edge \( x_i x_j \), \( i \neq j \), define
\[ N_{i,j} := \frac{1}{(\nu_i^T t_{i,j})(\nu_j^T t_{i,j})} (\nu_i \nu_j^T + \nu_j \nu_i^T), \quad 0 \leq i < j \leq n. \]
Here recall that \( t_{i,j} \) is a unit tangent vector of edge \( x_i x_j \) and \( \nu_i \) is the unit outwards normal vector of face \( F_i \). Due to the shape regularity of the triangulation, it holds that
\[ \nu_i^T t_{i,j} \approx 1, \quad 0 \leq i < j \leq n. \]
By direct manipulation, we have the following results about \( T_{i,j} \) and \( N_{i,j} \):
\[ (3.2) \quad T_{i,j} : N_{k,l} = \delta_{ik} \delta_{jl}, \quad 0 \leq i < j \leq n, \quad 0 \leq k < l \leq n, \]
Lemma 3.1. For any \( q_{ij} \in L^2(K) \), \( 0 \leq i < j \leq n \), let \( \tau_1 = \sum_{0 \leq i<j \leq n} q_{ij} T_{i,j} \) and let \( \tau_2 = \sum_{0 \leq i<j \leq n} q_{ij} N_{i,j} \), then it holds that
\[
\| \tau_1 \|_{0,K}^2 \leq \| \tau_2 \|_{0,K}^2 \sim \sum_{0 \leq i<j \leq n} \| q_{ij} \|_{0,K}^2 .
\]

Proof. Using the Cauchy-Schwarz inequality and (3.3), we have
\[
\| \tau_1 \|_{0,K}^2 \leq \frac{(n+1)n}{2} \sum_{0 \leq i<j \leq n} \| q_{ij} \|_{0,K}^2 = \frac{(n+1)n}{2} \sum_{0 \leq i<j \leq n} \| q_{ij} \|_{0,K}^2 ,
\]
\[
\| \tau_2 \|_{0,K}^2 \leq \frac{(n+1)n}{2} \sum_{0 \leq i<j \leq n} \| q_{ij} N_{i,j} \|_{0,K}^2 \sim \frac{(n+1)n}{2} \sum_{0 \leq i<j \leq n} \| q_{ij} \|_{0,K}^2 .
\]

On the other side, it follows from Cauchy-Schwarz inequality and (3.2) that
\[
\sum_{0 \leq i<j \leq n} \| q_{ij} \|_{0,K}^2 = \sum_{0 \leq i<j \leq n} \int_K q_{ij}^2 \, dx = \sum_{0 \leq i<j \leq n} \sum_{0 \leq k<l \leq n} \int_K q_{ij} q_{kl} \delta_{ik} \delta_{jl} \, dx
\]
\[
= \sum_{0 \leq i<j \leq n} \sum_{0 \leq k<l \leq n} \int_K q_{ij} T_{i,j} : q_{kl} N_{k,l} \, dx
\]
\[
= \int_K \tau_1 : \tau_2 \, dx \leq \| \tau_1 \|_{0,K} \| \tau_2 \|_{0,K}.
\]

Hence we conclude the result by combining the last three inequalities. \( \square \)

We then map \( \varepsilon_h(V_h) \) into the \( H(\text{div}, K; S) \) bubble function space. For each element \( K \in T_h \), introduce a bijective connection operator \( E_K : P_{k-2}(K; S) \to B_{K,k} \) with \( k \geq 2 \) as follows: For any \( \tau = \sum_{0 \leq i<j \leq n} q_{ij} N_{i,j} \) with \( q_{ij} \in P_{k-2}(K) \), \( 0 \leq i < j \leq n \), define
\[
E_K \tau := \sum_{0 \leq i<j \leq n} \lambda_i \lambda_j q_{ij} T_{i,j} .
\]
Applying Lemma 3.1 and the scaling argument, we get for any \( \tau \in P_{k-2}(K; S) \),
\[
\| E_K \tau \|_{0,K}^2 \sim \sum_{0 \leq i<j \leq n} \| \lambda_i \lambda_j q_{ij} \|_{0,K}^2 \sim \sum_{0 \leq i<j \leq n} \| q_{ij} \|_{0,K}^2 \sim \| \tau \|_{0,K}^2 .
\]
(3.4)
\[
\int_K E_K \tau : \tau \, dx \sim \sum_{0 \leq i<j \leq n} \int_K \lambda_i \lambda_j q_{ij}^2 \, dx \sim \sum_{0 \leq i<j \leq n} \| q_{ij} \|_{0,K}^2 \sim \| \tau \|_{0,K}^2 .
\]
(3.5)
Denote by \( E \) the elementwise global version of \( E_K \), i.e., \( E|_K := E_K \) for each \( K \in T_h \).

Next, we give an equivalent formulation of the mesh-dependent norm \( | \cdot |_{1,h} \). For each \( F \in F_h \), denote by \( \pi_F \) the orthogonal projection operator from \( L^2(F; \mathbb{R}^n) \) onto \( P_1(F; \mathbb{R}^n) \). Define the broken \( H^1 \) space as
\[
H^1(T_h; \mathbb{R}^n) := \{ v \in L^2(\Omega; \mathbb{R}^n) : v|_K \in H^1(K; \mathbb{R}^n) \quad \forall K \in T_h \} .
\]
The domain of the mesh-dependent norm \( | \cdot |_{1,h} \) can be extended from \( V_h \) to \( H^1(T_h; \mathbb{R}^n) \).
Lemma 3.2. We have the norm equivalence:

\[(3.6)\quad |v|^2_{1,h} \simeq \|v\|_0^2 + \sum_{F \in \mathcal{T}_h} h_F^{-1} \|\pi_F[v]\|_{0,F}^2 \quad \forall \ v \in H^1(T_h; \mathbb{R}^n).\]

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

\[(3.7)\quad \sum_{F \in \mathcal{T}_h} h_F^{-1} \|v - \pi_F[v]\|_{0,F}^2 = \sum_{F \in \mathcal{T}_h} h_F^{-1} \|v - \pi_F[v]\|_{0,F}^2 \leq \sum_{F \in \mathcal{T}_h} h_F^{-1} \|v - \pi_v\|_{0,F}^2 \lesssim \|\varepsilon_h(v)\|_0^2.

Then the equivalence (3.6) follows from the triangle inequality. \(\square\)

We shall also use the following discrete Korn’s inequality (cf. (1.22) in [22] and (34) in [27])

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

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

3.2. **inf-sup condition in mesh-dependent norms.** The inf-sup condition we need is actually for the subspace \(\Sigma_h\) 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 \(\hat{\tau}_h = \tau_h - (\frac{1}{n} \int_{\Omega} \tau_h \, dx)\delta\). Therefore the inf-sup conditions (3.1) and (3.9) are equivalent.

**Lemma 3.3.** For \(k \geq n + 1\), we have the following inf-sup condition,

\[(3.9)\quad |v_h|_{1,h} \lesssim \sup_{\tau_h \in \Sigma_h} \frac{b(\tau_h, v_h)}{\|\tau_h\|_{0,h}},\]

for any \(v_h \in V_h\).

**Proof.** Given a \(v_h \in V_h\), we shall construct a \(\hat{\tau}_h \in \Sigma_h\) to verify (3.9).

We first control the norm \(\|\varepsilon_h(v_h)\|_0\). For any \(v_h \in V_h\), take \(\tau_1 = E\varepsilon_h(v_h)\). It follows from (3.4) that

\[(3.10)\quad \|\tau_1\|_0 \simeq \|\varepsilon_h(v_h)\|_0.

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

\[(3.11)\quad b(\tau_1, v_h) = \int_{\Omega} \tau_1 : \varepsilon_h(v_h) \, dx \geq C_1 \|\varepsilon_h(v_h)\|_0^2.

Next we control the jump term. Choose \(\tau_2 \in \Sigma_h\) such that all the degrees of freedom (cf. Lemma 2.1 in [20]) for \(\tau_2\) vanish except the following one:

\[\int_{F} (\tau_2 \nu_F) \cdot w \, ds = h_F^{-1} \int_{F} [v_h] \cdot w \, ds \quad \forall \ w \in P_1(F; \mathbb{R}^n)\] on each face \(F\).
Then we have
\[ \int_K \tau_2 : \varepsilon_h(v_h) \, dx = 0, \quad \int_F (\tau_2 \cdot \nabla v_h) \cdot \pi_F[v_h] \, ds = h_F^{-1} \| \pi_F[v_h] \|^2_{0,F}, \]
(3.12)
\[ \| \tau_2 \|^2_0 \lesssim \sum_{F \in T_h} h_F^{-1} \| \pi_F[v_h] \|^2_{0,F}. \]

Thus by (3.7) and (3.12), there exists a constant \( C_2 > 0 \) such that
\[ b(\tau_2, v_h) = \sum_{F \in T_h} \int_F (\tau_2 \cdot \nabla v_h) \cdot \pi_F[v_h] \, ds \]
\[ = \sum_{F \in T_h} \int_F (\tau_2 \cdot \nabla v_h) \cdot ((v_h) - \pi_F[v_h]) \, ds + \sum_{F \in T_h} h_F^{-1} \| \pi_F[v_h] \|^2_{0,F} \]
(3.13)
\[ \geq -C_2 \| \varepsilon_h(v_h) \|^2_0 + \frac{1}{2} \sum_{F \in T_h} h_F^{-1} \| \pi_F[v_h] \|^2_{0,F}. \]

Now taking \( \tau_h = \tau_1 + \frac{C_1}{2C_2} \tau_2 \), it holds from (3.11) and (3.13) that
\[ b(\tau_h, v_h) = b(\tau_1, v_h) + \frac{C_1}{2C_2} b(\tau_2, v_h) \]
\[ \geq \frac{C_1}{2} \| \varepsilon_h(v_h) \|^2_0 + \frac{C_1}{4C_2} \sum_{F \in T_h} h_F^{-1} \| \pi_F[v_h] \|^2_{0,F}. \]

Thanks to (3.6), we get
\[ |v_h|_{1,h}^2 \leq b(\tau_h, v_h). \]

On the other hand, it follows from the inverse inequality, (3.10) and (3.12) that
\[ \| \tau_h \|_{0,h} \lesssim \| \tau_h \|_0 \lesssim |v_h|_{1,h}. \]

Finally, the inf-sup condition (3.11) is the result of the last two inequalities and consequently (3.12) holds by taking \( \tau_h = \tau_1 - \left( \frac{1}{3} \int_{\Omega} \text{tr} \tau_h \, dx \right) \delta \).

3.3. Coercivity in the null space of the \( \text{div} \) 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 \( \sigma \in \Sigma_\text{tr} \),
\[ a(\sigma, \sigma) \geq \frac{1}{2n\lambda + 2\mu} \| \sigma \|^2_0. \]

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

Lemma 3.4 (Proposition 9.1.1 in [15]). For \( \tau \in H(\text{div}, \Omega; \mathbb{S}) \) satisfying \( \int_{\Omega} \text{tr} \tau \, dx = 0 \), we have
\[ \| \tau \|_0 \lesssim \| \tau \|_a + \| \text{div} \tau \|_{-1}, \]
where \( \| \tau \|^2_a := a(\tau, \tau) \) and \( \| \text{div} \tau \|_{-1} := \sup_{\nu \in H^1_0(\Omega; \mathbb{R}^n)} \frac{b(\tau, \nu)}{|\nu|_1}. \)

We then move to the discrete case. Define discrete norms
\[ \| \text{div} \tau \|_{-1,h} := \sup_{v_h \in V_h} \frac{b(\tau, v_h)}{|v_h|_{1,h}}, \]
\[ \| h \text{div} \tau \|^2 := \sum_{K \in T_h} h_K^2 \| \text{div} \tau \|^2_{0,K}. \]
Let $Q_h^{k-1}$ be the $L^2$ orthogonal projection from $L^2(\Omega; \mathbb{R}^n)$ onto $V_h$, which will be abbreviated as $Q_h$. It holds the following error estimate (cf. \[ \text{[26,34]} \]):

\begin{equation}
\| v - Q_h v \|_{0,K} + h^{k/2}_K \| v - Q_h v \|_{0,\partial K} \lesssim h^{\min(k,m)}_K \| v \|_{m,K} \quad \forall v \in H^m(\Omega; \mathbb{R}^n),
\end{equation}

with integer $m \geq 1$.

**Lemma 3.5.** For any $v \in H^1(\Omega; \mathbb{R}^n)$ satisfying $\int_{\Omega} \text{tr} v \, dx = 0$, we have

\[ \| v \|_0 \lesssim \| v \|_a + h \| \text{div} v \| + \| \text{div} v \|_{-1,h}. \]

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

\[ b(\tau, v) = b(\tau, v - Q_h v) + b(\tau, Q_h v) \lesssim \| h \text{div} \|_{1} + b(\tau, Q_h v). \]

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

\begin{equation}
|Q_h v|_{1,h}^2 = \sum_{F \in F_h} h_F^{-1} \| Q_h v \|_{0,F}^2 = \sum_{F \in F_h} h_F^{-1} \| Q_h v - v \|_{0,F}^2 \lesssim |v|_1^2.
\end{equation}

Hence we get from the last two inequalities that

\[ \| \text{div} v \|_{-1} = \sup_{v \in H^1(\Omega; \mathbb{R}^n)} \frac{b(\tau, v)}{|v|_1} \lesssim h \| \text{div} v \| + \sup_{v_h \in V_h} \frac{b(\tau, v_h)}{|v_h|_{1,h}}. \]

Therefore we can end the proof by using Lemma 3.4. \qed

### 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 $\text{div} \Sigma_h \subset V_h$, then $\ker(\text{div}) \cap \Sigma_h \subset \ker(\text{div}) \cap \Sigma$. The stability follows from Lemma 3.3 and inf-sup condition (3.10).

**Theorem 3.6.** For $k \geq n + 1$, it follows that for any $(\bar{\sigma}_h, \bar{u}_h) \in \Sigma_h \times V_h$,

\begin{equation}
\| \bar{\sigma}_h \|_{0,h} + |\bar{u}_h|_{1,h} \lesssim \sup_{(\tau_h, u_h) \in \Sigma_h \times V_h} \frac{A(\bar{\sigma}_h, \bar{u}_h; \tau_h, u_h)}{\| \tau_h \|_{0,h} + |u_h|_{1,h}}.
\end{equation}

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

\begin{equation}
\| \sigma - \sigma_h \|_{0,h} + |Q_h u - u_h|_{1,h} \lesssim h^{k+1} \| \sigma \|_{k+1},
\end{equation}

\begin{equation}
|u - u_h|_{1,h} \lesssim h^{k+1} \left( \| \sigma \|_{k+1} + \| u \|_{k+1} \right).
\end{equation}

Moreover, when $\Omega$ is convex, we have

\begin{equation}
|Q_h u - u_h|_{0} \lesssim h^{k+2} \| \sigma \|_{k+1}.
\end{equation}

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

\begin{align*}
(\sigma - \sigma_h, \tau_h) + b(\tau_h, u - u_h) &= 0 \quad \forall \tau_h \in \Sigma_h, \\
b(\sigma - \sigma_h, v_h) &= 0 \quad \forall v_h \in V_h.
\end{align*}

Let $I_h^{HZ}$ be the standard interpolation from $H^1(\Omega; \mathbb{S})$ to $\Sigma_h$ defined in \[ \text{[10]} \] Remark 3.1, and it holds that

\[ \text{div}(I_h^{HZ} \sigma) = Q_h(\text{div} \sigma). \]
Thus we have from (3.22) that
\[ b(I^H \sigma - \sigma_h, v_h) = b(\sigma - \sigma_h, v_h) = 0. \]
By the definition of \( Q_h \) and (3.21),
\[ b(\tau_h, Q_h u - u_h) = b(\tau_h, u - u_h) = -a(\sigma - \sigma_h, \tau_h). \]
Combining the last two equalities, it holds that
\[ a(I^H \sigma - \sigma_h, Q_h u - u_h; \tau_h, v_h) \]
\[ = a(I^H \sigma - \sigma_h, \tau_h) + b(\tau_h, Q_h u - u_h) + b(I^H \sigma - \sigma_h, v_h) \]
\[ = a(I^H \sigma - \sigma_h, \tau_h), \]
which together with (3.17) implies
\[ \|I^H \sigma - \sigma_h\|_{0, h} + |Q_h u - u_h|_{1, h} \lesssim \|I^H \sigma - \sigma\|_{0, h}. \]
Therefore we will achieve (3.18)-(3.19) by using the last inequality, and the error estimate of \( I^H \sigma \) and \( Q_h \). 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 \( \|\sigma - \sigma_h\|_{0, h} \) has been proved in [46] Remarks 3.1-3.2 and [49] Remarks 3.6, but the second order higher superconvergent rates of \( |Q_h u - u_h|_{1, h} \) and \( \|Q_h u - u_h\|_0 \) are new which can be used to reconstruct a better approximation of displacement. The convergence rate of \( |u - u_h|_{1, h} \) 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 \( 1 \leq k \leq n \), which is similar to the one in [39] Lemma 3.2.

**Theorem 3.9.** For \( 1 \leq k \leq n \), it holds for any \((\hat{\sigma}_h, \hat{u}_h) \in \hat{\Sigma}_h \times V_h\) that
\[ \|\sigma_h\|_{0, h} + |u_h|_{1, h} \lesssim \sup_{(\tau_h, v_h) \in \Sigma_h \times V_h} \frac{A(\sigma_h, \hat{u}_h; \tau_h, v_h)}{\|\tau_h\|_{0, h} + |v_h|_{1, h}}. \]

**Proof.** As demonstrated in Lemma 3.3 it is equivalent to prove that
\[ \|\sigma_h\|_{0, h} + |u_h|_{1, h} \lesssim \sup_{(\tau_h, v_h) \in \Sigma_h \times V_h} \frac{A(\sigma_h, \hat{u}_h; \tau_h, v_h)}{\|\tau_h\|_{0, h} + |v_h|_{1, h}} := \beta. \]
The notation \( \beta \) is introduced just for ease of presentation. Let \( \tau_1 = E \varepsilon_h(\hat{u}_h) \) for \( k \geq 2 \) and let \( \tau_1 = 0 \) for \( k = 1 \), then it holds from Cauchy-Schwarz inequality that
\[ A(\sigma_h, \hat{u}_h; \tau_1, 0) = a(\sigma_h, \tau_1) + b(\tau_1, \hat{u}_h) \geq -\|\sigma_h\|_a \|\tau_1\|_a + b(\tau_1, \hat{u}_h). \]
Using (3.10)-(3.11), there exists a constant \( C_3 > 0 \) such that
\[ A(\sigma_h, \hat{u}_h; \tau_1, 0) \geq C_1 \|\varepsilon_h(\hat{u}_h)\|_a^2 - C_3 \|\sigma_h\|_a \|\varepsilon_h(\hat{u}_h)\|_0 \]
\[ \geq \frac{C_1}{2} \|\varepsilon_h(\hat{u}_h)\|_a^2 - \frac{C_3}{2C_1} \|\sigma_h\|_a^2. \]
Let \( v_1 \in V_h \) such that \( v_1|_K = h^2 \frac{\text{div} \sigma_h}{|K|} \) for each \( K \in T_h \). Applying inverse inequality, we have
\[ |v_1|_{1, h} \lesssim \|h \text{div} \sigma_h\| \lesssim \|\sigma_h\|_0. \]
Thus there exists a constant $C_4 > 0$ such that
\[
\mathcal{A}(\overline{\sigma}_h, \overline{u}_h; 0, v_1) = b(\overline{\sigma}_h, v_1) - c(\overline{u}_h, v_1) \geq \|h \text{ div } \overline{\sigma}_h\|^2 - \|\overline{u}_h\|_c \|v_1\|_c \\
\geq \|h \text{ div } \overline{\sigma}_h\|^2 - C_4 \|\overline{u}_h\|_c \|h \text{ div } \overline{\sigma}_h\| \\
\geq \frac{1}{2}\|h \text{ div } \overline{\sigma}_h\|^2 - \frac{C_4^2}{2} \|\overline{u}_h\|_c^2.
\]
(3.28)

Now taking $\tau_h = \overline{\sigma}_h + \frac{C_4}{C_3} \tau_1$ and $v_h = -\overline{u}_h + \frac{1}{C_4} v_1$, we have from (3.20) and (3.28)
\[
\mathcal{A}(\overline{\sigma}_h, \overline{u}_h; \tau_h, v_h) \\
= \mathcal{A}(\overline{\sigma}_h, \overline{u}_h; \overline{\sigma}_h, -\overline{u}_h) + \frac{C_1}{C_3} \mathcal{A}(\overline{\sigma}_h, \overline{u}_h; \tau_1, 0) + \frac{1}{C_4} \mathcal{A}(\overline{\sigma}_h, \overline{u}_h; 0, v_1) \\
\geq \frac{1}{2}\|\overline{\sigma}_h\|^2_\alpha + \frac{1}{2}\|\overline{u}_h\|^2_\alpha + \frac{1}{2C^2_4}\|h \text{ div } \overline{\sigma}_h\|^2 + \frac{C_4^2}{2C^2_3}\|\varepsilon_h(\overline{u}_h)\|^2_0,
\]
which together with (3.10) and (3.27) indicates
\[
\|\overline{\sigma}_h\|^2_\alpha + \|h \text{ div } \overline{\sigma}_h\|^2 + \|\overline{u}_h\|^2_1 \lesssim \mathcal{A}(\overline{\sigma}_h, \overline{u}_h; \tau_h, v_h) \lesssim \beta(\|\overline{\sigma}_h\|_0 + \|\overline{u}_h\|_1,h).
\]

Thus we obtain from the last two inequalities that
\[
\|\overline{\sigma}_h\|^2_0 + \|\overline{u}_h\|^2_1 \lesssim \|\overline{\sigma}_h\|^2_\alpha + \|h \text{ div } \overline{\sigma}_h\|^2 + \|\overline{u}_h\|^2_1 + \beta^2 \lesssim \beta(\|\overline{\sigma}_h\|_0 + \|\overline{u}_h\|_1,h) + \beta^2,
\]
which implies inf-sup condition (3.29). \(\square\)

**Corollary 3.10.** Let $1 \leq k \leq n$. Assume that $\sigma \in H^{k+1}(\Omega; \mathbb{S})$ and $u \in H^k(\Omega; \mathbb{R}^n)$, then
\[
\|\sigma - \sigma_h\|_{0,h} + \|u - u_h\|_{1,h} \lesssim h^{k-1}(\|\sigma\|_{k+1} + \|u\|_k).
\]

The convergence rate of $|u - u_h|_{1,h}$ is optimal. But the $L^2$-type error of $\|\sigma - \sigma_h\|_{0,h}$ 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 \((\sigma_h, u_h)\) for the higher order case \(k \geq n + 1\) in this subsection.

To this end, let

\[
V_h^* := \{v \in L^2(\Omega; \mathbb{R}^n) : v|_K \in P_{k+1}(K; \mathbb{R}^n) \quad \forall K \in \mathcal{T}_h\}.
\]

Then a postprocessed displacement can be defined as follows: Find \(u_h^* \in V_h^*\) such that

\[
Q_h u_h^* = u_h,
\]

(3.29)

\[
(\varepsilon(u_h^*), \varepsilon(v))_K = (A\sigma_h, \varepsilon(v))_K \quad \forall v \in (I - Q_h) V_h^*|_K,
\]

(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 \(u_h^* \in V_h^*\) and \(\phi_h \in V_h\) satisfying

\[
(\varepsilon(u_h^*), \varepsilon(v))_K + (v, \phi_h)_K = (A\sigma_h, \varepsilon(v))_K \quad \forall v \in V_h^*|_K,
\]

(3.31)

\[
(u_h^*, \psi)_K = (u_h, \psi)_K \quad \forall \psi \in V_h|_K.
\]

To derive the error estimate for the postprocessed displacement \(u_h^*\), 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 \((\sigma_h, u_h^*) \in \Sigma_h \times V_h^*\) such that

\[
\mathcal{A}_h(\sigma_h, u_h^*; \tau_h, v_h^*) = -(Q_h f, v_h^*) \quad \forall (\tau_h, v_h^*) \in \Sigma_h \times V_h^*,
\]

(3.32)

where

\[
\mathcal{A}_h(\sigma_h, u_h^*; \tau_h, v_h^*) := \mathcal{A}_h(\sigma_h, u_h^*; \tau_h, v_h^*) + (\varepsilon_h(u_h^*) - A\sigma_h, \varepsilon_h(v_h^* - Q_h v_h^*)).
\]

**Lemma 3.12.** The mixed finite element method (2.3)–(2.4) and the problem (3.31) are equivalent in the following sense: If \((\sigma_h, u_h^*) \in \Sigma_h \times V_h^*\) is the solution of the problem (3.31) and \(u_h = Q_h u_h^*\), then \((\sigma_h, u_h) \in \Sigma_h \times V_h\) solves the mixed finite element method (2.3)–(2.4); Conversely, if \((\sigma_h, u_h) \in \Sigma_h \times V_h\) is the solution of the mixed finite element method (2.3)–(2.4) and \(u_h^* \in V_h^*\) is the postprocessed displacement defined by (3.29)–(3.30), then \((\sigma_h, u_h^*) \in \Sigma_h \times V_h^*\) solves the problem (3.31).

**Proof.** Taking any \((\tau_h, v_h) \in \Sigma_h \times V_h\), and noting the fact that \(v_h = Q_h v_h\) and \(\text{div} \Sigma_h \subset V_h\), we have

\[
\mathcal{A}_h(\sigma_h, u_h^*; \tau_h, v_h) = \mathcal{A}_h(\sigma_h, u_h^*; \tau_h, v_h) = a(\sigma_h, \tau_h) + b(\tau_h, u_h^*) + b(\sigma_h, v_h)
\]

(3.33)

\[
= a(\sigma_h, \tau_h) + b(\tau_h, Q_h u_h^*) + b(\sigma_h, v_h) = \mathcal{A}_h(\sigma_h, Q_h u_h^*; \tau_h, v_h).
\]

Hence we can see from (3.32) that \((\sigma_h, u_h)\) solves the mixed finite element method (2.3)–(2.4) if \((\sigma_h, u_h^*)\) is the solution of problem (3.31).

Conversely, since \(\text{div} \Sigma_h \subset V_h\) and \((I - Q_h)^2 = I - Q_h\), it follows from (3.32) and (3.29) that

\[
\mathcal{A}_h(\sigma_h, u_h^*; \tau_h, v_h) = \mathcal{A}_h(\sigma_h, u_h^*; \tau_h, Q_h v_h^*) + \mathcal{A}_h(\sigma_h, u_h^*; 0, v_h^* - Q_h v_h^*)
\]

(3.34)

\[
= \mathcal{A}_h(\sigma_h, u_h^*; \tau_h, Q_h v_h^*) + (\varepsilon_h(u_h^*) - A\sigma_h, \varepsilon_h(v_h^* - Q_h v_h^*))
\]

which together with (3.30) means that \((\sigma_h, u_h^*)\) solves problem (3.31). •
Lemma 3.13. For any $v \in H^1(T_h; \mathbb{R}^n)$, it holds that
\begin{equation}
|v - Q_h v|_{1,h} \approx \|\varepsilon_h(v - Q_h v)\|_0.
\end{equation}

Proof. It is sufficient to prove that
\begin{equation}
\sum_{F \in \mathcal{F}_h} h_F^{-1} \|v - Q_h v\|^2_{0,F} \lesssim \|\varepsilon_h(v - Q_h v)\|^2_0.
\end{equation}

Let $\pi$ be defined as in Lemma 3.12 and $w = v - Q_h v$. It follows from (3.3) in [22]
\begin{equation}
\sum_{F \in \mathcal{F}_h} h_F^{-1} \|w - Q_h w\|^2_{0,F} = \sum_{F \in \mathcal{F}_h} h_F^{-1} \|((w - \pi w) - Q_h (w - \pi w))\|^2_{0,F}
\end{equation}
\begin{equation}
\lesssim \sum_{K \in \mathcal{T}_h} |w - \pi w|^2_{1,K} \lesssim \|\varepsilon_h(w)\|^2_0.
\end{equation}

On the other hand,
\begin{equation}
\sum_{F \in \mathcal{F}_h} h_F^{-1} \|v - Q_h v\|^2_{0,F} = \sum_{F \in \mathcal{F}_h} h_F^{-1} \|[w - \pi w]\|^2_{0,F}.
\end{equation}

Therefore (3.35) follows from (3.30).

Theorem 3.14. For any $(\tilde{\sigma}_h, \tilde{u}_h) \in \Sigma_h \times V_h$, it follows that
\begin{equation}
\|\tilde{\sigma}_h\|_{0,h} + |\tilde{u}_h|_{1,h} \lesssim \sup_{(\tau_h, v_h) \in \Sigma_h \times V_h} A_h(\tilde{\sigma}_h, \tilde{u}_h; \tau_h, v_h).
\end{equation}

Proof. For any $v_h \in V_h$, we have from (3.32) that
\begin{equation}
A_h(\tilde{\sigma}_h, \tilde{u}_h; \tau_h, v_h) = \mathcal{A}(\tilde{\sigma}_h, Q_h \tilde{u}_h; \tau_h, v_h).
\end{equation}

Since $(\tilde{\sigma}_h, Q_h \tilde{u}_h) \in \Sigma_h \times V_h$, it holds from (3.17) that
\begin{equation}
\|\tilde{\sigma}_h\|_{0,h} + |Q_h \tilde{u}_h|_{1,h} \lesssim \sup_{(\tau_h, v_h) \in \Sigma_h \times V_h} \mathcal{A}(\tilde{\sigma}_h, Q_h \tilde{u}_h; \tau_h, v_h)
\end{equation}
\begin{equation}
= \sup_{(\tau_h, v_h) \in \Sigma_h \times V_h} A_h(\tilde{\sigma}_h, \tilde{u}_h; \tau_h, v_h)
\end{equation}
\begin{equation}
\lesssim \sup_{(\tau_h, v_h) \in \Sigma_h \times V_h} A_h(\tilde{\sigma}_h, \tilde{u}_h; \tau_h, v_h).
\end{equation}

If $\varepsilon_h(\tilde{u}_h - Q_h \tilde{u}_h) = 0$, then (3.37) is the immediate result of the triangle inequality, (3.38) and (3.34). Next we only need to focus on $\varepsilon_h(\tilde{u}_h - Q_h \tilde{u}_h) \neq 0$. Similarly, as in (3.33), we get
\begin{equation}
A_h(\tilde{\sigma}_h, \tilde{u}_h; 0, \tilde{u}_h - Q_h \tilde{u}_h) = (\varepsilon_h(\tilde{u}_h) - \mathcal{A}\tilde{\sigma}_h, \varepsilon_h(\tilde{u}_h - Q_h \tilde{u}_h)).
\end{equation}

Then we rewrite it as
\begin{equation}
\|\varepsilon_h(\tilde{u}_h - Q_h \tilde{u}_h)\|^2_0 = (\mathcal{A}\tilde{\sigma}_h - \varepsilon_h(\tilde{u}_h - Q_h \tilde{u}_h), \varepsilon_h(\tilde{u}_h - Q_h \tilde{u}_h))
\end{equation}
\begin{equation}
+ A_h(\tilde{\sigma}_h, \tilde{u}_h; 0, \tilde{u}_h - Q_h \tilde{u}_h).
\end{equation}

By the triangle inequality and (3.38), it holds that
\begin{equation}
\|\mathcal{A}\tilde{\sigma}_h - \varepsilon_h(\tilde{u}_h - Q_h \tilde{u}_h)\|_0 \leq \|\mathcal{A}\tilde{\sigma}_h\|_0 + \|\varepsilon_h(\tilde{u}_h - Q_h \tilde{u}_h)\|_0 \lesssim \|\tilde{\sigma}_h\|_0 + |Q_h \tilde{u}_h|_{1,h}
\end{equation}
\begin{equation}
\lesssim \sup_{(\tau_h, v_h) \in \Sigma_h \times V_h} A_h(\tilde{\sigma}_h, \tilde{u}_h; \tau_h, v_h) \|\tau_h\|_{0,h} + |v_h|_{1,h}.
\end{equation}
Due to (3.34), we have
\[
\mathcal{A}_h(\sigma_h, \bar{u}_h^*; 0, \bar{u}_h^* - Q_h \bar{u}_h^*) \leq \|\varepsilon_h(\bar{u}_h^* - Q_h \bar{u}_h^*)\|_0 \sup_{v_h^* \in V_h^*} \frac{\mathcal{A}_h(\bar{\sigma}_h, \bar{u}_h^*; 0, v_h^* - Q_h v_h^*)}{\|\varepsilon_h(v_h^* - Q_h v_h^*)\|_0}
\]
\[
\leq \|\varepsilon_h(\bar{u}_h^* - Q_h \bar{u}_h^*)\|_0 \sup_{v_h^* \in V_h^*, \varepsilon_h(v_h^* - Q_h v_h^*) \neq 0} \frac{\mathcal{A}_h(\bar{\sigma}_h, \bar{u}_h^*; 0, v_h^* - Q_h v_h^*)}{|v_h^* - Q_h v_h^*|_{1,h}}
\]
\[
\leq \|\varepsilon_h(\bar{u}_h^* - Q_h \bar{u}_h^*)\|_0 \sup_{(\tau_h, v_h^*) \in \Sigma_h \times V_h^*} \frac{\mathcal{A}_h(\bar{\sigma}_h, \bar{u}_h^*; \tau_h, v_h^*)}{\|\tau_h\|_{0,h} + |v_h^*|_{1,h}}.
\]

Using the last two inequalities and the Cauchy-Schwarz inequality, we get from (3.39) that
\[
\|\varepsilon_h(\bar{u}_h^* - Q_h \bar{u}_h^*)\|_0 \lesssim \sup_{(\tau_h, v_h^*) \in \Sigma_h \times V_h^*} \frac{\mathcal{A}_h(\bar{\sigma}_h, \bar{u}_h^*; \tau_h, v_h^*)}{\|\tau_h\|_{0,h} + |v_h^*|_{1,h}},
\]
which together with (3.34) implies
\[
(\bar{u}_h^* - Q_h \bar{u}_h^*)_{1,h} \lesssim \sup_{(\tau_h, v_h^*) \in \Sigma_h \times V_h^*} \frac{\mathcal{A}_h(\bar{\sigma}_h, \bar{u}_h^*; \tau_h, v_h^*)}{\|\tau_h\|_{0,h} + |v_h^*|_{1,h}}.
\]
Finally, we can finish the proof by combining (3.38) and (3.40).

**Theorem 3.15.** Assume that \(\sigma \in H^{k+1}(\Omega; \mathbb{S})\) and \(u \in H^{k+2}(\Omega; \mathbb{R}^n)\), then
\[
\|\sigma - \sigma_h\|_{0,h} + |u - u_h^*|_{1,h} \lesssim h^{k+1}(\|\sigma\|_{k+1} + \|u\|_{k+2}).
\]
Moreover, when \(\Omega\) is convex, we have
\[
\|u - u_h^*\|_0 \lesssim h^{k+2}(\|\sigma\|_{k+1} + \|u\|_{k+2}).
\]
**Proof.** By direct computation, we have
\[
\mathcal{A}_h(\sigma - I_h^{HZ} \sigma, u - Q_h^* u; \tau_h, v_h^*) = -(f, v_h^*) \quad \forall (\tau_h, v_h^*) \in \Sigma_h \times V_h^*.
\]
Combining with (3.31), we get the error equation
\[
(\sigma - \sigma_h, u - u_h^*; \tau_h, v_h^*) = (Q_h f - f, v_h^*) \quad \forall (\tau_h, v_h^*) \in \Sigma_h \times V_h^*.
\]
Let \(Q_h^*\) be the \(L^2\) orthogonal projection from \(L^2(\Omega; \mathbb{R}^n)\) onto \(V_h^*\). It holds from (3.23) that
\[
\mathcal{A}_h(\sigma - I_h^{HZ} \sigma, u - Q_h^* u; \tau_h, v_h^*) = a(\sigma - I_h^{HZ} \sigma, \tau_h) + b(\sigma - I_h^{HZ} \sigma, v_h^*) + (\varepsilon_h(u - Q_h^* u) - \mathcal{A}(\sigma - I_h^{HZ} \sigma), \varepsilon_h(v_h^* - Q_h v_h^*))
\]
\[
= a(\sigma - I_h^{HZ} \sigma, \tau_h) + (Q_h f - f, v_h^*) + (\varepsilon_h(u - Q_h^* u) - \mathcal{A}(\sigma - I_h^{HZ} \sigma), \varepsilon_h(v_h^* - Q_h v_h^*)).
\]
Then we obtain from (3.43), Cauchy-Schwarz inequality and the error estimates of \(I_h^{HZ}, Q_h^*\) and \(Q_h\) that
\[
\mathcal{A}_h(I_h^{HZ} \sigma - \sigma, u - u_h^*; \tau_h, v_h^*) = a(I_h^{HZ} \sigma - \sigma, \tau_h) + (\varepsilon_h(u - Q_h^* u) - \mathcal{A}(\sigma - I_h^{HZ} \sigma), \varepsilon_h(v_h^* - Q_h v_h^*))
\]
\[
\lesssim h^{k+1}(\|\sigma\|_{k+1} + \|u\|_{k+2})(\|\tau_h\|_0 + \|\varepsilon_h(v_h^*)\|_0).
Here we indeed have used the following estimate derived from the inverse inequality and (3.3)-(3.4) in [22]
\[
\|\varepsilon_h (v_h^* - Q_h v_h^*)\|_{0,K} \lesssim h_K^{-1} \|v_h^* - Q_h v_h^*\|_{0,K} \lesssim h_K^{-1} \|v_h^* - \pi_K v_h^*\|_{0,K} \lesssim \|\varepsilon_h (v_h^*)\|_{0,K}
\]
for any \(K \in T_h\). Applying the inf-sup condition (3.37), it follows that
\[
\|I_h^{HZ} \sigma - \sigma_h\|_{0,h} + |Q_h\ u - u_h^*|_{1,h} \lesssim h^{k+1} (\|\sigma\|_{k+1} + \|u\|_{k+2}).
\]
Hence we will achieve (3.41) by using the triangle inequality, and the error estimates of \(I_h^{HZ}\) and \(Q_h^*\).

When \(\Omega\) is convex, we have from the triangle inequality, the error estimate of \(Q_h\) and (3.29) that
\[
\|u - u_h^*\|_0 \leq \|(I - Q_h)(u - u_h^*)\|_0 + \|Q_h u - Q_h u_h^*\|_0 \lesssim h\|u - u_h^*|_{1,h} + \|Q_h u - u_h\|_0.
\]
Finally (3.42) is achieved by using (3.41) and (3.20). \(\square\)

Remark 3.16. The convergence rate of \(|u - u_h^*|_{1,h}\) is of second order higher than that of \(|u - u_h|_{1,h}\). This superconvergent postprocessed displacement \(u_h^*\) 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
\[
\|\tau_h\|_{0,h} \approx \|\tau_h\|_0 \quad \forall \ \tau_h \in \Sigma_h,
\]
which implies that we can use the weighted mass matrix \(M_h^\lambda\) with \(\lambda = 0\), i.e., \(M_h\).

For each \(v_h \in V_h\), denote by \(v_h\) the vector representation of \(v_h\) based on the basis of \(V_h\) 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., \(S_h := B_h M_h^{-1} B_h^T + C_h\). It is easy to see \(S_h\) is symmetric and positive definite (SPD) and induces a norm \(\|\cdot\|_{S_h}\) on \(V_h\), i.e.,
\[
\|v_h\|^2_{S_h} := v_h^T \tau_h \tau_h v_h \quad \forall \ v_h \in V_h.
\]

Lemma 4.1. We have the norm equivalence:
\[
|v_h|_{1,h} \approx \|v_h\|_{S_h} \quad \forall \ v_h \in V_h.
\]

Proof. We focus on the case \(k \geq n+1\) first. The low order case \(1 \leq k \leq n\) can be proved similarly by adding the stabilization term.

The inf-sup condition (4.1) implies \(B_h^T\) is injective and thus \(S_h\) is SPD and defines an inner product on \(V_h\). The identity
\[
(v_h^T S_h v_h)^{1/2} = \|M_h^{-1/2} B_h^T v_h\| = \sup_{\tau_h \in \Sigma_h} \frac{b(\tau_h, v_h)}{\|\tau_h\|_0} \quad \forall \ v_h \in V_h
\]
follows from the Riesz representation. Here \(\|\cdot\|\) denotes the Euclidean norm of a vector. The inequality \(|v_h|_{1,h} \lesssim \|v_h\|_{S_h}\) is a combination of (3.1), (4.1), and (4.2).
From an integration by parts, we can easily get $b(\tau_h, v_h) \lesssim \|\tau_h\|_{0,h} |v_h|_{1,h}$. Then the inequality $\|v_h\|_{S_h} \lesssim |v_h|_{1,h}$ follows from (4.1) and (4.2).

We define the operator $\mathcal{P}_h : \Sigma'_h \times \mathbf{V}'_h \rightarrow \Sigma_h \times \mathbf{V}_h$ with the matrix representation

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

and denote

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

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

$$\|\mathcal{P}_h \mathcal{L}_h^\lambda \|_{\Sigma_h \times \mathbf{V}_h \rightarrow \Sigma_h \times \mathbf{V}_h}, \| (\mathcal{P}_h \mathcal{L}_h^\lambda)^{-1} \|_{\Sigma_h \times \mathbf{V}_h \rightarrow \Sigma_h \times \mathbf{V}_h}$$

are bounded and independent of parameters $h$ and $\lambda$.

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

$$\begin{pmatrix} \tilde{M}_h & \tilde{B}_h^T \\ \tilde{B}_h & -\tilde{C}_h \end{pmatrix} \begin{pmatrix} I & \tilde{M}_h^{-1} \tilde{B}_h^T \\ 0 & -I \end{pmatrix} = \begin{pmatrix} \tilde{M}_h & 0 \\ \tilde{B}_h & \tilde{S}_h \end{pmatrix}$$

(4.4)

to obtain an approximate block-factorization preconditioner, where

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

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

We define the operator $\mathcal{G}_h : \Sigma'_h \times \mathbf{V}'_h \rightarrow \Sigma_h \times \mathbf{V}_h$ as

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

If we denote

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

it is trivial to verify that $\mathcal{G}_h = \tilde{\mathcal{L}}_h^{-1}$. Since $\tilde{M}_h$ 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

$$\|\mathcal{G}_h \mathcal{L}_h^\lambda \|_{\Sigma_h \times \mathbf{V}_h \rightarrow \Sigma_h \times \mathbf{V}_h}, \| (\mathcal{G}_h \mathcal{L}_h^\lambda)^{-1} \|_{\Sigma_h \times \mathbf{V}_h \rightarrow \Sigma_h \times \mathbf{V}_h}$$

are bounded and independent of parameters $h$ and $\lambda$.

**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 \( \zeta \) is an eigenvalue of \( \mathcal{G}_h \mathcal{L}_h^\lambda \), then there exists nonzero complex vector \((y, z)\) such that

\[
\begin{pmatrix} M_h^\lambda & B_h^T \\ B_h & -C_h \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} = \zeta \begin{pmatrix} M_h & B_h^T \\ B_h & -C_h \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix}.
\]

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

\[
\overline{z}^T B_h y = \overline{z}^T C_h z = \zeta (\overline{z}^T B_h y - \overline{z}^T C_h z).
\]

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

\[
\overline{y}^T B_h^T z = \overline{z}^T C_h z.
\]

Thus

\[
\overline{y}^T B_h^T z = \overline{z}^T C_h z.
\]

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

\[
\overline{y}^T M_h^\lambda y + \overline{y}^T B_h^T z = \zeta (\overline{y}^T M_h y + \overline{y}^T B_h^T z).
\]

Then we get from the last two equalities 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. Therefore we get from (4.7) that \( \zeta \) is positive. \( \square \)

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

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

\[ \zeta \approx 1. \]

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 \( \tau_h \in \Sigma_h \) satisfying \( \int_\Omega tr \tau_h \, dx = 0 \), we have

\[
\|\tau_h\|_0 \lesssim \|\tau_h\|_a + \|\text{div} \, \tau_h\|_{-1,h}.
\]
Proof. First we can see that

$$\|h \text{ div } \tau_h\| = \sup_{v_h \in V_h} \frac{\sum_{K \in T_h} (h_K \text{ div } \tau_h, v_h)_K}{\|v_h\|_0} = \sup_{v_h \in V_h} \frac{\text{div } \tau_h, v_h}{\left(\sum_{K \in T_h} \|h_K^{-1} v_h\|_{0,K}^2\right)^{1/2}}.$$ 

By the inverse inequality, it follows that

$$\|v_h\|_{1,h}^2 \leq \sum_{K \in T_h} \|h_K^{-1} v_h\|_{0,K}^2.$$ 

Hence

$$\|h \text{ div } \tau_h\| \leq \sup_{v_h \in V_h} \frac{\text{div } \tau_h, v_h}{|v_h|_{1,h}} = \|\text{div } \tau_h\|_{-1,h},$$

which together with Lemma 3.5 will end the proof. 

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

$$C_1 M_h \leq \hat{M}_h + B_h^T S_h^{-1} B_h. \tag{4.8}$$

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

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

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

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

Assumption (4.9) is easily satisfied by scaling $\hat{M}_h$ 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 $S_h$, take $\hat{M}_h = \omega_1 M_h$ and $S_{Fasp} = \omega_2 \hat{S}_h$ with constants $\omega_1$ and $\omega_2$ being independent of $h$, then (4.9) can be verified only if constants $\omega_1$ and $\omega_2$ are chosen large enough.

Obviously

$$(1 - \alpha') \hat{S}_h \leq \hat{S}_h \leq S_h. \tag{4.10}$$

$$\alpha \hat{M}_h \leq \hat{M}_h - M_h^{-1} \hat{M}_h. \tag{4.11}$$

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

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

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

By direct computation, we get

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

Let

$$\mathcal{D}_h := \begin{pmatrix} \hat{M}_h - M_h^\lambda & 0 \\ 0 & S_{Fasp} \end{pmatrix}.$$
Lemma 4.7 (Field-of-value equivalence). Assume (1.9) holds, then we have
\[
\frac{(\hat{\gamma}_h L^\alpha_h x, x)_{\mathcal{D}_h}}{(x,x)_{\mathcal{D}_h}} \geq \gamma \quad \text{and} \quad \frac{\|\hat{\gamma}_h L^\alpha_h x\|^2_{\mathcal{D}_h}}{\|x\|^2_{\mathcal{D}_h}} \leq \Gamma
\]
for all nonzero vectors \(x^T := (y^T, z^T)\), where
\[
\gamma := \min \left\{ C_1\alpha^3(1-\alpha'), \ C_1\alpha^2\beta'(1-\alpha')(1-\frac{\sqrt{3}}{2}), \beta'(1-\frac{\sqrt{3}}{2}) \right\},
\]
\[
\Gamma := \max \left\{ \frac{3}{\alpha^2}((1-\alpha)^2 + \beta^2) + \frac{2\beta}{\alpha}, \ \frac{3\beta}{\alpha} (1-\beta)^2 + 2\beta^2, \ \frac{3\beta}{\alpha} (1-\beta')^2 + 2\beta^2 \right\}.
\]
Here \(C_1\) is the constant from (4.8).

Proof. It follows from the Cauchy-Schwarz inequality and (4.9) that
\[
(B_h E_M y, S^{-1}_{Fasp} \tilde{S}_h z) \leq \|B_h E_M y\|_{S^{-1}_{Fasp}} \|\tilde{S}_h z\|_{S^{-1}_{Fasp}} \leq \sqrt{\beta} \|B_h E_M y\|_{S^{-1}_{Fasp}} \|z\|_{\tilde{S}_h} \leq \frac{\sqrt{\beta}}{2} (\|B_h E_M y\|_{S^{-1}_{Fasp}}^2 + \|z\|^2_{\tilde{S}_h}).
\]
Due to (4.12), it holds that
\[
(\hat{\gamma}_h L^\alpha_h x, x)_{\mathcal{D}_h} = (\hat{\gamma}_h L^\alpha_h y, y) + \|B_h E_M y\|_{S^{-1}_{Fasp}}^2 + \|z\|^2_{\tilde{S}_h} - (B_h E_M y, S^{-1}_{Fasp} \tilde{S}_h z)
\]
(4.14) \[\geq \alpha(M_h^\lambda y, y) + (1 - \frac{\sqrt{\beta}}{2}) (\|B_h E_M y\|_{S^{-1}_{Fasp}}^2 + \|z\|^2_{\tilde{S}_h}).\]
By (4.9) and (4.8),
\[
C_1(1-\alpha')\|E_M y\|_{M_h}^2 \leq C_1\|E_M y\|_{M_h}^2 \leq (M_h^\lambda E_M y, E_M y) + \|B_h E_M y\|_{S^{-1}_{Fasp}}^2.
\]
Combining with (4.11), (4.12) and (4.8), we get
\[
\|y\|_{M_h}^2 \leq \frac{1}{\alpha^2} \|E_M y\|_{M_h}^2 \leq \frac{1}{C_1\alpha^2(1-\alpha')} \left( (M_h^\lambda E_M y, E_M y) + \|B_h E_M y\|_{S^{-1}_{Fasp}}^2 \right) \leq \frac{1}{C_1\alpha^2(1-\alpha')} \left( (M_h^\lambda y, y) + \|B_h E_M y\|_{S^{-1}_{Fasp}}^2 \right) \leq \frac{1}{C_1\alpha^2(1-\alpha')} \left( (M_h^\lambda y, y) + \|B_h E_M y\|_{S^{-1}_{Fasp}}^2 \right).
\]
According to (4.9) and (4.10) and (4.14), we obtain
\[
\gamma(x,x)_{\mathcal{D}_h} \leq \gamma\|y\|_{M_h}^2 + \gamma\|z\|^2_{Fasp} \leq \gamma\|y\|_{M_h}^2 + \frac{\gamma}{\beta'}\|z\|^2_{\tilde{S}_h} \leq \frac{\gamma}{C_1\alpha^2(1-\alpha')}(M_h^\lambda y, y) + \frac{\gamma}{C_1\alpha^2(1-\alpha')\beta'}\|B_h E_M y\|_{S^{-1}_{Fasp}}^2 + \frac{\gamma}{\beta'}\|z\|^2_{\tilde{S}_h} \leq (\hat{\gamma}_h L^\alpha_h x, x)_{\mathcal{D}_h}.
\]
On the other hand, by (4.12) and (4.9),
\[
\|M_h^\lambda y\|_{M_h}^2 = (M_h^\lambda y, y) - (E_M^2 M_h^\lambda y, y) \leq (1-\alpha)(M_h^\lambda y, y) \leq (1-\alpha)^2\|y\|_{M_h}^2.
\]
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 \( B_h^T \tilde{S}_h^{-1} B_h y = \lambda M_s h y \) are not bigger than one. We get from (4.9) and (4.13) that
\[
\| B_h^T S_{Fasp}^{-1} B_h E_M y \|_{\tilde{M}_h}^2 \leq y^T E_M^T B_h^T S_{Fasp}^{-1} B_h \tilde{M}_h^{-1} B_h^T \tilde{S}_h^{-1} B_h E_M y \\
\leq y^T E_M^T B_h^T S_{Fasp}^{-1} \tilde{S}_h S_{Fasp}^{-1} B_h E_M y \\
\leq \beta y^T E_M^T B_h^T S_{Fasp}^{-1} B_h E_M y \\
\leq \beta^2 y^T E_M^T \tilde{M}_h E_M y \leq \beta^2 \| y \|_{\tilde{M}_h}^2.
\]

Similarly, we obtain from (4.9) and (4.13) that
\[
\| B_h^T (I - S_{Fasp}^{-1} \tilde{S}_h) z \|_{\tilde{M}_h}^2 \leq z^T (I - \tilde{S}_h S_{Fasp}^{-1}) B_h \tilde{M}_h^{-1} B_h^T (I - S_{Fasp}^{-1} \tilde{S}_h) z \\
\leq z^T (I - \tilde{S}_h S_{Fasp}^{-1}) \tilde{S}_h (I - S_{Fasp}^{-1} \tilde{S}_h) z \\
\leq \max \{(1 - \beta)^2, (1 - \beta')^2\} \| z \|_{\tilde{S}_h}^2,
\]
\[
\| B_h E_M y + \tilde{S}_h z \|_{\tilde{S}_h}^2 \leq 2\| B_h E_M y \|_{\tilde{S}_h}^2 + \| \tilde{S}_h z \|_{\tilde{S}_h}^2 \\
= 2y^T E_M^T B_h^T \tilde{S}_h^{-1} B_h E_M y + 2\| z \|_{\tilde{S}_h}^2 \\
\leq 2y^T E_M^T \tilde{M}_h E_M y + 2\| z \|_{\tilde{S}_h}^2 \leq 2\| y \|_{\tilde{M}_h}^2 + 2\| z \|_{\tilde{S}_h}^2.
\]

Combining the last four inequalities, we achieve from (4.9), (4.10), and (4.13) that
\[
\| \tilde{G}_h E_h x \|_{\tilde{D}_h}^2 \leq \| M_h^\lambda y + B_h^T S_{Fasp}^{-1} B_h E_M y + B_h^T (I - S_{Fasp}^{-1} \tilde{S}_h) z \|_{E_M (\tilde{M}_h - M_h^\lambda)^{-1} E_M}^2 \\
+ \| B_h E_M y + \tilde{S}_h z \|_{\tilde{S}_{Fasp}}^2 \\
\leq \frac{1}{\alpha} \| M_h^\lambda y + B_h^T S_{Fasp}^{-1} B_h E_M y + B_h^T (I - S_{Fasp}^{-1} \tilde{S}_h) z \|_{\tilde{M}_h}^2 \\\n+ \beta \| B_h E_M y + \tilde{S}_h z \|_{\tilde{S}_h}^2 \\
\leq \frac{3}{\alpha} \| M_h^\lambda y \|_{\tilde{M}_h}^2 + \frac{3}{\alpha} \| B_h^T S_{Fasp}^{-1} B_h E_M y \|_{\tilde{M}_h}^2 \\
+ \frac{3}{\alpha} \| B_h^T (I - S_{Fasp}^{-1} \tilde{S}_h) z \|_{\tilde{M}_h}^2 + \beta \| B_h E_M y + \tilde{S}_h z \|_{\tilde{S}_h}^2 \\
\leq (\frac{3}{\alpha} (1 - \beta)^2 + 2\beta) \| y \|_{\tilde{M}_h}^2 \\
+ (\frac{3}{\alpha} \max \{(1 - \beta)^2, (1 - \beta')^2\} + 2\beta) \| z \|_{\tilde{S}_h}^2 \leq \Gamma \| x \|_{\tilde{D}_h}^2,
\]
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 \( \| \cdot \|_{\tilde{D}_h} \)-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 \( \tilde{G}_h \) for linear system (2.5) converges uniformly in the \( \| \cdot \|_{\tilde{D}_h} \)-norm with respect to the parameters \( h \) and \( \lambda \).
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

\[ \mathcal{V}_h := \{ v \in H^1_0(\Omega; \mathbb{R}^n) : v|_K \in P_1(K; \mathbb{R}^n) \quad \forall K \in \mathcal{T}_h \} . \]

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

\[ |v|_{1,h} = \|\varepsilon(v_h)\|_0 + |v|_1 \quad \forall v_h \in \mathcal{V}_h. \]

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

\[ 2\mu(\varepsilon(u_h), \varepsilon(v_h)) = (f, v_h) \quad \forall v_h \in \mathcal{V}_h. \]

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

\[ (Aw_h, v_h) := 2\mu(\varepsilon(w_h), \varepsilon(v_h)) \quad \forall w_h, v_h \in \mathcal{V}_h. \]

It is apparent that the operator $A$ is SPD.

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

\[ (M\varsigma_h, \tau_h) = \frac{1}{2\mu}(\varsigma_h, \tau_h) \quad \forall \varsigma_h, \tau_h \in \Sigma_h, \]

operator $B : \Sigma_h \to \mathcal{V}_h$ by

\[ (B\tau_h, v_h) = b(\tau_h, v_h) \quad \forall \tau_h \in \Sigma_h, v_h \in \mathcal{V}_h, \]

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

\[ (Cw_h, v_h) = c(u_h, v_h) \quad \forall w_h, v_h \in \mathcal{V}_h. \]

Then the Schur complement operator $S = BM^{-1}B^T + C$. Based on the norm equivalence [51], we can easily derive the estimate of spectral radius and condition number of the Schur complement operator $S$ that

\[ \rho_S = \lambda_{\text{max}}(S) \approx h^{-2}, \quad \kappa(S) = \frac{\lambda_{\text{max}}(S)}{\lambda_{\text{min}}(S)} \approx h^{-2}. \]

The relation between $S$ and $S_h$ is given by

\[ S_h = M_{u,h}S \]

with $S$ 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 $\mathcal{V}_h$ as the “fine” space and $\mathcal{V}_h$ as the “coarse” space. Denote $B : \mathcal{V}_h \to \mathcal{V}_h$ 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 $R : \mathcal{V}_h \to \mathcal{V}_h$, 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 $\Pi : \mathcal{V}_h \to \mathcal{V}_h$. A “restriction” operator $\Pi^t : \mathcal{V}_h \to \mathcal{V}_h$ is consequently defined by

\[ (\Pi^t v, w) = (v, \Pi w) \quad \text{for } v \in \mathcal{V}_h \text{ and } w \in \mathcal{V}_h. \]
It is also well-known that the matrix representation of the restriction operator $\Pi^t$ is just the transpose of the matrix representation of the prolongation operator $\Pi$. Then, the auxiliary space preconditioner $X : V_h \rightarrow V_h$, following the definition in \cite{79}, is given by

\begin{align}
\text{(5.4)} & \quad \text{Additive} \quad X = R + \Pi S \Pi^t, \\
\text{(5.5)} & \quad \text{Multiplicative} \quad I - XS = (I - R^t S)(I - \Pi S \Pi^t S)(I - RS).
\end{align}

According to \cite{79}, the following theorem holds.

**Theorem 5.1** (Xu \cite{79}). Assume that for all $v \in V_h$, $w \in V_h$,

\begin{align}
(5.6) & \quad (Sv, v) \lesssim (R^{-1} v, v) \lesssim \rho_S(v, v), \\
(5.7) & \quad (Aw, w) \lesssim (BAw, Aw) \lesssim (Aw, w), \\
(5.8) & \quad \|w\|_1, h \lesssim |w|_1 \quad \text{(stability of $\Pi$)},
\end{align}

and furthermore, assume that there exists a linear operator $P : V_h \rightarrow V_h$ such that

\begin{align}
(5.9) & \quad |Pv|_1 \lesssim |v|_1, h \quad \text{(stability of $P$)}, \\
(5.10) & \quad \|v - \Pi P v\|_0^2 \lesssim \rho^{-1}_S |v|_{1, h}^2 \quad \text{(approximability)}.
\end{align}

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 $B$ that satisfies condition (5.7). For example, $B$ can be either the direct solver, for which $B \sim A^{-1}$, 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 \cite{18} will satisfy the condition (5.6). Operator $\Pi$ is the natural inclusion for $k \geq 2$ and the $L^2$ projection $Q_h^k$ 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 $P : V_h \rightarrow V_h$ 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 $B$, $R$, and $\Pi$ only.

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

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

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

Define

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

Obviously for interior nodes we have (cf. \cite{21})

$$\text{(5.11)} \quad (Pv)(x_i) = \sum_{K \in \mathcal{T}_{h,i}} \frac{|K|}{|\Omega_i|} Q^0_K v.$$
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

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

**Proof.** According to (5.11), it holds for each interior node $x_i$ that

$$
|Q^0_K v - (P v)(x_i)|^2 \lesssim \sum_{K \in T_h} |Q^0_K v - Q^0_K v|_2^2 \lesssim \sum_{F \in F_{h,i}} \|Q^0_K v\|^2.
$$

For each boundary node $x_i$, we obtain by a similar technique and the definition of jump on the boundary,

$$
|Q^0_K v - (P v)(x_i)|^2 \lesssim \sum_{F \in F_{h,i}} \|Q^0_K v\|^2.
$$

Then using the scaling argument, we have

$$
\sum_{K \in T_h} h^{-2}_K \|Q^0_K v - P v\|^2_{0,K} = \sum_{K \in T_h} \sum_{i=0}^n h^{-2}_K \|Q^0_K v - (P v)(x_{K,i})\|^2 \lesssim \sum_{F \in F_h} h^{-1}_F \|Q^0_K v\|^2_{0,F}.
$$

From the $L^2$ error estimate (3.14), the discrete Korn’s inequality (3.8), and the norm equivalence (3.9), we get

$$
\sum_{K \in T_h} h^{-2}_K \|v - P v\|^2_{0,K} \lesssim \sum_{K \in T_h} h^{-2}_K \|v - Q^0_K v\|^2_{0,K} + \sum_{K \in T_h} h^{-2}_K \|Q^0_K v - P v\|^2_{0,K} \lesssim \|v\|^2_{1,h} + \sum_{F \in F_h} h^{-1}_F \|Q^0_K v - v\|^2_{0,F} \lesssim \|v\|^2_{1,h}.
$$

(5.12)

It follows from (5.12) and (3.15) that

$$
|P v|^2 = \sum_{K \in T_h} |P v - Q^0_K v|^2_{1,K} \lesssim \sum_{K \in T_h} h^{-2}_K \|P v - Q^0_K v\|^2_{0,K} \lesssim \sum_{K \in T_h} h^{-2}_K \|v - P v\|^2_{0,K} + \sum_{K \in T_h} h^{-2}_K \|v - Q^0_K v\|^2_{0,K} \lesssim \|v\|^2_{1,h}.
$$

Therefore we can finish the proof by combining the last two inequalities. \hfill \Box

**Lemma 5.3.** For any $v \in V_h$, it holds that

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

**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

$$
\|v - \Pi P v\|_0^2 = \|v - Q_h P v\|_0^2 \lesssim \|v - P v\|_0^2 + \|P v - Q_h P v\|_0^2 \lesssim \|v - P v\|_0^2 + h^2 |P v|^2 \lesssim h^2 |v|^2_{1,h} \lesssim \rho_S^{-1} \|v\|^2_{1,h},
$$

as required. \hfill \Box

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

Theorem 5.4. Let \( R \) be a Jacobi or a symmetric Gauss-Seidel smoother, let \( B \) be one step of classical multigrid iteration, and let \( \Pi = Q_h \). 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 \( \Omega = (-1, 1)^2 \), \( \mu = 0 \), and the load \( f = 1 \). We use the uniform triangulation \( T_h \) of \( \Omega \) 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

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

where \( D_h \) is 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 \( B \).

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 \( \lambda \). 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 \( \omega_1 \) and \( \omega_2 \).

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

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

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

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

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

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

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

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

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 $k = 1$

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

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

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

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

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

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

$$
\sum_{K \in T_h} h_K^4 \| \text{rot} (\mathbb{A} \sigma_h) \|_{0,K}^2 + \sum_{K \in T_h} h_K^2 \| t^T ((\mathbb{A} \sigma_h) t) \|_{0,\partial K}^2 + \sum_{K \in T_h} h_K^2 \| \partial_i (\nu^T ((\mathbb{A} \sigma_h) t) - t^T \text{rot} (\mathbb{A} \sigma_h)) \|_{0,\partial K}^2
$$
Table 7. The iteration steps and CPU time (in seconds) of the approximate block-factorization preconditioned GMRES method for $k = 4$

<table>
<thead>
<tr>
<th>#dofs</th>
<th>$\lambda = 0$</th>
<th>$\lambda = 10$</th>
<th>$\lambda = 100$</th>
<th>$\lambda = 1000$</th>
<th>$\lambda = +\infty$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>steps</td>
<td>time</td>
<td>steps</td>
<td>time</td>
<td>steps</td>
</tr>
<tr>
<td>427</td>
<td>26</td>
<td>0.02</td>
<td>34</td>
<td>0.03</td>
<td>32</td>
</tr>
<tr>
<td>1627</td>
<td>28</td>
<td>0.14</td>
<td>35</td>
<td>0.17</td>
<td>36</td>
</tr>
<tr>
<td>6355</td>
<td>28</td>
<td>0.67</td>
<td>36</td>
<td>0.86</td>
<td>37</td>
</tr>
<tr>
<td>25123</td>
<td>29</td>
<td>2.90</td>
<td>37</td>
<td>3.70</td>
<td>38</td>
</tr>
<tr>
<td>99907</td>
<td>30</td>
<td>12.2</td>
<td>38</td>
<td>15.5</td>
<td>40</td>
</tr>
<tr>
<td>398467</td>
<td>31</td>
<td>50.9</td>
<td>40</td>
<td>65.8</td>
<td>42</td>
</tr>
</tbody>
</table>

To design the adaptive algorithm. The Dörfler marking strategy with bulk parameter $\theta = 0.1$ and the newest vertex bisection are employed in the adaptive algorithm. For consideration of unstructured meshes, we replace the auxiliary discrete problem with the following vectorial Poisson equation discretized by the conforming linear element: Find $u_h \in V_h$ such that

$$2\mu(\nabla u_h, \nabla v_h) = (f, v_h) \quad \forall v_h \in V_h.$$ 

This discrete problem will be solved efficiently by algebraic multigrid method. 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 $\lambda$ 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

<table>
<thead>
<tr>
<th>k = 2</th>
<th>k = 3</th>
<th>k = 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>#dofs</td>
<td>steps</td>
<td>#dofs</td>
</tr>
<tr>
<td>6226</td>
<td>34</td>
<td>5531</td>
</tr>
<tr>
<td>35190</td>
<td>35</td>
<td>36026</td>
</tr>
<tr>
<td>175532</td>
<td>40</td>
<td>158348</td>
</tr>
<tr>
<td>386054</td>
<td>42</td>
<td>309797</td>
</tr>
<tr>
<td>1102619</td>
<td>46</td>
<td>805682</td>
</tr>
<tr>
<td>1394239</td>
<td>46</td>
<td>1028794</td>
</tr>
<tr>
<td>1756771</td>
<td>49</td>
<td>1311058</td>
</tr>
<tr>
<td>2198618</td>
<td>43</td>
<td>1669303</td>
</tr>
</tbody>
</table>

References


