

# TAYLOR-HOOD LIKE FINITE ELEMENTS FOR NEARLY INCOMPRESSIBLE STRAIN GRADIENT ELASTICITY PROBLEMS

YULEI LIAO, PINGBING MING, AND YUN XU

**ABSTRACT.** We propose a family of mixed finite elements that are robust for the nearly incompressible strain gradient model, which is a fourth-order singular perturbed elliptic system. The element is similar to [C. Taylor and P. Hood, *Comput. & Fluids*, 1(1973), 73-100] in the Stokes flow. Using a uniform discrete B-B inequality for the mixed finite element pairs, we show the optimal rate of convergence that is robust in the incompressible limit. By a new regularity result that is uniform in both the materials parameter and the incompressibility, we prove the method converges with 1/2 order to the solution with strong boundary layer effects. Moreover, we estimate the convergence rate of the numerical solution to the unperturbed second-order elliptic system. Numerical results for both smooth solutions and the solutions with sharp layers confirm the theoretical prediction.

## 1. INTRODUCTION

The strain gradient models have drawn great attention recently because they capture the size effect of nano-materials for plasticity [FH97] as well as for mechanical meta-materials [KN20] by incorporating the higher-order strain gradient and the intrinsic material length scale into the constitutive relations. Studies from the perspective of modeling, mechanics and mathematics may date back to 1960s [Koi64, Min64, HH69], while large-scale simulations are relatively recent [PZV09, ZMKZ12, RVdVG14, PB15]. Different methods such as  $H^2$  conforming finite element methods [ZPV09, FMS10],  $H^1$  conforming mixed finite element methods [AA02a, PB15], nonconforming finite element methods [LMS17, LM19, LMW21], discontinuous Galerkin methods [EGH<sup>+</sup>02], isogeometric analysis [FKM<sup>+</sup>11, NKBN16], and meshless methods [AA02b] have been used to simulate the strain gradient elastic models with different complexity, just to mention a few. One of the numerical difficulties is that the number of the materials parameters is too large [Min64], another is that the materials parameters may cause boundary layer or numerical instability when they tend to certain critical values [EGH<sup>+</sup>02].

The strain gradient elasticity model proposed by Altan and Aifantis [AA92] seems the simplest one among them because it contains only one material parameter besides the Lamé constants, while it still models the size effect adequately [Aif99]. This model is described by the following boundary value problem:

$$(1.1) \quad \begin{cases} (\iota^2 \Delta - \mathbf{I})(\mu \Delta \mathbf{u} + (\lambda + \mu) \nabla \operatorname{div} \mathbf{u}) = \mathbf{f} & \text{in } \Omega, \\ \mathbf{u} = \partial_{\mathbf{n}} \mathbf{u} = \mathbf{0} & \text{on } \partial \Omega, \end{cases}$$

where  $\Omega \subset \mathbb{R}^2$  is a smooth domain,  $\mathbf{u} : \Omega \rightarrow \mathbb{R}^2$  is the displacement,  $\partial_{\mathbf{n}} \mathbf{u}$  is the normal derivative of  $\mathbf{u}$ ,  $\lambda$  and  $\mu$  are the Lamé constants, and  $\iota$  is the microscopic parameter such that  $0 < \iota \leq 1$ , which stands for the intrinsic length scale. Besides modeling the strain gradient elasticity, the system (1.1) may also be regarded as a vector analog of the fourth-order singular perturbed problem, which usually models

---

*Date:* January 27, 2023.

*2020 Mathematics Subject Classification.* Primary 65N15, 65N30; Secondary 74K20.

*Key words and phrases.* Mixed finite elements; Nearly incompressible strain gradient elasticity; Uniform error estimate.

The work of Liao and Ming were supported by National Natural Science Foundation of China through Grant No. 11971467. The work of Xu was supported by National Natural Science Foundation of China through Grant No. 11772067.

a clamped plate problem [Joh72, Sch76, Sem92, Sem94, NTW01, BN11], or arises from a fourth-order perturbation of the fully nonlinear Monge-Ampère equation [BGNS11, FN14]. System of the form (1.1) may also come from the linearized model in MEMS [LW17].

In the present work, we are interested in (1.1) for the nearly incompressible materials. Such materials are commonly used in industry and a typical example is natural rubber. To the best of our knowledge, the studies on the approximation of incompressible and nearly incompressible strain gradient elasticity have not been sufficiently addressed in the literature, although vast efforts have been devoted to finite element approximation of the incompressible and nearly incompressible elasticity problems; See e.g., [Her65, Vog83, STP85, BS92a, BS92b, BM05, ABdVL<sup>+</sup>13]. In [FH97, §III. C], the authors studied the incompressible limit of the strain gradient model. Mixed finite elements for the incompressible Fleck-Hutchinson strain gradient model have been designed and tested in [SKF99]. A finite element method has been tested for the nearly incompressible strain gradient model in [Wei06]. A mixed finite element, which approximated the displacement with Bogner-Fox-Schmidt element [BFS65] and approximated the pressure with the 9-node quadrilateral element, was constructed for the coupled stress model in [Fis11], and bore certain similarities with problem (1.1). Recently, Hu and Tian [Tia21] have proposed several robust elements for the two-dimensional strain gradient model in the framework of reduced integration. Unfortunately, none of the above work proved the robustness of the proposed elements rigorously in the incompressible limit.

Following the classical approach dealing with the nearly incompressible elasticity problem [Her65], we introduce an auxiliary variable “pressure”  $p$  and recast (1.1) into a displacement-pressure mixed variational problem, i.e., the so-called  $(\mathbf{u}, p)$ -formulation. We approximate the displacement by augmenting the finite element space in [GLN12] with certain new bubble functions. The original motivation for the bubble functions is to design the stable finite element pair for the Stokes problem [ABF84, BG85]. The augmented bubble functions help out in dealing with the extra constraints such as the divergence stability in Stokes problem and the high order consistency error [NTW01, GLN12, WZZ12]. Such idea has been exploited by one of the authors to design robust finite elements for the strain gradient elasticity model [LMS17]. Besides, we employ the standard continuous Lagrangian finite element of one order lower than that for the displacement to approximate the pressure. Such a finite element pair is similar to the Taylor-Hood element in the Stokes flows [TH73] which is  $\mathbb{P}_r - \mathbb{P}_{r-1}$  scheme and continuous pressure approximation. For both smooth solutions and solutions with strong boundary layer effects, these mixed finite element pairs are robust in the incompressible limit, here the robustness is understood in the sense that the rate of convergence is uniform in both  $\iota$  and  $\lambda$ . The bubble function spaces in approximating the displacement are defined by certain orthogonal constraints, and the explicit representations of these spaces are desired for the sake of implementation. We achieve this with the aid of the Jacobi polynomial. In addition to perspicuous results in view of analytics, such representation lends itself to the construction of the analytical shape functions for the approximating space of the displacement. Though we focus on the two-dimensional problem, the element may be readily extended to the three-dimensional problem. cf., Remark 3.3.

By standard mixed finite element theory [BBF13], a discrete B-B inequality that is uniform in  $\iota$  is needed for the well-posedness of the mixed  $(\mathbf{u}, p)$ -discretization problem. This B-B inequality reduces to the remarkable B-B inequality for the Stokes problem when  $\iota$  tends to zero. A natural way to prove the discrete B-B inequality is to construct a uniformly stable Fortin operator [For77, MSW13, MTW02], which does not seem easy due to the complication of the constraints. To this end, we construct a quasi-Fortin operator that takes different forms for small  $\iota$  as well as large  $\iota$ . This quasi-Fortin operator is bounded in a weighted energy norm in the corresponding regimes of  $\iota$ . Besides the discrete B-B inequality, another ingredient in proving the robustness is a new regularity result for the solution of (1.1) that is uniform in

both  $\lambda$  and  $\iota$ , which is crucial to prove the convergence rate for the layered solution. The proof combines the method dealing with the nearly incompressible linear elasticity [Vog83] and the regularity estimate for the fourth-order singular perturbed problem [NTW01, LMW21].

The outline of the paper is as follows. In §2, we introduce Altan and Aifantis' strain gradient model and its mixed variational formulation. We demonstrate the numerical difficulty caused by large  $\lambda$ , and prove the uniform regularity estimate for problem (1.1). In §3, we construct a family of nonconforming finite elements, and derive the explicit formulations for the bubble spaces. In §4, we use the nonconforming elements proposed in §3 together with the continuous Lagrangian finite elements to discretize the mixed variational problem and prove the optimal rate of convergence. In the last section, we report the numerical results, which are consistent with the theoretical prediction.

Throughout this paper, the constant  $C$  may differ from line to line, while it is independent of the mesh size  $h$ , the materials parameter  $\iota$  and the Lamé constant  $\lambda$ .

## 2. THE MIXED VARIATIONAL FORMULATION AND REGULARITY ESTIMATES

First we fix some notations. The space  $L^2(\Omega)$  of the square-integrable functions defined on a smooth domain  $\Omega$  is equipped with the inner product  $(\cdot, \cdot)$  and the norm  $\|\cdot\|_{L^2(\Omega)}$ , while  $L_0^2(\Omega)$  is the subspace of  $L^2(\Omega)$  with mean value zero. Let  $H^m(\Omega)$  be the standard Sobolev space [AF03] with the norm  $\|\cdot\|_{H^m(\Omega)}$ , while  $H_0^m(\Omega)$  is the closure in  $H^m(\Omega)$  of  $C_0^\infty(\Omega)$ . We may drop  $\Omega$  in  $\|\cdot\|_{H^m(\Omega)}$  when no confusion may occur. For any vector-valued function  $\mathbf{v}$ , its gradient is a matrix-valued function with components  $(\nabla \mathbf{v})_{ij} = \partial_j v_i$ , and the symmetric part of  $\nabla \mathbf{v}$  is defined by  $\boldsymbol{\epsilon}(\mathbf{v}) = (\nabla \mathbf{v} + [\nabla \mathbf{v}]^T)/2$ . The divergence operator is defined as  $\operatorname{div} \mathbf{v} = \partial_1 v_1 + \partial_2 v_2$ . The Sobolev spaces  $[H^m(\Omega)]^2$ ,  $[H_0^m(\Omega)]^2$  and  $[L^2(\Omega)]^2$  of a vector-valued function may be defined similarly as their scalar counterpart. This rule equally applies to the inner products and the norms. The double inner product between tensors  $\mathbf{A} = (A_{ij})_{i,j=1}^2$ ,  $\mathbf{B} = (B_{ij})_{i,j=1}^2$  equals  $\mathbf{A} : \mathbf{B} = \sum_{i,j=1}^2 A_{ij} B_{ij}$ .

We recast (1.1) into a variational problem: Find  $\mathbf{u} \in V := [H_0^2(\Omega)]^2$  such that

$$(2.1) \quad a(\mathbf{u}, \mathbf{v}) = (\mathbf{f}, \mathbf{v}) \quad \text{for all } \mathbf{v} \in V,$$

where  $a(\mathbf{u}, \mathbf{v}) := (\mathbb{C}\boldsymbol{\epsilon}(\mathbf{u}), \boldsymbol{\epsilon}(\mathbf{v})) + \iota^2(\mathbb{D}\nabla\boldsymbol{\epsilon}(\mathbf{u}), \nabla\boldsymbol{\epsilon}(\mathbf{v}))$ , and the fourth-order tensor  $\mathbb{C}$  and the sixth-order tensor  $\mathbb{D}$  are defined as

$$\mathbb{C}_{ijkl} = \lambda\delta_{ij}\delta_{kl} + 2\mu\delta_{ik}\delta_{jl} \quad \text{and} \quad \mathbb{D}_{ijklmn} = \lambda\delta_{il}\delta_{jk}\delta_{mn} + 2\mu\delta_{il}\delta_{jm}\delta_{kn},$$

respectively. Here  $\delta_{ij}$  is the Kronecker delta function. The strain gradient  $\nabla\boldsymbol{\epsilon}(\mathbf{v})$  is a third-order tensor that is defined by  $(\nabla\boldsymbol{\epsilon}(\mathbf{v}))_{ijk} = \partial_i(\boldsymbol{\epsilon}(\mathbf{v}))_{jk}$ .

We are interested in the regime when  $\lambda \rightarrow \infty$ , which means that the material is nearly incompressible. Proceeding along the same line that leads to [LMW21, Theorem 5], the tensor product of the element (NTW) proposed in [NTW01] may be used to approximate (1.1), and the error estimate reads as

$$\|\mathbf{u} - \mathbf{u}_h\| \leq C\lambda(h^2 + \iota h)\|\mathbf{u}\|_{H^3},$$

where  $\|\mathbf{v}\|^2 := a(\mathbf{v}, \mathbf{v})$ , and  $C$  is independent of the mesh size  $h$ , and  $\iota$  and  $\lambda$ . Therefore, the error bound degenerates when  $\lambda$  is large, and the NTW element does not seem a good candidate for the nearly incompressible material. The following numerical example confirms this observation.

**Example 2.1.** Let  $\Omega = (0, 1)^2$ , and  $\mathbf{u} = (u_1, u_2)$  with

$$u_1 = -\sin^3(\pi x) \sin(2\pi y) \sin(\pi y), \quad u_2 = \sin(2\pi x) \sin(\pi x) \sin^3(\pi y).$$

It is clear that  $\operatorname{div} \mathbf{u} = 0$ , hence the material is completely incompressible. The details of the numerical experiment such as the mesh generation, are the same as those in § 5. The relative error  $\|\mathbf{u} - \mathbf{u}_h\|/\|\mathbf{u}\|$  in Table 1 shows that the rate of convergence deteriorates when  $\lambda$  is large.

TABLE 1. Relative errors and rate of convergence for NTW

$\iota \backslash h$	1/8	1/16	1/32	1/64
$\nu = 0.3000, \lambda = 0.5769, \mu = 0.3846$				
1e+00	2.681e-01	1.373e-01	6.698e-02	3.334e-02
rate		0.97	1.04	1.01
1e-06	4.550e-02	1.244e-02	3.001e-03	7.467e-04
rate		1.87	2.05	2.01
$\nu = 0.499999, \lambda = 1.6667e5, \mu = 0.3333$				
1e+00	9.995e-01	9.979e-01	9.916e-01	9.682e-01
rate		0.00	0.01	0.03
1e-06	9.752e-01	7.233e-01	2.502e-01	6.561e-02
rate		0.43	1.53	1.93

**2.1. The mixed variational formulation.** We introduce an auxiliary variable  $p = \lambda \operatorname{div} \mathbf{u}$ , and  $p \in P := L_0^2(\Omega) \cap H_0^1(\Omega)$ . We write Problem (2.1) into a mixed variational problem as

$$(2.2) \quad \begin{cases} a_\iota(\mathbf{u}, \mathbf{v}) + b_\iota(\mathbf{v}, p) = (\mathbf{f}, \mathbf{v}) & \text{for all } \mathbf{v} \in V, \\ b_\iota(\mathbf{u}, q) - \lambda^{-1} c_\iota(p, q) = 0 & \text{for all } q \in P, \end{cases}$$

where

$$\begin{aligned} a_\iota(\mathbf{v}, \mathbf{w}) &:= 2\mu \left( (\boldsymbol{\epsilon}(\mathbf{v}), \boldsymbol{\epsilon}(\mathbf{w})) + \iota^2 (\nabla \boldsymbol{\epsilon}(\mathbf{v}), \nabla \boldsymbol{\epsilon}(\mathbf{w})) \right), & \mathbf{v}, \mathbf{w} \in V, \\ b_\iota(\mathbf{v}, q) &:= (\operatorname{div} \mathbf{v}, q) + \iota^2 (\nabla \operatorname{div} \mathbf{v}, \nabla q), & \mathbf{v} \in V, q \in P, \\ c_\iota(s, q) &:= (s, q) + \iota^2 (\nabla s, \nabla q), & s, q \in P. \end{aligned}$$

It is convenient to define the weighted norm for all  $q \in P$  as

$$\|q\|_\iota := \|q\|_{L^2} + \iota \|\nabla q\|_{L^2}.$$

$\|q\|_\iota$  is a norm over  $P$  for any  $q \in P$  and any finite  $\iota$ . By Poincaré's inequality,  $\|\nabla \mathbf{v}\|_\iota$  is a norm over  $V$  for any  $\mathbf{v} \in V$ . To study the well-posedness of Problem (2.2), we start with the following B-B inequality that is uniform for any  $\iota$ .

**Lemma 2.2.** *For any  $q \in P$ , there exists  $\mathbf{v} \in V$  such that*

$$(2.3) \quad \operatorname{div} \mathbf{v} = q \quad \text{and} \quad \|\nabla \mathbf{v}\|_\iota \leq C \|q\|_\iota,$$

where  $C$  only depends on  $\Omega$  but is independent of  $\iota$ .

*Proof.* By [Gal11, Theorem III 3.3] and [CM10, Proposition 4.1], for any  $q \in P$ , there exists  $\mathbf{v} \in V$  such that  $\operatorname{div} \mathbf{v} = q$  and

$$(2.4) \quad \|\mathbf{v}\|_{H^1} \leq C \|q\|_{L^2} \quad \text{and} \quad \|\mathbf{v}\|_{H^2} \leq C \|q\|_{H^1},$$

where the constant  $C$  only depends on  $\Omega$ .

Because the mean of  $q$  is zero for any  $q \in P$ , by Poincaré's inequality, there exists  $C$  such that

$$\|q\|_{H^1} \leq C \|\nabla q\|_{L^2}.$$

Combining the above two inequalities, we obtain

$$\|\nabla \mathbf{v}\|_\iota = \|\nabla \mathbf{v}\|_{L^2} + \iota \|\nabla^2 \mathbf{v}\|_{L^2} \leq C \|q\|_\iota.$$

This gives (2.3). □

**Lemma 2.3.** *There exists a unique  $\mathbf{u} \in V$  and  $p \in P$  satisfying (2.2), and there exists  $C$  independent of  $\iota$  and  $\lambda$  such that*

$$(2.5) \quad \|\nabla \mathbf{u}\|_\iota + \|p\|_\iota \leq C \|\mathbf{f}\|_{H^{-1}}.$$

*Proof.* By the first Korn's inequality [Kor08, Kor09],

$$(2.6) \quad \|\boldsymbol{\epsilon}(\mathbf{v})\|_{L^2}^2 \geq \frac{1}{2} \|\nabla \mathbf{v}\|_{L^2}^2 \quad \text{for all } \mathbf{v} \in [H_0^1(\Omega)]^2,$$

and the  $H^2$  Korn's inequality [LMW21, Theorem 1],

$$(2.7) \quad \|\nabla \boldsymbol{\epsilon}(\mathbf{v})\|_{L^2}^2 \geq \left(1 - \frac{1}{\sqrt{2}}\right) \|\nabla^2 \mathbf{v}\|_{L^2}^2 \quad \text{for all } \mathbf{v} \in [H^2(\Omega)]^2,$$

we obtain

$$a_\iota(\mathbf{v}, \mathbf{v}) \geq 2\mu \left( \frac{1}{2} \|\nabla \mathbf{v}\|_{L^2}^2 + \left(1 - \frac{1}{\sqrt{2}}\right) \iota^2 \|\nabla^2 \mathbf{v}\|_{L^2}^2 \right) \geq \frac{\mu}{2} \|\nabla \mathbf{v}\|_\iota^2.$$

Using (2.3), for any  $p \in P$ , there exists  $\mathbf{v}_0 \in V$  such that  $\operatorname{div} \mathbf{v}_0 = p$  and  $\|\nabla \mathbf{v}_0\|_\iota \leq C \|p\|_\iota$ . This implies

$$\sup_{\mathbf{v} \in V} \frac{b_\iota(\mathbf{v}, p)}{\|\nabla \mathbf{v}\|_\iota} \geq \frac{b_\iota(\mathbf{v}_0, p)}{\|\nabla \mathbf{v}_0\|_\iota} = \frac{\|p\|_\iota^2}{\|\nabla \mathbf{v}_0\|_\iota} \geq C \|p\|_\iota.$$

By [Bra96, Theorem 2], we immediately obtain the well-posedness of (2.2) and the estimate (2.5) by noting

$$|(\mathbf{f}, \mathbf{v})| \leq \|\mathbf{f}\|_{H^{-1}} \|\mathbf{v}\|_{H^1} \leq C \|\mathbf{f}\|_{H^{-1}} \|\nabla \mathbf{v}\|_{L^2} \leq C \|\mathbf{f}\|_{H^{-1}} \|\nabla \mathbf{v}\|_\iota.$$

□

By the standard regularity theory for the elliptic system, we find  $\mathbf{u} \in [H^4(\Omega)]^2$  and  $p \in H^3(\Omega)$  provided that  $\mathbf{f} \in [L^2(\Omega)]^2$ , while we are interested in whether the shift estimate  $\|\nabla^2 \mathbf{u}\|_\iota + \|\nabla p\|_\iota \leq C(\iota) \|\mathbf{f}\|_{L^2}$  is true with a  $\lambda$ -independent constant  $C(\iota)$ , this is the objective of the next part.

**2.2. Regularity estimates.** We aim to study the regularity of the solution of (1.1). Letting  $\iota \rightarrow 0$ , we find  $\mathbf{u}_0 \in [H_0^1(\Omega)]^2$  satisfying

$$(2.8) \quad -\mathcal{L}\mathbf{u}_0 = \mathbf{f} \quad \text{in } \Omega, \quad \mathbf{u}_0 = \mathbf{0} \quad \text{on } \partial\Omega,$$

in the sense of distribution, where  $\mathcal{L}\mathbf{u}_0 := \mu \Delta \mathbf{u}_0 + (\lambda + \mu) \nabla \operatorname{div} \mathbf{u}_0$ . The  $H^1$ -error for  $\mathbf{u} - \mathbf{u}_0$  will be given in Theorem 2.9, which is crucial for the regularity estimate of problem (1.1). We reshape (2.8) into a variational problem: Find  $\mathbf{u}_0 \in [H_0^1(\Omega)]^2$  such that

$$(2.9) \quad (\mathbb{C}\boldsymbol{\epsilon}(\mathbf{u}_0), \boldsymbol{\epsilon}(\mathbf{v})) = (\mathbf{f}, \mathbf{v}) \quad \text{for all } \mathbf{v} \in [H_0^1(\Omega)]^2.$$

By [BB03], we have the following shift estimate for  $\mathbf{u}_0$ : There exists  $C$  independent of  $\lambda$  such that

$$(2.10) \quad \|\mathbf{u}_0\|_{H^2} + \lambda \|\operatorname{div} \mathbf{u}_0\|_{H^1} \leq C \|\mathbf{f}\|_{L^2}.$$

Next we study an auxiliary boundary value problem:

$$(2.11) \quad \begin{cases} \Delta \mathcal{L}\mathbf{w} = \mathbf{F}, & \text{in } \Omega, \\ \mathbf{w} = \partial_n \mathbf{w} = \mathbf{0}, & \text{on } \partial\Omega. \end{cases}$$

The a-priori estimate for the solution of the above boundary value problem reads as

**Lemma 2.4.** *There exists a unique  $\mathbf{w} \in V$  satisfying (2.11), and there exists  $C$  independent of  $\lambda$  such that*

$$(2.12) \quad \|\mathbf{w}\|_{H^2} + \lambda \|\operatorname{div} \mathbf{w}\|_{H^1} \leq C \|\mathbf{F}\|_{H^{-2}}.$$

*Proof.* We recast (2.11) into a variational problem: Find  $\mathbf{w} \in V$  such that

$$A(\mathbf{w}, \mathbf{v}) = (\mathbf{F}, \mathbf{v}) \quad \text{for all } \mathbf{v} \in V,$$

where  $A(\mathbf{v}, \mathbf{z}) := 2\mu(\nabla\epsilon(\mathbf{v}), \nabla\epsilon(\mathbf{z})) + \lambda(\nabla\operatorname{div} \mathbf{v}, \nabla\operatorname{div} \mathbf{z})$  for any  $\mathbf{v}, \mathbf{z} \in V$ .

For any  $\mathbf{v} \in V$ , by the  $H^2$ -Korn's inequality (2.7) and Poincaré's inequality, there exists  $C$  such that

$$A(\mathbf{v}, \mathbf{v}) \geq 2\mu \|\nabla\epsilon(\mathbf{v})\|_{L^2}^2 \geq \frac{\mu}{2} \|\nabla^2\mathbf{v}\|_{L^2}^2 \geq C \|\mathbf{v}\|_{H^2}^2.$$

The existence and uniqueness of  $\mathbf{w} \in V$  follow from the Lax-Milgram theorem, and

$$(2.13) \quad \|\nabla^2\mathbf{w}\|_{L^2} \leq \|\mathbf{w}\|_{H^2} \leq C \|\mathbf{F}\|_{H^{-2}}.$$

Noting that  $\operatorname{div} \mathbf{w} \in P$ , using (2.4), we obtain that, there exists  $\mathbf{v}_0 \in V$  such that  $\operatorname{div} \mathbf{v}_0 = \operatorname{div} \mathbf{w}$ , and

$$\|\nabla^2\mathbf{v}_0\|_{L^2} \leq C \|\operatorname{div} \mathbf{w}\|_{H^1} \leq C \|\nabla\operatorname{div} \mathbf{w}\|_{L^2}.$$

A combination of the above two inequalities gives

$$\begin{aligned} \lambda \|\nabla\operatorname{div} \mathbf{w}\|_{L^2}^2 &= \lambda(\nabla\operatorname{div} \mathbf{w}, \nabla\operatorname{div} \mathbf{v}_0) = A(\mathbf{w}, \mathbf{v}_0) - 2\mu(\nabla\epsilon(\mathbf{w}), \nabla\epsilon(\mathbf{v}_0)) \\ &= (\mathbf{F}, \mathbf{v}_0) - 2\mu(\nabla\epsilon(\mathbf{w}), \nabla\epsilon(\mathbf{v}_0)) \\ &\leq \|\mathbf{F}\|_{H^{-2}} \|\mathbf{v}_0\|_{H^2} + 2\mu \|\nabla^2\mathbf{w}\|_{L^2} \|\nabla^2\mathbf{v}_0\|_{L^2} \\ &\leq C (\|\mathbf{F}\|_{H^{-2}} + 2\mu \|\nabla^2\mathbf{w}\|_{L^2}) \|\nabla^2\mathbf{v}_0\|_{L^2} \\ &\leq C \|\mathbf{F}\|_{H^{-2}} \|\nabla\operatorname{div} \mathbf{w}\|_{L^2}. \end{aligned}$$

This implies  $\lambda \|\nabla\operatorname{div} \mathbf{w}\|_{L^2} \leq C \|\mathbf{F}\|_{H^{-2}}$ , which together with (2.13) and Poincaré's inequality gives (2.12).  $\square$

Now we turn to prove the regularity estimate of problem (2.11). We consider an auxiliary elliptic system: For any  $\tilde{\mathbf{F}} \in [L^2(\Omega)]^2$  and  $\tilde{G} \in H^1(\Omega)$ , find  $\mathbf{z} \in V$  and  $q \in P$  such that the following boundary value problem is valid in the sense of distribution:

$$(2.14) \quad \begin{cases} \mu\Delta^2\mathbf{z} + \nabla\Delta q = \tilde{\mathbf{F}} & \text{in } \Omega, \\ \Delta\operatorname{div} \mathbf{z} = \tilde{G} & \text{in } \Omega, \\ \mathbf{z} = \partial_n\mathbf{z} = \mathbf{0} & \text{on } \partial\Omega, \\ q = 0 & \text{on } \partial\Omega. \end{cases}$$

**Lemma 2.5.** *Let  $\mathbf{z} \in [H^4(\Omega)]^2$  and  $q \in H^3(\Omega)$  be the solution of (2.14). Assume that  $m$  is a nonnegative integer, then there exists  $C$  depending only on  $\Omega$  and  $\mu$  such that*

$$(2.15) \quad \|\mathbf{z}\|_{H^{m+4}} + \|q\|_{H^{m+3}} \leq C \left( \|\tilde{\mathbf{F}}\|_{H^m} + \|\tilde{G}\|_{H^{m+1}} + \|\mathbf{z}\|_{L^2} + \|q\|_{L^2} \right).$$

*Proof.* We write (2.14)<sub>1</sub> and (2.14)<sub>2</sub> as

$$\begin{pmatrix} \mu\Delta^2 & 0 & \partial_x\Delta \\ 0 & \mu\Delta^2 & \partial_y\Delta \\ \Delta\partial_x & \Delta\partial_y & 0 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ q \end{pmatrix} = \begin{pmatrix} \tilde{F}_1 \\ \tilde{F}_2 \\ \tilde{G} \end{pmatrix}.$$

The symbol of the above system is

$$\mathcal{L}(\xi) = \begin{pmatrix} \mu|\xi|^4 & 0 & \xi_1|\xi|^2 \\ 0 & \mu|\xi|^4 & \xi_2|\xi|^2 \\ \xi_1|\xi|^2 & \xi_2|\xi|^2 & 0 \end{pmatrix}.$$

A direct calculation gives

$$|\det \mathcal{L}(\xi)| = \mu|\xi|^{10} > 0 \quad \text{if } \xi \neq 0.$$

This means that the boundary value problem (2.14) is elliptic in the sense of Agmon-Douglis-Nirenberg [ADN64]. Moreover, the boundary condition is pure Dirichlet, and it is straightforward to verify that the boundary condition satisfies the complementing condition [ADN64]. The regularity estimate (2.15) follows from [ADN64].  $\square$

A direct consequence of the above lemma is the following regularity estimate for problem (2.11).

**Lemma 2.6.** *Let  $\mathbf{w} \in V$  be the solution of (2.11), there exists  $C$  independent of  $\lambda$  such that*

$$(2.16) \quad \|\mathbf{w}\|_{H^3} + \lambda \|\operatorname{div} \mathbf{w}\|_{H^2} \leq C \|\mathbf{F}\|_{H^{-1}}.$$

*Proof.* Using the standard elliptic regularity estimate, there exists a unique solution  $\mathbf{w} \in [H^4(\Omega)]^2$  when  $\mathbf{F} \in [L^2(\Omega)]^2$ . We introduce  $\mathbf{z} = \mathbf{w}$  and  $q = (\lambda + \mu) \operatorname{div} \mathbf{w}$ , hence  $\mathbf{z} \in [H^4(\Omega)]^2$  and  $q \in H^3(\Omega)$  satisfy (2.14) with  $\tilde{\mathbf{F}} = \mathbf{F}$  and  $\tilde{G} = \Delta \operatorname{div} \mathbf{w}$ .

By (2.15) with  $m = 0$ , we obtain

$$\|\mathbf{w}\|_{H^4} + (\lambda + \mu) \|\operatorname{div} \mathbf{w}\|_{H^3} \leq C (\|\mathbf{F}\|_{L^2} + \|\operatorname{div} \mathbf{w}\|_{H^3} + \|\mathbf{w}\|_{L^2} + (\lambda + \mu) \|\operatorname{div} \mathbf{w}\|_{L^2}).$$

Using the a-priori estimate (2.12), we obtain

$$\|\mathbf{w}\|_{H^4} + (\lambda + \mu) \|\operatorname{div} \mathbf{w}\|_{H^3} \leq C_0 (\|\mathbf{F}\|_{L^2} + \|\operatorname{div} \mathbf{w}\|_{H^3}).$$

Now for  $\lambda + \mu > 2C_0$ , it follows from

$$\|\mathbf{w}\|_{H^4} + (\lambda + \mu) \|\operatorname{div} \mathbf{w}\|_{H^3} \leq C_0 \|\mathbf{F}\|_{L^2} + \frac{\lambda + \mu}{2} \|\operatorname{div} \mathbf{w}\|_{H^3}$$

that

$$\|\mathbf{w}\|_{H^4} + \lambda \|\operatorname{div} \mathbf{w}\|_{H^3} \leq 2C_0 \|\mathbf{F}\|_{L^2}.$$

Interpolating the above inequality with (2.12), we obtain (2.16).

If  $\lambda + \mu \leq 2C_0$ , then (2.16) follows from the standard regularity estimates [ADN64] for problem (2.11).  $\square$

We turn to prove the regularity of problem (1.1) when  $\mathbf{f} \in [L^2(\Omega)]^2$ . Let  $\mathbf{u}$  and  $\mathbf{u}_0$  be the solutions of (1.1) and (2.8), respectively. For any  $\mathbf{v} \in [H_0^1(\Omega) \cap H^2(\Omega)]^2$ , integration by parts gives

$$(2.17) \quad (\mathbb{C}\boldsymbol{\epsilon}(\mathbf{u}), \boldsymbol{\epsilon}(\mathbf{v})) = -(\mathcal{L}\mathbf{u}, \mathbf{v}),$$

and

$$(2.18) \quad \iota^2(\mathbb{D}\nabla\boldsymbol{\epsilon}(\mathbf{u}), \nabla\boldsymbol{\epsilon}(\mathbf{v})) = \iota^2(\Delta\mathcal{L}\mathbf{u}, \mathbf{v}) + \iota^2 \int_{\partial\Omega} (\partial_{\mathbf{n}}\boldsymbol{\sigma}\mathbf{n}) \cdot \partial_{\mathbf{n}}\mathbf{v} d\sigma(\mathbf{x}),$$

where  $\boldsymbol{\sigma} := 2\mu\boldsymbol{\epsilon}(\mathbf{u}) + \lambda \operatorname{div} \mathbf{u} \mathbf{I}$  and  $(\partial_{\mathbf{n}}\boldsymbol{\sigma})_{ij} := \partial_{\mathbf{n}}\sigma_{ij}$ . The boundary term in (2.18) is derived by the fact  $\partial_j v_i = n_j \partial_{\mathbf{n}} v_i + t_j \partial_{\mathbf{t}} v_i = n_j \partial_{\mathbf{n}} v_i$  and

$$(2.19) \quad \begin{aligned} 2\mu \int_{\partial\Omega} \partial_{\mathbf{n}}\boldsymbol{\epsilon}(\mathbf{u}) : \boldsymbol{\epsilon}(\mathbf{v}) d\sigma(\mathbf{x}) + \lambda \int_{\partial\Omega} \partial_{\mathbf{n}} \operatorname{div} \mathbf{u} \operatorname{div} \mathbf{v} d\sigma(\mathbf{x}) &= \int_{\partial\Omega} \partial_{\mathbf{n}}\boldsymbol{\sigma} : \nabla\mathbf{v} d\sigma(\mathbf{x}) \\ &= \int_{\partial\Omega} \partial_{\mathbf{n}}\sigma_{ij} \partial_j v_i d\sigma(\mathbf{x}) = \int_{\partial\Omega} \partial_{\mathbf{n}}\sigma_{ij} n_j \partial_{\mathbf{n}} v_i d\sigma(\mathbf{x}) = \int_{\partial\Omega} (\partial_{\mathbf{n}}\boldsymbol{\sigma}\mathbf{n}) \cdot \partial_{\mathbf{n}}\mathbf{v} d\sigma(\mathbf{x}). \end{aligned}$$

A combination of (2.17) and (2.18) leads to

$$(\mathbb{C}\boldsymbol{\epsilon}(\mathbf{u}), \boldsymbol{\epsilon}(\mathbf{v})) + \iota^2(\mathbb{D}\nabla\boldsymbol{\epsilon}(\mathbf{u}), \nabla\boldsymbol{\epsilon}(\mathbf{v})) = \iota^2 \int_{\partial\Omega} (\partial_{\mathbf{n}}\boldsymbol{\sigma}\mathbf{n}) \cdot \partial_{\mathbf{n}}\mathbf{v} d\sigma(\mathbf{x}) + (\mathbf{f}, \mathbf{v}),$$

which together with (2.9) yields

$$(2.20) \quad (\mathbb{C}\boldsymbol{\epsilon}(\mathbf{u} - \mathbf{u}_0), \boldsymbol{\epsilon}(\mathbf{v})) + \iota^2(\mathbb{D}\nabla\boldsymbol{\epsilon}(\mathbf{u}), \nabla\boldsymbol{\epsilon}(\mathbf{v})) = \iota^2 \int_{\partial\Omega} (\partial_{\mathbf{n}}\boldsymbol{\sigma}\mathbf{n}) \cdot \partial_{\mathbf{n}}\mathbf{v} d\sigma(\mathbf{x}).$$

This identity is the starting point of the proof.

We shall frequently use the following multiplicative trace inequality. There exists  $C$  that depends only on  $\Omega$  such that

$$(2.21) \quad \|\psi\|_{L^2(\partial\Omega)} \leq C \|\psi\|_{L^2(\Omega)}^{1/2} \|\psi\|_{H^1(\Omega)}^{1/2}, \quad \psi \in H^1(\Omega).$$

Using Lemma 2.6, we condense the regularity of problem (1.1) to estimating  $\mathbf{u} - \mathbf{u}_0$  and  $p - p_0$  in various norms, with  $p = \lambda \operatorname{div} \mathbf{u}$  and  $p_0 = \lambda \operatorname{div} \mathbf{u}_0$ .

**Lemma 2.7.** *There exists  $C$  independent of  $\iota$  and  $\lambda$  such that*

$$(2.22) \quad \|\mathbf{u}\|_{H^3} + \|p\|_{H^2} \leq C\iota^{-2} (\|\epsilon(\mathbf{u} - \mathbf{u}_0)\|_{L^2} + \|p - p_0\|_{L^2}).$$

*Proof.* We rewrite (1.1) as

$$\begin{cases} \Delta \mathcal{L}\mathbf{u} = \iota^{-2} \mathcal{L}(\mathbf{u} - \mathbf{u}_0) & \text{in } \Omega, \\ \mathbf{u} = \partial_n \mathbf{u} = \mathbf{0} & \text{on } \partial\Omega. \end{cases}$$

Applying the regularity estimate (2.16) to the elliptic system (2.11), and using (2.17), we obtain

$$\|\mathbf{u}\|_{H^3} + \|p\|_{H^2} \leq C\iota^{-2} \|\mathcal{L}(\mathbf{u} - \mathbf{u}_0)\|_{H^{-1}} \leq C\iota^{-2} (\|\epsilon(\mathbf{u} - \mathbf{u}_0)\|_{L^2} + \|p - p_0\|_{L^2}).$$

□

The next lemma is crucial to prove Theorem 2.9, which transforms the estimate of  $\|p - p_0\|_{\iota}$  in terms of  $\|\nabla(\mathbf{u} - \mathbf{u}_0)\|_{\iota}$  besides a term concerning  $\mathbf{f}$ .

**Lemma 2.8.** *There exists  $C$  independent of  $\iota$  and  $\lambda$  such that*

$$(2.23) \quad \|p - p_0\|_{\iota} \leq C (\|\nabla(\mathbf{u} - \mathbf{u}_0)\|_{\iota} + \iota^{1/2} \|\mathbf{f}\|_{L^2}).$$

*Proof.* By [DM13, Theorem 3.1] and Poincaré's inequality, there exists  $\mathbf{v}_0 \in [H^2(\Omega) \cap H_0^1(\Omega)]^2$  such that  $\operatorname{div} \mathbf{v}_0 = \operatorname{div}(\mathbf{u} - \mathbf{u}_0)$ , and

$$(2.24) \quad \|\mathbf{v}_0\|_{H^1} \leq C \|\operatorname{div}(\mathbf{u} - \mathbf{u}_0)\|_{L^2}, \quad \|\mathbf{v}_0\|_{H^2} \leq C \|\nabla \operatorname{div}(\mathbf{u} - \mathbf{u}_0)\|_{L^2}.$$

Substituting  $\mathbf{v} = \mathbf{v}_0$  into (2.20), multiplying the resulting identity by  $\lambda$ , we obtain

$$(2.25) \quad \begin{aligned} \|p - p_0\|_{\iota}^2 &= -\iota^2 (\nabla p_0, \nabla(p - p_0)) - 2\lambda\mu \left( (\epsilon(\mathbf{u} - \mathbf{u}_0), \epsilon(\mathbf{v}_0)) + \iota^2 (\nabla\epsilon(\mathbf{u}), \nabla\epsilon(\mathbf{v}_0)) \right) \\ &\quad + \lambda\iota^2 \int_{\partial\Omega} (\partial_n \boldsymbol{\sigma} \mathbf{n}) \cdot \partial_n \mathbf{v}_0 d\sigma(\mathbf{x}). \end{aligned}$$

By the regularity estimate (2.10), the first term may be bounded as

$$\iota^2 |(\nabla p_0, \nabla(p - p_0))| \leq \frac{\iota^2}{8} \|\nabla(p - p_0)\|_{L^2}^2 + 2\iota^2 \|\nabla p_0\|_{L^2}^2 \leq \frac{1}{8} \|p - p_0\|_{\iota}^2 + C\iota^2 \|\mathbf{f}\|_{L^2}^2.$$

Using the triangle inequality and (2.10) again, we obtain

$$(2.26) \quad \iota \|\nabla\epsilon(\mathbf{u})\|_{L^2} \leq \iota \|\nabla\epsilon(\mathbf{u} - \mathbf{u}_0)\|_{L^2} + \iota \|\nabla\epsilon(\mathbf{u}_0)\|_{L^2} \leq \|\nabla(\mathbf{u} - \mathbf{u}_0)\|_{\iota} + C\iota \|\mathbf{f}\|_{L^2}.$$

Using (2.24) and the above inequality, we bound the second term as

$$\begin{aligned} 2\lambda\mu |(\epsilon(\mathbf{u} - \mathbf{u}_0), \epsilon(\mathbf{v}_0)) + \iota^2 (\nabla\epsilon(\mathbf{u}), \nabla\epsilon(\mathbf{v}_0))| &\leq C (\|\epsilon(\mathbf{u} - \mathbf{u}_0)\|_{L^2} + \iota \|\nabla\epsilon(\mathbf{u})\|_{L^2}) \|p - p_0\|_{\iota} \\ &\leq \frac{1}{8} \|p - p_0\|_{\iota}^2 + C (\|\nabla(\mathbf{u} - \mathbf{u}_0)\|_{\iota}^2 + \iota^2 \|\mathbf{f}\|_{L^2}^2). \end{aligned}$$

Using (2.19) and the definition of  $\mathbf{v}_0$ , we rewrite the boundary term as

$$\begin{aligned} \lambda^2 \int_{\partial\Omega} (\partial_{\mathbf{n}} \boldsymbol{\sigma} \mathbf{n}) \cdot \partial_{\mathbf{n}} \mathbf{v}_0 d\sigma(\mathbf{x}) &= 2\lambda\mu^2 \int_{\partial\Omega} \partial_{\mathbf{n}} \boldsymbol{\epsilon}(\mathbf{u}) : \boldsymbol{\epsilon}(\mathbf{v}_0) d\sigma(\mathbf{x}) + \lambda^2 \iota^2 \int_{\partial\Omega} \partial_{\mathbf{n}} \operatorname{div} \mathbf{u} \operatorname{div} \mathbf{v}_0 d\sigma(\mathbf{x}) \\ &= 2\lambda\mu^2 \int_{\partial\Omega} \partial_{\mathbf{n}} \boldsymbol{\epsilon}(\mathbf{u}) : \boldsymbol{\epsilon}(\mathbf{v}_0) d\sigma(\mathbf{x}) + \iota^2 \int_{\partial\Omega} \partial_{\mathbf{n}} p (p - p_0) d\sigma(\mathbf{x}) \\ &= 2\lambda\mu^2 \int_{\partial\Omega} \partial_{\mathbf{n}} \boldsymbol{\epsilon}(\mathbf{u}) : \boldsymbol{\epsilon}(\mathbf{v}_0) d\sigma(\mathbf{x}) - \iota^2 \int_{\partial\Omega} \partial_{\mathbf{n}} p p_0 d\sigma(\mathbf{x}). \end{aligned}$$

Recalling the trace inequality (2.21), we estimate the first boundary term as

$$\begin{aligned} 2\lambda\mu^2 \left| \int_{\partial\Omega} \partial_{\mathbf{n}} \boldsymbol{\epsilon}(\mathbf{u}) : \boldsymbol{\epsilon}(\mathbf{v}_0) d\sigma(\mathbf{x}) \right| &\leq 2\lambda\mu^2 \|\partial_{\mathbf{n}} \boldsymbol{\epsilon}(\mathbf{u})\|_{L^2(\partial\Omega)} \|\boldsymbol{\epsilon}(\mathbf{v}_0)\|_{L^2(\partial\Omega)} \\ &\leq C\lambda\iota^2 \|\nabla \boldsymbol{\epsilon}(\mathbf{u})\|_{L^2}^{1/2} \|\nabla \boldsymbol{\epsilon}(\mathbf{u})\|_{H^1}^{1/2} \|\boldsymbol{\epsilon}(\mathbf{v}_0)\|_{L^2}^{1/2} \|\boldsymbol{\epsilon}(\mathbf{v}_0)\|_{H^1}^{1/2}. \end{aligned}$$

Using (2.24), there exists  $C$  independent of  $\lambda$  and  $\iota$  such that

$$\lambda^2 \|\boldsymbol{\epsilon}(\mathbf{v}_0)\|_{L^2} \|\boldsymbol{\epsilon}(\mathbf{v}_0)\|_{H^1} \leq C \|p - p_0\|_{L^2} \|\nabla(p - p_0)\|_{L^2} \leq C\iota^{-1} \|p - p_0\|_{L^2}^2.$$

Using (2.26) to estimate  $\|\nabla \boldsymbol{\epsilon}(\mathbf{u})\|_{L^2}$  and using (2.22) to bound  $\|\nabla \boldsymbol{\epsilon}(\mathbf{u})\|_{H^1}$ , we obtain

$$\begin{aligned} \|\nabla \boldsymbol{\epsilon}(\mathbf{u})\|_{L^2} \|\nabla \boldsymbol{\epsilon}(\mathbf{u})\|_{H^1} &\leq C\iota^{-3} (\|\nabla(\mathbf{u} - \mathbf{u}_0)\|_{L^2} + \iota \|\mathbf{f}\|_{L^2}) (\|\boldsymbol{\epsilon}(\mathbf{u} - \mathbf{u}_0)\|_{L^2} + \|p - p_0\|_{L^2}) \\ &\leq C\iota^{-3} \left( \|\nabla(\mathbf{u} - \mathbf{u}_0)\|_{L^2}^2 + \iota^2 \|\mathbf{f}\|_{L^2}^2 + (\|\nabla(\mathbf{u} - \mathbf{u}_0)\|_{L^2} + \iota \|\mathbf{f}\|_{L^2}) \|p - p_0\|_{L^2} \right). \end{aligned}$$

A combination of the above three inequalities gives

$$\begin{aligned} 2\lambda\mu^2 \left| \int_{\partial\Omega} \partial_{\mathbf{n}} \boldsymbol{\epsilon}(\mathbf{u}) : \boldsymbol{\epsilon}(\mathbf{v}_0) d\sigma(\mathbf{x}) \right| &\leq C \left( (\|\nabla(\mathbf{u} - \mathbf{u}_0)\|_{L^2} + \iota \|\mathbf{f}\|_{L^2}) \|p - p_0\|_{L^2} \right. \\ &\quad \left. + (\|\nabla(\mathbf{u} - \mathbf{u}_0)\|_{L^2} + \iota \|\mathbf{f}\|_{L^2})^{1/2} \|p - p_0\|_{L^2}^{3/2} \right) \\ &\leq \frac{1}{8} \|p - p_0\|_{L^2}^2 + C (\|\nabla(\mathbf{u} - \mathbf{u}_0)\|_{L^2}^2 + \iota^2 \|\mathbf{f}\|_{L^2}^2). \end{aligned}$$

Using the trace inequality (2.21) and the regularity estimate (2.10) again, we bound

$$\iota^2 \left| \int_{\partial\Omega} \partial_{\mathbf{n}} p p_0 d\sigma(\mathbf{x}) \right| \leq C\iota^2 \|\nabla p\|_{L^2}^{1/2} \|\nabla p\|_{H^1}^{1/2} \|p_0\|_{H^1} \leq C\iota^2 \|\nabla p\|_{L^2}^{1/2} \|\nabla p\|_{H^1}^{1/2} \|\mathbf{f}\|_{L^2}.$$

Using the triangle inequality and (2.10) again, we obtain

$$\iota \|\nabla p\|_{L^2} \leq \|p - p_0\|_{L^2} + \iota \|\nabla p_0\|_{L^2} \leq \|p - p_0\|_{L^2} + C\iota \|\mathbf{f}\|_{L^2},$$

which together with (2.22) implies

$$\|\nabla p\|_{L^2} \|\nabla p\|_{H^1} \leq C\iota^{-3} \left( \|p - p_0\|_{L^2}^2 + (\|\nabla(\mathbf{u} - \mathbf{u}_0)\|_{L^2} + \iota \|\mathbf{f}\|_{L^2}) \|p - p_0\|_{L^2} + \iota \|\nabla(\mathbf{u} - \mathbf{u}_0)\|_{L^2} \|\mathbf{f}\|_{L^2} \right).$$

Combining the above three inequalities, we bound the second boundary term as

$$\begin{aligned} \iota^2 \left| \int_{\partial\Omega} \partial_{\mathbf{n}} p p_0 d\sigma(\mathbf{x}) \right| &\leq C \left( \iota^{1/2} \|p - p_0\|_{L^2} \|\mathbf{f}\|_{L^2} + (\|\nabla(\mathbf{u} - \mathbf{u}_0)\|_{L^2} + \iota^{1/2} \|\mathbf{f}\|_{L^2})^{3/2} \|p - p_0\|_{L^2}^{1/2} \right. \\ &\quad \left. + \iota \|\nabla(\mathbf{u} - \mathbf{u}_0)\|_{L^2}^{1/2} \|\mathbf{f}\|_{L^2}^{3/2} \right) \\ &\leq \frac{1}{8} \|p - p_0\|_{L^2}^2 + C (\|\nabla(\mathbf{u} - \mathbf{u}_0)\|_{L^2}^2 + \iota \|\mathbf{f}\|_{L^2}^2). \end{aligned}$$

Substituting the above inequalities into (2.25), we obtain

$$\|p - p_0\|_{L^2}^2 \leq \frac{1}{2} \|p - p_0\|_{L^2}^2 + C (\|\nabla(\mathbf{u} - \mathbf{u}_0)\|_{L^2}^2 + \iota \|\mathbf{f}\|_{L^2}^2).$$

This immediately gives (2.23).  $\square$

We are ready to prove the regularity of problem (1.1).

**Theorem 2.9.** *There exists  $C$  independent of  $\iota$  and  $\lambda$  such that*

$$(2.27) \quad \|\mathbf{u} - \mathbf{u}_0\|_{H^1} + \|p - p_0\|_{L^2} \leq C\iota^{1/2} \|\mathbf{f}\|_{L^2},$$

and

$$(2.28) \quad \|\mathbf{u}\|_{H^{2+k}} + \|p\|_{H^{1+k}} \leq C\iota^{-1/2-k} \|\mathbf{f}\|_{L^2}, \quad k = 0, 1.$$

The estimate (2.27) improves the known results [LMW21, Lemma 1] in two aspects. It clarifies the fact that the estimate is  $\lambda$ -independent and it gives the convergence rate for the pressure. The rate  $\iota^{1/2}$  is optimal even for the scalar counterpart; cf., [NTW01, Lemma 5.1].

*Proof.* Substituting  $\mathbf{v} = \mathbf{u} - \mathbf{u}_0$  into (2.20), we get

$$a(\mathbf{u} - \mathbf{u}_0, \mathbf{u} - \mathbf{u}_0) = -\iota^2 (\mathbb{D}\nabla\epsilon(\mathbf{u}), \nabla\epsilon(\mathbf{u} - \mathbf{u}_0)) - \iota^2 \int_{\partial\Omega} (\partial_{\mathbf{n}}\boldsymbol{\sigma}\mathbf{n}) \cdot \partial_{\mathbf{n}}\mathbf{u}_0 d\sigma(\mathbf{x}).$$

Using the regularity estimate (2.10), we bound the first term as

$$\iota^2 |(\mathbb{D}\nabla\epsilon(\mathbf{u}), \nabla\epsilon(\mathbf{u} - \mathbf{u}_0))| \leq \frac{\iota^2}{4} (\mathbb{D}\nabla\epsilon(\mathbf{u} - \mathbf{u}_0), \nabla\epsilon(\mathbf{u} - \mathbf{u}_0)) + \iota^2 (\mathbb{D}\nabla\epsilon(\mathbf{u}_0), \nabla\epsilon(\mathbf{u}_0)) \leq \frac{1}{4} a(\mathbf{u} - \mathbf{u}_0, \mathbf{u} - \mathbf{u}_0) + C\iota^2 \|\mathbf{f}\|_{L^2}^2.$$

To bound the second term, we let  $\mathbf{v} = \mathbf{u}_0$  in (2.19) and obtain

$$\iota^2 \int_{\partial\Omega} (\partial_{\mathbf{n}}\boldsymbol{\sigma}\mathbf{n}) \cdot \partial_{\mathbf{n}}\mathbf{u}_0 d\sigma(\mathbf{x}) = 2\mu^2 \int_{\partial\Omega} \partial_{\mathbf{n}}\epsilon(\mathbf{u}) : \epsilon(\mathbf{u}_0) d\sigma(\mathbf{x}) + \iota^2 \int_{\partial\Omega} \partial_{\mathbf{n}} \operatorname{div} \mathbf{u} p_0 d\sigma(\mathbf{x}).$$

Invoking the trace inequality, using the fact  $\|\nabla \operatorname{div} \mathbf{u}\|_{L^2} \leq \|\nabla\epsilon(\mathbf{u})\|_{L^2}$  and (2.10), we obtain

$$\begin{aligned} \iota^2 \left| \int_{\partial\Omega} (\partial_{\mathbf{n}}\boldsymbol{\sigma}\mathbf{n}) \cdot \partial_{\mathbf{n}}\mathbf{u}_0 d\sigma(\mathbf{x}) \right| &\leq C\iota^2 \left( \|\nabla\epsilon(\mathbf{u})\|_{L^2}^{1/2} \|\nabla\epsilon(\mathbf{u})\|_{H^1}^{1/2} \|\epsilon(\mathbf{u}_0)\|_{H^1} + \|\nabla \operatorname{div} \mathbf{u}\|_{L^2}^{1/2} \|\nabla \operatorname{div} \mathbf{u}\|_{H^1}^{1/2} \|p_0\|_{H^1} \right) \\ &\leq C\iota^2 \|\nabla\epsilon(\mathbf{u})\|_{L^2}^{1/2} \|\mathbf{u}\|_{H^3}^{1/2} (\|\mathbf{u}_0\|_{H^2} + \|p_0\|_{H^1}) \\ &\leq C\iota^2 \|\nabla\epsilon(\mathbf{u})\|_{L^2}^{1/2} \|\mathbf{u}\|_{H^3}^{1/2} \|\mathbf{f}\|_{L^2}. \end{aligned}$$

Substituting (2.23) into (2.22), we obtain,

$$\|\mathbf{u}\|_{H^3} + \|p\|_{H^2} \leq C\iota^{-2} (\|\nabla(\mathbf{u} - \mathbf{u}_0)\|_{\iota} + \iota^{1/2} \|\mathbf{f}\|_{L^2}).$$

Invoking (2.26) again, we get

$$\iota^2 \|\nabla\epsilon(\mathbf{u})\|_{L^2}^{1/2} \|\mathbf{u}\|_{H^3}^{1/2} \leq C\iota^{1/2} (\|\nabla(\mathbf{u} - \mathbf{u}_0)\|_{\iota} + \iota^{1/2} \|\mathbf{f}\|_{L^2}).$$

A combination of the above three inequalities gives

$$\iota^2 \left| \int_{\partial\Omega} (\partial_{\mathbf{n}}\boldsymbol{\sigma}\mathbf{n}) \cdot \partial_{\mathbf{n}}\mathbf{u}_0 d\sigma(\mathbf{x}) \right| \leq \frac{1}{4} \|\nabla(\mathbf{u} - \mathbf{u}_0)\|_{\iota}^2 + C\iota \|\mathbf{f}\|_{L^2}^2.$$

Combining the above inequalities, we obtain

$$\|\nabla(\mathbf{u} - \mathbf{u}_0)\|_{\iota} \leq C\iota^{1/2} \|\mathbf{f}\|_{L^2},$$

which together with (2.23) leads to

$$\|p - p_0\|_{\iota} \leq C\iota^{1/2} \|\mathbf{f}\|_{L^2}.$$

The above two estimates immediately give (2.27) and

$$\|\nabla^2(\mathbf{u} - \mathbf{u}_0)\|_{L^2} + \|\nabla(p - p_0)\|_{L^2} \leq C\iota^{-1/2} \|\mathbf{f}\|_{L^2},$$

which together with (2.10) gives (2.28) with  $k = 0$ .

Substituting (2.27) into (2.22), we obtain the higher regularity estimate (2.28) with  $k = 1$ .  $\square$

A direct consequence of the above theorem is

**Corollary 2.10.** *There exists  $C$  independent of  $\iota$  and  $\lambda$  such that*

$$(2.29) \quad \|\mathbf{u}\|_{H^{3/2}} + \|p\|_{H^{1/2}} \leq C \|\mathbf{f}\|_{L^2},$$

and

$$(2.30) \quad \|\mathbf{u}\|_{H^{5/2}} + \|p\|_{H^{3/2}} \leq C\iota^{-1} \|\mathbf{f}\|_{L^2}.$$

*Proof.* Using triangle inequality, (2.10) and (2.28), we obtain

$$\|\mathbf{u} - \mathbf{u}_0\|_{H^2} + \|p - p_0\|_{H^1} \leq \|\mathbf{u}\|_{H^2} + \|p\|_{H^1} + \|\mathbf{u}_0\|_{H^2} + \|p_0\|_{H^1} \leq C\iota^{-1/2} \|\mathbf{f}\|_{L^2}.$$

Interpolating the above inequality and (2.27), we obtain

$$\|\mathbf{u} - \mathbf{u}_0\|_{H^{3/2}} + \|p - p_0\|_{H^{1/2}} \leq C \|\mathbf{f}\|_{L^2}.$$

Using (2.10), we get

$$\|\mathbf{u}_0\|_{H^{3/2}} + \|p_0\|_{H^{1/2}} \leq C \|\mathbf{u}_0\|_{H^2} + \|p_0\|_{H^1} \leq C \|\mathbf{f}\|_{L^2}.$$

A combination of the above two inequalities and the triangle inequality leads to (2.29).

Interpolating (2.28) with  $k = 0$  and (2.28) with  $k = 1$ , we obtain (2.30).  $\square$

### 3. A FAMILY OF NONCONFORMING FINITE ELEMENTS

We introduce a family of finite elements to approximate the mixed variational problem (2.2). Let  $\mathcal{T}_h$  be a triangulation of  $\Omega$  with maximum mesh size  $h$ . We assume all elements in  $\mathcal{T}_h$  are shape-regular in the sense of Ciarlet and Raviart [Cia78]. We also assume that  $\mathcal{T}_h$  satisfies the inverse assumption: there exists  $\sigma_0$  such that  $h/h_K \leq \sigma_0$  for all  $K \in \mathcal{T}_h$ . The space of piecewise vector fields is defined by

$$[H^m(\Omega, \mathcal{T}_h)]^2 := \{ \mathbf{v} \in [L^2(\Omega)]^2 \mid \mathbf{v}|_K \in [H^m(T)]^2 \text{ for all } K \in \mathcal{T}_h \},$$

which is equipped with the broken norm

$$\|\mathbf{v}\|_{H^m(\Omega, \mathcal{T}_h)} := \|\mathbf{v}\|_{L^2} + \sum_{k=1}^m \|\nabla_h^k \mathbf{v}\|_{L^2},$$

where  $\|\nabla_h^k \mathbf{v}\|_{L^2}^2 = \sum_{K \in \mathcal{T}_h} \|\nabla^k \mathbf{v}\|_{L^2(K)}^2$ . For an interior edge  $e$  shared by the triangles  $K^+$  and  $K^-$ , we define the jump of  $\mathbf{v}$  across  $e$  as

$$[[\mathbf{v}]] := \mathbf{v}^+ \mathbf{n}^+ + \mathbf{v}^- \mathbf{n}^- \quad \text{with} \quad \mathbf{v}^\pm = \mathbf{v}|_{K^\pm},$$

where  $\mathbf{n}^\pm$  is the unit normal vector of  $e$  towards the outside of  $K^\pm$ . For  $e \cap \partial\Omega \neq \emptyset$ , we set  $[[\mathbf{v}]] = \mathbf{v}\mathbf{n}$ .

**3.1. A family of finite elements.** Our construction is motivated by the element proposed in [GLN12]. Define the element with a triple  $(K, P_K, \Sigma_K)$  by specifying  $K$  as a triangle, and

$$(3.1) \quad P_K := \mathbb{P}_r(K) + b_K \sum_{i=1}^3 b_i Q_i^{r-2}(K) + b_K^2 R^{r-2}(K),$$

where  $b_K = \prod_{i=1}^3 \lambda_i$  and  $b_i = b_K/\lambda_i$  with  $\{\lambda_i\}_{i=1}^3$  the barycentric coordinates of  $K$ .

Define

$$(3.2) \quad Q_i^{r-2}(K) := \left\{ v \in \mathbb{P}_{r-2}(K) \mid \int_K b_K b_i v q \, d\mathbf{x} = 0 \text{ for all } q \in \mathbb{P}_{r-3}(K) \right\},$$

and

$$(3.3) \quad R^{r-2}(K) := \left\{ v \in \mathbb{P}_{r-2}(K) \mid \int_K b_K^2 v q \, d\mathbf{x} = 0 \text{ for all } q \in \mathbb{P}_{r-3}(K) \right\}.$$

The degrees of freedom (DoFs) for  $P_K$  are given by

$$(3.4) \quad v \mapsto \begin{cases} v(a) & \text{for all vertices } a, \\ \int_e vq \, d\sigma(\mathbf{x}) & \text{for all edges } e \text{ and } q \in \mathbb{P}_{r-2}(e), \\ \int_e \partial_{\mathbf{n}} vq \, d\sigma(\mathbf{x}) & \text{for all edges } e \text{ and } q \in \mathbb{P}_{r-2}(e), \\ \int_K vq \, d\mathbf{x} & \text{for all } q \in \mathbb{P}_{r-2}(K). \end{cases}$$

We plot the DoFs for  $r = 2, 3$  in Figure 1.

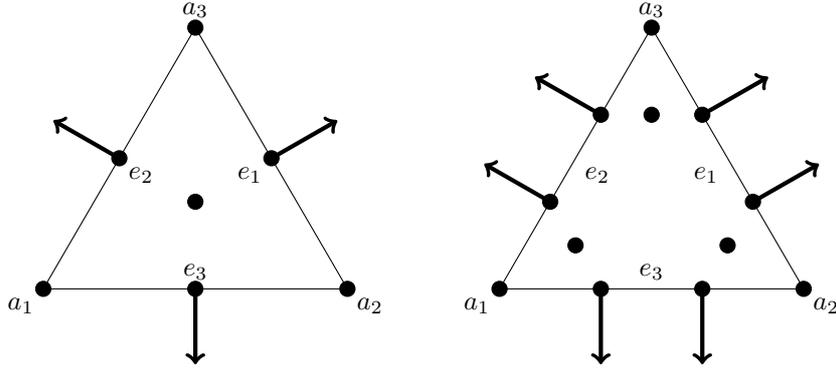


FIGURE 1. Diagram for DoFs. Left: DoFs for  $r = 2$  are point evaluations of the function values at the vertex, the mean of the function along each edge, the mean of the normal derivative along each edge, and the mean of the function over the element; Right: DoFs for  $r = 3$  are point evaluations of the function values at the vertex, the means of the function against  $\mathbb{P}_1$  along each edge, the means of the normal derivative against  $\mathbb{P}_1$  along each edge, and the means of the function against  $\mathbb{P}_1$  over the element

**Lemma 3.1.** *The set  $(K, P_K, \Sigma_K)$  is unisolvent.*

*Proof.* Firstly we show that if the DoFs (3.4) can determine an element in  $P_K$ , then the element is unique. Suppose  $v \in P_K$  vanishes at the DoFs listed in (3.4), it suffices to show  $v \equiv 0$ . Assume that

$$v = p_r + b_K \sum_{i=1}^3 b_i q_i + b_K^2 q_r,$$

where  $p_r \in \mathbb{P}_r(K)$ , and  $q_i \in Q_i^{r-2}(K)$ , and  $q_r \in R^{r-2}(K)$ . DoFs associated with  $\mathbb{P}_r(K)$  are

$$v \mapsto \begin{cases} v(a) & \text{for all vertices } a, \\ \int_e vq \, d\sigma(\mathbf{x}) & \text{for all edges } e \text{ and } q \in \mathbb{P}_{r-2}(e), \\ \int_K vq \, d\mathbf{x} & \text{for all } q \in \mathbb{P}_{r-3}(K). \end{cases}$$

The bubble space vanishes on this subset of DoFs by (3.2) and (3.3). The number of the DoFs is  $3 + 3(r-1) + (r-1)(r-2)/2 = (r+1)(r+2)/2$ , which equals to the cardinality of  $\mathbb{P}_r(K)$ . Hence  $p_r \equiv 0$ .

A direct calculation gives

$$\int_{e_i} \partial_{\mathbf{n}} vq \, d\sigma(\mathbf{x}) = -|\nabla \lambda_i| \int_{e_i} b_i^2 q_i q \, d\sigma(\mathbf{x}) = 0 \quad \text{for all } q \in \mathbb{P}_{r-2}(e_i).$$

Taking  $q = q_i$  in the above identity, we obtain  $q_i = 0$  on  $e_i$ . Therefore, we write  $q_i = \lambda_i p_{r-3}$  for certain  $p_{r-3} \in \mathbb{P}_{r-3}(K)$ . Using (3.2), we get

$$\int_K b_K b_i q_i q \, d\mathbf{x} = \int_K b_K^2 p_{r-3} q \, d\mathbf{x} = 0 \quad \text{for all } q \in \mathbb{P}_{r-3}(K).$$

Taking  $q = p_{r-3}$  in the above identity, we obtain  $p_{r-3} \equiv 0$ . Therefore  $v = b_K^2 q_r$  for certain  $q_r \in R^{r-2}(K)$ . The last set of DoFs equals zero, i.e.,

$$\int_K b_K^2 q_r q \, d\mathbf{x} = 0 \quad \text{for all } q \in \mathbb{P}_{r-2}(K).$$

Taking  $q = q_r$  in the above identity, we obtain  $q_r \equiv 0$ . So does  $v$ .

It remains to show the dimension of  $P_K$  equals the number of DoFs (3.4). Proceeding along the same line as above, the element  $v \equiv 0$  has a unique representation. Therefore (3.1) is a direct sum, and

$$\dim P_K = \dim \mathbb{P}_r(K) + 4(\dim \mathbb{P}_{r-2}(K) - \dim \mathbb{P}_{r-3}(K)) = \frac{1}{2}(r^2 + 11r - 6),$$

which equals to the number of DoFs (3.4) exactly.  $\square$

We define a local interpolation operator  $\pi_K : H^2(K) \mapsto P_K$  as:

$$(3.5) \quad \left\{ \begin{array}{ll} \pi_K v(a) = v(a) & \text{for all vertices } a, \\ \int_e \pi_K v q \, d\sigma(\mathbf{x}) = \int_e v q \, d\sigma(\mathbf{x}) & \text{for all edges } e \text{ and } q \in \mathbb{P}_{r-2}(e), \\ \int_e \partial_{\mathbf{n}} \pi_K v q \, d\sigma(\mathbf{x}) = \int_e \partial_{\mathbf{n}} v q \, d\sigma(\mathbf{x}) & \text{for all edges } e \text{ and } q \in \mathbb{P}_{r-2}(e), \\ \int_K \pi_K v q \, d\mathbf{x} = \int_K v q \, d\mathbf{x} & \text{for all } q \in \mathbb{P}_{r-2}(K). \end{array} \right.$$

**Lemma 3.2.** *There exists  $C$  independent of  $h_K$  such that for  $v \in H^k(K)$  with  $2 \leq k \leq r+1$ , there holds*

$$(3.6) \quad \|\nabla^j(v - \pi_K v)\|_{L^2(K)} \leq C h_K^{k-j} \|\nabla^k v\|_{L^2(K)},$$

where  $0 \leq j \leq k$ .

*Proof.* For any  $v \in \mathbb{P}_r(K) \subset P_K$ , the definition (3.5) shows that  $v - \pi_K v \in P_K$  and all DoFs of  $v - \pi_K v$  vanish, then  $v = \pi_K v$ . The estimate (3.6) immediately follows from the  $\mathbb{P}_r(K)$ -invariance of the local interpolation operator  $\pi_K$  [CR72].  $\square$

*Remark 3.3.* The element has a natural extension to three-dimensions by specifying  $K$  as a tetrahedron, and

$$P_K := \mathbb{P}_r(K) + b_K \sum_{i=1}^4 b_i Q_i^{r-2}(K) + b_K^2 R^{r-2}(K),$$

where  $b_K = \prod_{i=1}^4 \lambda_i$  is the element bubble function with  $\lambda_i$  the barycentric coordinates associated with the vertices  $a_i$  for  $i = 1, \dots, 4$ .  $b_i = b_K / \lambda_i$  is the face bubble function associated with the face  $f_i$ .

Define

$$Q_i^{r-2}(K) := \left\{ v \in \mathbb{P}_{r-2}(K) \mid \int_K b_K b_i v q \, d\mathbf{x} = 0 \text{ for all } q \in \mathbb{P}_{r-3}(K) \right\},$$

and

$$R^{r-2}(K) := \left\{ v \in \mathbb{P}_{r-2}(K) \mid \int_K b_K^2 v q \, d\mathbf{x} = 0 \text{ for all } q \in \mathbb{P}_{r-4}(K) \right\}.$$

The DoFs for  $P_K$  are given by

$$v \mapsto \begin{cases} v(a) & \text{for all vertices,} \\ \int_e vq \, d\sigma(\mathbf{x}) & \text{for all edges } e \text{ and } q \in \mathbb{P}_{r-2}(e), \\ \int_f vq \, d\sigma(\mathbf{x}) & \text{for all faces } f \text{ and } q \in \mathbb{P}_{r-3}(f), \\ \int_f \partial_{\mathbf{n}} vq \, d\sigma(\mathbf{x}) & \text{for all faces } f \text{ and } q \in \mathbb{P}_{r-2}(f), \\ \int_K vq \, d\mathbf{x} & \text{for all } q \in \mathbb{P}_{r-2}(K). \end{cases}$$

Similar to Lemma 3.1, the set  $(K, P_K, \Sigma_K)$  is also unisolvent.

**3.2. Explicit representation for the bubble space.** We clarify the structures of (3.2) and (3.3) associated with the set of DoFs (3.4)<sub>3</sub> and the subset of (3.4)<sub>4</sub> respectively, and derive the explicit formulations of the corresponding shape functions, which seems missing in the literature, while such explicit representations are useful for implementation. We firstly recall the following facts about the Jacobi polynomials [Sze75]. For any  $\alpha, \beta > -1$  and nonnegative integers  $n, m$ , there holds

$$(3.7) \quad \int_{-1}^1 (1-t)^\alpha (1+t)^\beta P_n^{(\alpha, \beta)}(t) P_m^{(\alpha, \beta)}(t) dt = h_n^{(\alpha, \beta)} \delta_{nm},$$

where

$$h_n^{(\alpha, \beta)} = \int_{-1}^1 (1-t)^\alpha (1+t)^\beta [P_n^{(\alpha, \beta)}(t)]^2 dt.$$

By [Sze75, Eq. (4.3.3)], we may write

$$(3.8) \quad h_n^{(\alpha, \beta)} = \frac{2^{\alpha+\beta+1}}{2n + \alpha + \beta + 1} \frac{\Gamma(n + \alpha + 1)\Gamma(n + \beta + 1)}{\Gamma(n + \alpha + \beta + 1)\Gamma(n + 1)},$$

where  $\Gamma$  is the Gamma function.

One of the explicit form for  $P_n^{(\alpha, \beta)}$  is

$$(1-t)^\alpha (1+t)^\beta P_n^{(\alpha, \beta)}(t) = \frac{(-1)^n}{2^n n!} \frac{d^n}{dt^n} ((1-t)^{n+\alpha} (1+t)^{n+\beta}).$$

In particular,

$$P_0^{(\alpha, \beta)}(t) = 1, \quad P_1^{(\alpha, \beta)}(t) = \frac{1}{2}(\alpha + \beta + 2)t + \frac{1}{2}(\alpha - \beta).$$

Next we list certain facts about the Jacobi polynomials on the triangle [DX14, Section 2.4]. For a triangle  $K$  with vertices  $a_1, a_2, a_3$ , any point  $x \in K$  is uniquely expressed as

$$x = \lambda_1 a_1 + \lambda_2 a_2 + \lambda_3 a_3, \quad \lambda_i \geq 0 \text{ and } \lambda_1 + \lambda_2 + \lambda_3 = 1.$$

Then  $(\lambda_1, \lambda_2, \lambda_3)$  is the barycentric coordinates of the point  $x$  with respect to  $K$ . For nonnegative integers  $k, n$  such that  $k \leq n$ , we define

$$(3.9) \quad P_{k,n}^{(\alpha, \beta, \gamma)}(\lambda_1, \lambda_2, \lambda_3) := (\lambda_2 + \lambda_3)^k P_{n-k}^{(2k+\beta+\gamma+1, \alpha)}(\lambda_1 - \lambda_2 - \lambda_3) P_k^{(\gamma, \beta)}((\lambda_2 - \lambda_3)/(\lambda_2 + \lambda_3)).$$

It is straightforward to verify  $P_{k,n}^{(\alpha, \beta, \gamma)} \in \mathbb{P}_n(K)$ . In particular,

$$P_{0,0}^{(\alpha, \beta, \gamma)}(\lambda_1, \lambda_2, \lambda_3) = 1, \quad P_{0,1}^{(\alpha, \beta, \gamma)}(\lambda_1, \lambda_2, \lambda_3) = (\beta + \gamma + 2)\lambda_1 - (\alpha + 1)(\lambda_2 + \lambda_3),$$

and

$$P_{1,1}^{(\alpha, \beta, \gamma)}(\lambda_1, \lambda_2, \lambda_3) = (\gamma + 1)\lambda_2 - (\beta + 1)\lambda_3.$$

For all  $\alpha, \beta, \gamma > -1$ , and nonnegative integers  $j, k, m, n$  such that  $j \leq m$  and  $k \leq n$ , there holds

$$\begin{aligned}
& \int_K \lambda_1^\alpha \lambda_2^\beta \lambda_3^\gamma P_{k,n}^{(\alpha,\beta,\gamma)}(\lambda_1, \lambda_2, \lambda_3) P_{j,m}^{(\alpha,\beta,\gamma)}(\lambda_1, \lambda_2, \lambda_3) d\mathbf{x} \\
(3.10) \quad & = 2 \int_{\widehat{K}} \lambda_1^\alpha \lambda_2^\beta (1 - \lambda_1 - \lambda_2)^\gamma P_{k,n}^{(\alpha,\beta,\gamma)}(\lambda_1, \lambda_2, 1 - \lambda_1 - \lambda_2) P_{j,m}^{(\alpha,\beta,\gamma)}(\lambda_1, \lambda_2, 1 - \lambda_1 - \lambda_2) d\lambda_1 d\lambda_2 \\
& = 2h_{k,n}^{(\alpha,\beta,\gamma)} \delta_{jk} \delta_{mn},
\end{aligned}$$

where  $\widehat{K} := \{(\lambda_1, \lambda_2) \mid \lambda_1 \geq 0, \lambda_2 \geq 0, \lambda_1 + \lambda_2 \leq 1\}$  is the standard reference triangle and

$$\begin{aligned}
(3.11) \quad h_{k,n}^{(\alpha,\beta,\gamma)} & = 2^{-(2k+\alpha+2\beta+2\gamma+3)} h_{n-k}^{(2k+\beta+\gamma+1,\alpha)} h_k^{(\gamma,\beta)} \\
& = \frac{1}{(2n + \alpha + \beta + \gamma + 2)(2k + \beta + \gamma + 1)} \\
& \quad \times \frac{\Gamma(n - k + \alpha + 1)\Gamma(n + k + \beta + \gamma + 2)\Gamma(k + \beta + 1)\Gamma(k + \gamma + 1)}{\Gamma(n - k + 1)\Gamma(n + k + \alpha + \beta + \gamma + 2)\Gamma(k + 1)\Gamma(k + \beta + \gamma + 1)}.
\end{aligned}$$

Using the notation  $(x)_n = \Gamma(x + n)/\Gamma(x)$ , we may find that the expression (3.11) is equivalent to the one in [DX14, Eq. (2.4.3)]. The identity (3.10) illustrates that  $\{P_{k,n}^{(\alpha,\beta,\gamma)}(\lambda_1, \lambda_2, \lambda_3) \mid 0 \leq k \leq n \leq r\}$  are mutually orthogonal bases of  $\mathbb{P}_r(K)$  with respect to the weight  $\lambda_1^\alpha \lambda_2^\beta \lambda_3^\gamma$ .

Next we study the structure of the bubble spaces. For the barycentric coordinate function  $\lambda_i$  such that  $\lambda_i \equiv 0$  on  $e_i$ , let  $\lambda_i^+$  and  $\lambda_i^-$  be the two other barycentric coordinates associated with the edges  $e_i^+$  and  $e_i^-$ , respectively.  $(e_i, e_i^+, e_i^-)$  are chosen in a counterclockwise manner. The space  $Q_i^{r-2}(K)$  can be clarified by the Jacobi polynomials with respect to the weight  $b_K b_i$ , while  $R^{r-2}(K)$  can be clarified by the Jacobi polynomials with respect to the weight  $b_K^2$ , which are formulated in the following lemmas.

**Lemma 3.4.** *The space  $Q_i^{r-2}(K)$  takes the form*

$$Q_i^{r-2}(K) = \text{span} \left\{ P_{k,r-2}^{(1,2,2)}(\lambda_i, \lambda_i^+, \lambda_i^-) \mid 0 \leq k \leq r-2 \right\}.$$

*Proof.* For any  $v \in Q_i^{r-2}(K) \subset \mathbb{P}_{r-2}(K)$ ,  $v$  may be expanded into

$$v = \sum_{0 \leq k \leq n \leq r-2} a_{kn} P_{k,n}^{(1,2,2)}(\lambda_i, \lambda_i^+, \lambda_i^-)$$

with unknown parameters  $a_{kn}$ . Using the above representation, we may write the constraint in the definition (3.2) as

$$\sum_{0 \leq k \leq n \leq r-2} a_{kn} \int_K b_K b_i P_{k,n}^{(1,2,2)}(\lambda_i, \lambda_i^+, \lambda_i^-) q d\mathbf{x} = 0 \quad \text{for all } q \in \mathbb{P}_{r-3}(K).$$

Substituting  $q = P_{j,m}^{(1,2,2)}(\lambda_i, \lambda_i^+, \lambda_i^-)$  for  $0 \leq j \leq m \leq r-3$  into the above equation, and using the orthogonal relation (3.10), we obtain  $a_{jm} = 0$  for  $0 \leq j \leq m \leq r-3$ . This concludes the lemma.  $\square$

Motivated by the above lemma, we change the definition of DoFs for the bubble space  $b_K \sum_{i=1}^3 b_i Q_i^{r-2}(K)$  from  $\int_{e_i} \partial_n v q d\sigma(\mathbf{x})$  for any  $q \in \mathbb{P}_{r-2}(e_i)$  to

$$\int_{e_i} \frac{\partial v}{\partial n} P_k^{(2,2)}(\lambda_i^+ - \lambda_i^-) d\sigma(\mathbf{x}), \quad k = 0, \dots, r-2.$$

**Lemma 3.5.** *The shape functions for the bubble space  $b_K \sum_{i=1}^3 b_i Q_i^{r-2}(K)$  associated with the above definition of DoFs are*

$$a_{k,r-2} b_K b_i P_{k,r-2}^{(1,2,2)}(\lambda_i, \lambda_i^+, \lambda_i^-)$$

with

$$(3.12) \quad a_{k,r-2} = \frac{(-1)^{r-k-1}}{|\nabla \lambda_i|} \frac{(k+3)(k+4)(2k+5)}{(r-k-1)(k+1)(k+2)} \quad k = 0, \dots, r-2.$$

*Proof.* A direct calculation gives

$$\frac{\partial}{\partial \mathbf{n}} \left( b_K b_i P_{k,r-2}^{(1,2,2)}(\lambda_i, \lambda_i^+, \lambda_i^-) \right) |_{e_i} = -|\nabla \lambda_i| b_i^2 P_{r-2-k}^{(2k+5,1)}(-1) P_k^{(2,2)}(\lambda_i^+ - \lambda_i^-),$$

For  $j = 1, \dots, r-2$ . Using the relation (3.7), and noting that  $b_i = \lambda_i^+ \lambda_i^- = \lambda_i^+(1 - \lambda_i^+)$  on  $e_i$ , we obtain

$$\begin{aligned} & \int_{e_i} \frac{\partial}{\partial \mathbf{n}} \left( b_K b_i P_{k,r-2}^{(1,2,2)}(\lambda_i, \lambda_i^+, \lambda_i^-) \right) P_j^{(2,2)}(\lambda_i^+ - \lambda_i^-) d\sigma(\mathbf{x}) \\ &= -|\nabla \lambda_i| P_{r-2-k}^{(2k+5,1)}(-1) \int_0^1 (\lambda_i^+(1 - \lambda_i^+))^2 P_k^{(2,2)}(2\lambda_i^+ - 1) P_j^{(2,2)}(2\lambda_i^+ - 1) d\lambda_i^+ \\ &= -\frac{|\nabla \lambda_i|}{32} P_{r-2-k}^{(2k+5,1)}(-1) \int_{-1}^1 (1-t)^2 (1+t)^2 P_k^{(2,2)}(t) P_j^{(2,2)}(t) dt \\ &= -\frac{|\nabla \lambda_i|}{32} P_{r-2-k}^{(2k+5,1)}(-1) h_k^{(2,2)} \delta_{jk}. \end{aligned}$$

This gives

$$a_{k,r-2} = -\frac{32}{|\nabla \lambda_i| P_{r-2-k}^{(2k+5,1)}(-1) h_k^{(2,2)}}.$$

By [Sze75, Eq. (4.1.1),(4.1.3)], we obtain  $P_{r-2-k}^{(2k+5,1)}(-1) = (-1)^{r-k}(r-k-1)$ . Using (3.8), we obtain

$$h_k^{(2,2)} = \frac{32(k+1)(k+2)}{(k+3)(k+4)(2k+5)}.$$

A combination of the above three identities leads to (3.12).  $\square$

Next we list the shape functions for the elements of low-order.

**Example 3.6.** The bubble space  $b_K \sum_{i=1}^3 b_i Q_i^{r-2}(K)$  for the lowest-order  $r = 2$  is

$$b_K \text{span} \{ b_i \mid i = 1, 2, 3 \}.$$

The shape functions associated with  $\int_{e_i} \partial_{\mathbf{n}} v d\sigma(\mathbf{x})$  is  $-\frac{30}{|\nabla \lambda_i|} b_K b_i$ .

The bubble space  $b_K \sum_{i=1}^3 b_i Q_i^{r-2}(K)$  for the case  $r = 3$  is

$$b_K \text{span} \{ b_i(3\lambda_i - \lambda_i^+ - \lambda_i^-), b_i(\lambda_i^+ - \lambda_i^-) \mid i = 1, 2, 3 \}.$$

The shape functions associated with  $\int_{e_i} \partial_{\mathbf{n}} v d\sigma(\mathbf{x})$  is

$$\frac{30}{|\nabla \lambda_i|} b_K b_i (3\lambda_i - \lambda_i^+ - \lambda_i^-).$$

The shape functions associated with  $3 \int_{e_i} \partial_{\mathbf{n}} v (\lambda_i^+ - \lambda_i^-) d\sigma(\mathbf{x})$  is

$$-\frac{70}{|\nabla \lambda_i|} b_K b_i (\lambda_i^+ - \lambda_i^-).$$

**Lemma 3.7.** *The space  $R^{r-2}(K)$  takes the forms*

$$R^{r-2}(K) = \text{span} \left\{ P_{k,r-2}^{(2,2,2)}(\lambda_1, \lambda_2, \lambda_3) \mid 0 \leq k \leq r-2 \right\}.$$

*Proof.* For any  $v \in R^{r-2}(K) \subset \mathbb{P}_{r-2}(K)$ , we expand  $v$  into

$$v = \sum_{0 \leq k \leq n \leq r-2} a_{kn} P_{k,n}^{(2,2,2)}(\lambda_1, \lambda_2, \lambda_3)$$

with  $a_{kn}$  to be determined later on. Using the above representation, we may write the constraint in the definition of the space (3.3) as: For all  $q \in \mathbb{P}_{r-3}(K)$ ,

$$\sum_{0 \leq k \leq n \leq r-2} a_{kn} \int_K b_K^2 P_{k,n}^{(2,2,2)}(\lambda_1, \lambda_2, \lambda_3) q dx = 0.$$

We substitute  $q = P_{j,m}^{(2,2,2)}(\lambda_1, \lambda_2, \lambda_3)$  for  $0 \leq j \leq m \leq r-3$  into the above equation. Using the orthogonal relation (3.10), we may obtain  $a_{jm} = 0$  for  $0 \leq j \leq m \leq r-3$ . This completes the proof.  $\square$

Motivated by the above lemma, we change the definition of DoFs for  $b_K^2 R^{r-2}(K)$  from  $\int_K vq \, d\mathbf{x}$  for any  $q \in \mathbb{P}_{r-2}(K) \setminus \mathbb{P}_{r-3}(K)$  to

$$\int_K v P_{k,r-2}^{(2,2,2)}(\lambda_1, \lambda_2, \lambda_3) \, d\mathbf{x}, \quad k = 0, \dots, r-2.$$

**Lemma 3.8.** *The shape functions for the bubble space  $b_K^2 R^{r-2}(K)$  associated with the above definition of DoFs are*

$$a_{k,r-2} b_K^2 P_{k,r-2}^{(2,2,2)}(\lambda_1, \lambda_2, \lambda_3)$$

with

$$a_{k,r-2} = \frac{(2r+4)(2k+5)(r+k+4)(r+k+5)(k+3)(k+4)}{2(r-k)(r-1-k)(k+1)(k+2)} \quad k = 0, \dots, r-2.$$

*Proof.* For  $j = 1, \dots, r-2$ , we obtain

$$\int_K b_K^2 P_{k,r-2}^{(2,2,2)}(\lambda_1, \lambda_2, \lambda_3) P_{j,r-2}^{(2,2,2)}(\lambda_1, \lambda_2, \lambda_3) \, d\mathbf{x} = 2h_{k,r-2}^{(2,2,2)} \delta_{jk},$$

which gives

$$a_{k,r-2} = \frac{1}{2h_{k,r-2}^{(2,2,2)}}.$$

Using (3.11), we obtain

$$h_{k,r-2}^{(2,2,2)} = \frac{(r-k)(r-1-k)(k+1)(k+2)}{(2r+4)(2k+5)(r+k+4)(r+k+5)(k+3)(k+4)}.$$

These give the simplified expression of  $a_{k,r-2}$ .  $\square$

According to the definition of bubble space, we may have

**Example 3.9.** The bubble space  $b_K^2 R^{r-2}(K)$  for the lowest-order case  $r = 2$  is  $\text{span}\{b_K^2\}$ .

The shape functions associated with  $\int_K v \, d\mathbf{x}$  is  $2520b_K^2$ .

The bubble space  $b_K^2 R^{r-2}(K)$  for the case  $r = 3$  is  $b_K^2 \text{span}\{2\lambda_1 - \lambda_2 - \lambda_3, \lambda_2 - \lambda_3\}$ .

The shape functions associated with  $3 \int_K v(2\lambda_1 - \lambda_2 - \lambda_3) \, d\mathbf{x}$  is  $4200b_K^2(2\lambda_1 - \lambda_2 - \lambda_3)$ .

The shape functions associated with  $3 \int_K v(\lambda_2 - \lambda_3) \, d\mathbf{x}$  is  $12600b_K^2(\lambda_2 - \lambda_3)$ .

*Remark 3.10.* Based on the above results, we give the details for the element of the lowest-order, i.e.,  $r = 2$ , which have been used in the numerical examples. The local finite element space

$$P_K = \mathbb{P}_2(K) + b_K \sum_{i=1}^3 \text{span}\{b_i\} + \text{span}\{b_K^2\},$$

and DoFs

$$\Sigma_K = \left\{ v(a_i), \int_{e_i} v \, d\sigma(\mathbf{x}), \int_{e_i} \partial_{\mathbf{n}} v \, d\sigma(\mathbf{x}), \int_K v \, d\mathbf{x} \mid i = 1, 2, 3 \right\}.$$

The shape functions associated with  $\{v(a_i)\}_{i=1,2,3}$  are

$$\phi_i = \lambda_i(3\lambda_i - 2) + 30b_K \left( 2b_i + \sum_{j \neq i} \frac{\nabla \lambda_i \cdot \nabla \lambda_j}{|\nabla \lambda_j|^2} b_j(4\lambda_j - 1) + 6b_K \right).$$

The shape functions associated with  $\{\int_{e_i} v \, d\sigma(\mathbf{x})\}_{i=1,2,3}$  are

$$\varphi_i = 6b_i + 90b_K \left( b_i - \sum_{j \neq i} b_j - 10b_K \right).$$

The shape functions associated with  $\{\int_{e_i} \partial_{\mathbf{n}} v d\sigma(\mathbf{x})\}_{i=1,2,3}$  are

$$\psi_i = \frac{30}{|\nabla \lambda_i|} b_K b_i (4\lambda_i - 1).$$

The shape functions associated with  $\int_K v d\mathbf{x}$  is  $\phi_0 = 2520b_K^2$ .

#### 4. THE MIXED FINITE ELEMENTS APPROXIMATION

In this part, we construct stable finite element pairs to approximate (2.2). We ignore the influence of the curved boundary for error estimates for brevity. Define

$$X_h := \left\{ v \in H_0^1(\Omega) \mid v|_K \in P_K \text{ for all } K \in \mathcal{T}_h, \int_e [\partial_{\mathbf{n}} v] q d\sigma(\mathbf{x}) = 0 \text{ for all } e \in \mathcal{E}_h, q \in \mathbb{P}_{r-2}(e) \right\},$$

and  $V_h := [X_h]^2$ . Let  $P_h \subset P$  be the continuous Lagrangian finite element of order  $r-1$ . We shall prove a uniform discrete B-B inequality for the pair  $(V_h, P_h)$ .

The following rescaled trace inequality will be used later on: There exists  $C$  independent of  $h_K$  such that

$$(4.1) \quad \|v\|_{L^2(\partial K)} \leq C \left( h_K^{-1/2} \|v\|_{L^2(K)} + \|v\|_{L^2(K)}^{1/2} \|\nabla \mathbf{v}\|_{L^2(K)}^{1/2} \right).$$

This inequality may be found in [BS08].

**4.1. The mixed finite element approximation.** We define the mixed finite element approximation problem as follows. Find  $\mathbf{u}_h \in V_h$  and  $p_h \in P_h$  such that

$$(4.2) \quad \begin{cases} a_{\iota,h}(\mathbf{u}_h, \mathbf{v}) + b_{\iota,h}(\mathbf{v}, p_h) = (\mathbf{f}, \mathbf{v}) & \text{for all } \mathbf{v} \in V_h, \\ b_{\iota,h}(\mathbf{u}_h, q) - \lambda^{-1} c_{\iota}(p_h, q) = 0 & \text{for all } q \in P_h, \end{cases}$$

where

$$\begin{aligned} a_{\iota,h}(\mathbf{v}, \mathbf{w}) &:= 2\mu \left( (\boldsymbol{\epsilon}(\mathbf{v}), \boldsymbol{\epsilon}(\mathbf{w})) + \iota^2 (\nabla_h \boldsymbol{\epsilon}(\mathbf{v}), \nabla_h \boldsymbol{\epsilon}(\mathbf{w})) \right) & \text{for all } \mathbf{v}, \mathbf{w} \in V_h, \\ b_{\iota,h}(\mathbf{v}, q) &:= (\operatorname{div} \mathbf{v}, q) + \iota^2 (\nabla_h \operatorname{div} \mathbf{v}, \nabla q) & \text{for all } \mathbf{v} \in V_h, q \in P_h. \end{aligned}$$

Note that  $V_h \not\subset V$ , and this is a nonconforming method, we introduce the broken norm

$$\|\nabla \mathbf{v}\|_{\iota,h} := \|\nabla \mathbf{v}\|_{L^2} + \iota \|\nabla_h^2 \mathbf{v}\|_{L^2} \quad \text{for all } \mathbf{v} \in V_h$$

Due to the continuity of  $\mathbf{v}$ ,  $\|\nabla \mathbf{v}\|_{\iota,h}$  is a norm over  $V_h$ .

The following broken Korn's inequality was proved in [LMW21, Theorem 2]:

$$\|\nabla_h \boldsymbol{\epsilon}(\mathbf{v})\|_{L^2} \geq \left(1 - 1/\sqrt{2}\right) \|\nabla_h^2 \mathbf{v}\|_{L^2},$$

which together with the first Korn's inequality (2.6) gives

$$(4.3) \quad a_{\iota,h}(\mathbf{v}, \mathbf{v}) \geq \frac{\mu}{2} \|\nabla \mathbf{v}\|_{\iota,h}^2 \quad \text{for all } \mathbf{v} \in V_h.$$

It remains to prove the discrete B-B inequality for the pair  $(V_h, P_h)$ . To this end, we construct a Fortin operator that is uniformly stable in the weighted norm  $\|\nabla \cdot\|_{\iota,h}$  [MSW13]. The key is to construct different Fortin operators for  $\iota/h$  in different regimes.

Firstly we define an interpolation operator  $\Pi_h : V \rightarrow V_h$  by  $\Pi_h|_K := \Pi_K = [\pi_K]^2$ , which satisfies

**Lemma 4.1.** *For all  $\mathbf{v} \in V$ , there holds*

$$(4.4) \quad b_{\iota,h}(\Pi_h \mathbf{v}, p) = b_{\iota}(\mathbf{v}, p) \quad \text{for all } p \in P_h.$$

*Proof.* Using the fact that  $\Pi_h \mathbf{v} \in V_h \subset [H_0^1(\Omega)]^2$ , an integration by parts gives

$$(4.5) \quad \int_{\Omega} \operatorname{div}(\mathbf{v} - \Pi_h \mathbf{v}) p \, d\mathbf{x} = - \sum_{K \in \mathcal{T}_h} \int_K (\mathbf{v} - \Pi_K \mathbf{v}) \cdot \nabla p \, d\mathbf{x} = 0,$$

where we have used the identity (3.5)<sub>4</sub> in the last step.

Next, integration by parts yields

$$\int_{\Omega} \nabla \operatorname{div}(\mathbf{v} - \Pi_h \mathbf{v}) \cdot \nabla p \, d\mathbf{x} = \sum_{K \in \mathcal{T}_h} \int_{\partial K} \operatorname{div}(\mathbf{v} - \Pi_K \mathbf{v}) \partial_{\mathbf{n}} p \, d\sigma(\mathbf{x}) - \sum_{K \in \mathcal{T}_h} \int_K \operatorname{div}(\mathbf{v} - \Pi_K \mathbf{v}) \Delta p \, d\mathbf{x} = 0.$$

The first term vanishes because  $\partial_j = t_j \partial_{\mathbf{t}} + n_j \partial_{\mathbf{n}}$  for components  $j = 1, 2$ , and using (3.5)<sub>2</sub>, we obtain that for each edge  $e \in \partial K$ ,

$$\int_e t_j \frac{\partial}{\partial \mathbf{t}} (v_j - \pi_K v_j) \partial_{\mathbf{n}} p \, d\sigma(\mathbf{x}) = - \int_e t_j (v_j - \pi_K v_j) \frac{\partial^2 p}{\partial \mathbf{t} \partial \mathbf{n}} \, d\sigma(\mathbf{x}) = 0.$$

Using (3.5)<sub>3</sub>, we obtain

$$\int_e n_j \frac{\partial}{\partial \mathbf{n}} (v_j - \pi_K v_j) \partial_{\mathbf{n}} p \, d\sigma(\mathbf{x}) = 0.$$

While the second term vanishes because

$$\int_K \operatorname{div}(\mathbf{v} - \Pi_K \mathbf{v}) \Delta p \, d\mathbf{x} = \int_{\partial K} (\mathbf{v} - \Pi_K \mathbf{v}) \cdot \mathbf{n} \Delta p \, d\sigma(\mathbf{x}) - \int_K (\mathbf{v} - \Pi_K \mathbf{v}) \cdot \nabla \Delta p \, d\mathbf{x} = 0,$$

where we have used (3.5)<sub>2</sub> and (3.5)<sub>4</sub>. □

The operator  $\Pi_h$  is not  $H^1$ -bounded by (3.6), and we construct an  $H^1$ -bounded regularized interpolation operator as follows.

**Lemma 4.2.** *There exists an operator  $I_h : V \mapsto V_h$  satisfying*

$$(4.6) \quad \int_{\Omega} \operatorname{div}(\mathbf{v} - I_h \mathbf{v}) p \, d\mathbf{x} = 0 \quad \text{for all } p \in P_h,$$

and if  $\mathbf{v} \in V \cap [H^s(\Omega)]^2$  with  $1 \leq s \leq r + 1$ , then

$$(4.7) \quad \|\nabla_h^j (\mathbf{v} - I_h \mathbf{v})\|_{L^2} \leq Ch^{s-j} \|\nabla^s \mathbf{v}\|_{L^2} \quad 0 \leq j \leq s.$$

The construction of  $I_h$  is based on a regularized interpolation operator in [GLN12] and the standard construction of the Fortin operator [For77]. The operator  $I_h$  is also well-defined for functions in  $[H^2(\Omega) \cap H_0^1(\Omega)]^2$ .

*Proof.* Define  $I_h : V \mapsto V_h$  with  $I_h := [\Pi_1]^2$  and

$$\Pi_1 := \Pi_0(I - \Pi_2) + \Pi_2,$$

where the regularized interpolation operator  $\Pi_2 : H_0^2(\Omega) \mapsto X_h$  was constructed in [GLN12, Lemma 2], which satisfies

$$(4.8) \quad \|\nabla_h^j (v - \Pi_2 v)\|_{L^2} \leq Ch^{s-j} \|\nabla^s v\|_{L^2}, \quad 1 \leq s \leq r + 1, 0 \leq j \leq s.$$

The operator  $\Pi_0 : H_0^1(\Omega) \mapsto X_h$  is defined for any element  $K \in \mathcal{T}_h$  as

$$\left\{ \begin{array}{l} \Pi_0 v(a) = 0 \quad \text{for all vertices } a, \\ \int_e \Pi_0 v q \, d\sigma(\mathbf{x}) = 0 \quad \forall q \in \mathbb{P}_{r-2}(e) \quad \text{for all edges } e, \\ \int_e \partial_{\mathbf{n}} \Pi_0 v q \, d\sigma(\mathbf{x}) = 0 \quad \forall q \in \mathbb{P}_{r-2}(e) \quad \text{for all edges } e, \\ \int_K \Pi_0 v q \, d\mathbf{x} = \int_K v q \, d\mathbf{x} \quad \text{for all } q \in \mathbb{P}_{r-2}(K). \end{array} \right.$$

On each element  $K$ , we have  $(\Pi_0 v)|_K \in H_0^1(K)$  for any  $v \in H_0^1(\Omega)$ , and hence  $\Pi_0 v \in X_h$ . A standard scaling argument gives

$$\|\Pi_0 v\|_{L^2} \leq C\|v\|_{L^2} \quad \text{for all } v \in H_0^1(\Omega).$$

For any  $\mathbf{v} \in V$ , using the fact that  $I_h \mathbf{v} \in V_h$  and  $p \in P_h$ , an integration by parts gives

$$\begin{aligned} \int_{\Omega} \operatorname{div}(\mathbf{v} - I_h \mathbf{v}) p \, d\mathbf{x} &= - \sum_{K \in \mathcal{T}_h} \int_K (\mathbf{v} - I_h \mathbf{v}) \cdot \nabla p \, d\mathbf{x} + \sum_{K \in \mathcal{T}_h} \int_{\partial K} (\mathbf{v} - I_h \mathbf{v}) \cdot \mathbf{n} p \, d\sigma(\mathbf{x}) \\ &= - \sum_{K \in \mathcal{T}_h} \int_K (\mathbf{v} - I_h \mathbf{v}) \cdot \nabla p \, d\mathbf{x} = - \sum_{K \in \mathcal{T}_h} \int_K (I - \Pi_0)(I - \Pi_2) \mathbf{v} \cdot \nabla p \, d\mathbf{x} \\ &= 0, \end{aligned}$$

where we have used the last property of  $\Pi_0$ . This gives (4.6).

Using (4.8), the  $L^2$ -stability of  $\Pi_0$  and the inverse inequality, we obtain

$$\begin{aligned} \|\nabla_h^j(\mathbf{v} - \Pi_1 \mathbf{v})\|_{L^2} &\leq \|\nabla_h^j(\mathbf{v} - \Pi_2 \mathbf{v})\|_{L^2} + \|\nabla_h^j \Pi_0(\mathbf{v} - \Pi_2 \mathbf{v})\|_{L^2} \\ &\leq Ch^{s-j} \|\nabla^s \mathbf{v}\|_{L^2} + Ch^{-j} \|\mathbf{v} - \Pi_2 \mathbf{v}\|_{L^2} \\ &\leq Ch^{s-j} \|\nabla^s \mathbf{v}\|_{L^2}. \end{aligned}$$

This implies (4.7) and completes the proof.  $\square$

We are ready to prove the following discrete B-B inequality.

**Theorem 4.3.** *There exists  $\beta$  independent of  $\iota$  and  $h$ , such that*

$$(4.9) \quad \sup_{\mathbf{v} \in V_h} \frac{b_{\iota, h}(\mathbf{v}, p)}{\|\nabla \mathbf{v}\|_{\iota, h}} \geq \beta \|p\|_{\iota} \quad \text{for all } p \in P_h.$$

*Proof.* Using (2.3), for any  $p \in P_h \subset P$ , there exists  $\mathbf{v}_0 \in V$  such that

$$b_{\iota}(\mathbf{v}_0, p) = \|p\|_{\iota}^2 \quad \text{and} \quad \|\nabla \mathbf{v}_0\|_{\iota} \leq C\|p\|_{\iota}.$$

First, we consider the case  $\iota/h \leq \gamma$  with  $\gamma$  to be determined later on. By (4.7), we obtain

$$\|\nabla I_h \mathbf{v}_0\|_{\iota, h} \leq \|\nabla \mathbf{v}_0\|_{\iota} + \|\nabla(\mathbf{v}_0 - I_h \mathbf{v}_0)\|_{\iota, h} \leq C\|\nabla \mathbf{v}_0\|_{\iota},$$

and

$$\|\nabla_h \operatorname{div}(\mathbf{v}_0 - I_h \mathbf{v}_0)\|_{L^2} \leq C\|\nabla^2 \mathbf{v}_0\|_{L^2} \leq C\|\nabla p\|_{L^2}.$$

Combining the above inequality and using the inverse inequality for any  $p \in P_h$ , we obtain

$$\iota^2 |(\nabla_h \operatorname{div}(\mathbf{v}_0 - I_h \mathbf{v}_0), \nabla p)| \leq C\iota^2 \|\nabla p\|_{L^2}^2 \leq C_*(\iota/h)^2 \|p\|_{L^2}^2 \leq \gamma^2 C_* \|p\|_{\iota}^2.$$

Fix  $\gamma$  such that  $\gamma^2 C_* < 1$ , we obtain

$$b_{\iota, h}(I_h \mathbf{v}_0, p) = b_{\iota}(\mathbf{v}_0, p) - \iota^2 (\nabla_h \operatorname{div}(\mathbf{v}_0 - I_h \mathbf{v}_0), \nabla p) \geq (1 - \gamma^2 C_*) \|p\|_{\iota}^2.$$

This gives

$$(4.10) \quad \sup_{\mathbf{v} \in V_h} \frac{b_{\iota, h}(\mathbf{v}, p)}{\|\nabla \mathbf{v}\|_{\iota, h}} \geq \frac{b_{\iota, h}(I_h \mathbf{v}_0, p)}{\|\nabla I_h \mathbf{v}_0\|_{\iota, h}} \geq \frac{1 - \gamma^2 C_*}{C} \|p\|_{\iota}.$$

Next, if  $\iota/h > \gamma$ , then we use (3.6) and obtain

$$\|\nabla(\mathbf{v}_0 - \Pi_h \mathbf{v}_0)\|_{L^2} \leq Ch\|\nabla^2 \mathbf{v}_0\|_{L^2}, \quad \text{and} \quad \|\nabla_h^2(\mathbf{v}_0 - \Pi_h \mathbf{v}_0)\|_{L^2} \leq C\|\nabla^2 \mathbf{v}_0\|_{L^2}.$$

Therefore,

$$\|\nabla(\mathbf{v}_0 - \Pi_h \mathbf{v}_0)\|_{\iota, h} \leq C(h + \iota)\|\nabla^2 \mathbf{v}_0\|_{L^2} \leq C(1 + h/\iota)\|\nabla \mathbf{v}_0\|_{\iota} \leq C(1 + 1/\gamma)\|\nabla \mathbf{v}_0\|_{\iota}.$$

Hence,

$$\|\nabla \Pi_h \mathbf{v}_0\|_{\iota, h} \leq \|\nabla \mathbf{v}_0\|_{\iota} + \|\nabla(\mathbf{v}_0 - \Pi_h \mathbf{v}_0)\|_{\iota, h} \leq C(2 + 1/\gamma)\|\nabla \mathbf{v}_0\|_{\iota},$$

which together with (4.4) gives

$$(4.11) \quad \sup_{v \in V_h} \frac{b_{\iota,h}(\mathbf{v}, p)}{\|\nabla \mathbf{v}\|_{\iota,h}} \geq \frac{b_{\iota,h}(\Pi_h \mathbf{v}_0, p)}{\|\nabla \Pi_h \mathbf{v}_0\|_{\iota,h}} = \frac{b_{\iota,h}(\mathbf{v}_0, p)}{\|\nabla \Pi_h \mathbf{v}_0\|_{\iota,h}} = \frac{\|p\|_{\iota}^2}{\|\nabla \Pi_h \mathbf{v}_0\|_{\iota,h}} \geq \frac{\gamma}{C(1+2\gamma)} \|p\|_{\iota}.$$

A combination of (4.10) and (4.11) shows that (4.9) holds true with  $\beta$  independent of  $\iota$  and  $h$ .  $\square$

The well-posedness of the mixed approximation problem (4.2) follows from the ellipticity of  $a_{\iota,h}$  and the discrete B-B inequality of  $b_{\iota,h}$ . We are ready to derive the error estimate.

**4.2. Error estimates.** To carry out the error estimate, we define the bilinear form  $\mathcal{A}$  as

$$\mathcal{A}(\mathbf{v}, q; \mathbf{w}, z) := a_{\iota,h}(\mathbf{v}, \mathbf{w}) + b_{\iota,h}(\mathbf{w}, q) + b_{\iota,h}(\mathbf{v}, z) - \lambda^{-1} c_{\iota}(q, z)$$

for all  $\mathbf{v}, \mathbf{w} \in V_h$  and  $q, z \in P_h$ .

We prove the following inf-sup inequality for  $\mathcal{A}$  with the aid of the discrete B-B inequality (4.9).

**Lemma 4.4.** *There exists  $\alpha$  depending on  $\mu$  and  $\beta$  such that*

$$(4.12) \quad \inf_{(\mathbf{v}, q) \in V_h \times P_h} \sup_{(\mathbf{w}, z) \in V_h \times P_h} \frac{\mathcal{A}(\mathbf{v}, q; \mathbf{w}, z)}{\|(\mathbf{w}, z)\|_{\iota,h} \|(\mathbf{v}, q)\|_{\iota,h}} \geq \alpha,$$

where  $\|(\mathbf{w}, z)\|_{\iota,h} := \|\nabla \mathbf{w}\|_{\iota,h} + \|z\|_{\iota} + \lambda^{-1/2} \|z\|_{\iota}$  and  $\beta$  has appeared in (4.9).

*Proof.* Noting that  $a_{\iota,h}$  is elliptic over  $V_h$  (4.3) and the discrete B-B inequality for  $b_{\iota,h}$  holds (4.9), we obtain (4.12) by [Bra96, Theorem 2].  $\square$

We are ready to prove error estimates.

**Theorem 4.5.** *There exists  $C$  independent of  $\iota, \lambda$  and  $h$  such that*

$$(4.13) \quad \|(\mathbf{u} - \mathbf{u}_h, p - p_h)\|_{\iota,h} \leq C(h^r + \iota h^{r-1})(\|\mathbf{u}\|_{H^{r+1}} + \|p\|_{H^r}),$$

and

$$(4.14) \quad \|(\mathbf{u} - \mathbf{u}_h, p - p_h)\|_{\iota,h} \leq Ch^{1/2} \|\mathbf{f}\|_{L^2}.$$

*Proof.* Let  $\mathbf{v} = \mathbf{u}_h - \mathbf{u}_I$  and  $q = p_h - p_I$  with  $\mathbf{u}_I \in V_h$  and  $p_I \in P_h$ , for any  $\mathbf{w} \in V_h$  and  $z \in P_h$ ,

$$\begin{aligned} \mathcal{A}(\mathbf{v}, q; \mathbf{w}, z) &= \mathcal{A}(\mathbf{u}_h, p_h; \mathbf{w}, z) - \mathcal{A}(\mathbf{u}, p; \mathbf{w}, z) + \mathcal{A}(\mathbf{u} - \mathbf{u}_I, p - p_I; \mathbf{w}, z) \\ &= (\mathbf{f}, \mathbf{w}) - \mathcal{A}(\mathbf{u}, p; \mathbf{w}, z) + \mathcal{A}(\mathbf{u} - \mathbf{u}_I, p - p_I; \mathbf{w}, z) \\ &= \mathcal{A}(\mathbf{u} - \mathbf{u}_I, p - p_I; \mathbf{w}, z) - \iota^2 \sum_{e \in \mathcal{E}_h} \int_e (\partial_n \boldsymbol{\sigma} \mathbf{n}) \cdot [[\partial_n \mathbf{w}]] d\sigma(\mathbf{x}). \end{aligned}$$

The boundedness of  $\mathcal{A}$  yields

$$|\mathcal{A}(\mathbf{u} - \mathbf{u}_I, p - p_I; \mathbf{w}, z)| \leq \max(1, 2\mu) \|(\mathbf{u} - \mathbf{u}_I, p - p_I)\|_{\iota,h} \|(\mathbf{w}, z)\|_{\iota,h}.$$

Let  $\mathbf{u}_I = \Pi_h \mathbf{u}$  be the interpolation of  $\mathbf{u}$  and  $p_I$  be the  $r-1$  order Lagrangian interpolation of  $p$ , respectively. The standard interpolation error estimates in (3.6) gives

$$\|(\mathbf{u} - \mathbf{u}_I, p - p_I)\|_{\iota,h} \leq C(h^r + \iota h^{r-1})(\|\mathbf{u}\|_{H^{r+1}} + \|p\|_{H^r}).$$

Note that

$$\int_e [[\partial_n \mathbf{w}]] q d\sigma(\mathbf{x}) = 0 \quad \text{for all } q \in \mathbb{P}_{r-2}(e).$$

A standard estimate for the consistency error functional with trace inequality (4.1) gives

$$\begin{aligned} \iota^2 \left| \sum_{e \in \mathcal{E}_h} \int_e (\partial_n \boldsymbol{\sigma} \mathbf{n}) \cdot [[\partial_n \mathbf{w}]] d\sigma(\mathbf{x}) \right| &\leq C \iota^2 h^{r-1} (\|\mathbf{u}\|_{H^{r+1}} + \|p\|_{H^r}) \|\nabla_h^2 \mathbf{w}\|_{L^2} \\ &\leq C \iota h^{r-1} (\|\mathbf{u}\|_{H^{r+1}} + \|p\|_{H^r}) \|\nabla \mathbf{w}\|_{\iota,h}. \end{aligned}$$

A combination of the above three inequalities, the discrete inf-sup condition (4.12) and the triangle inequalities gives (4.13).

Next, let  $\mathbf{u}_I = I_h \mathbf{u}$  and let  $p_I$  be the Clément interpolation [Clé75] of  $p$ , respectively. The interpolation error (4.7) and the error estimates for the Clément interpolation give

$$\|(\mathbf{u} - \mathbf{u}_I, p - p_I)\|_{\iota, h} \leq Ch^{1/2} \left( \|\mathbf{u}\|_{H^{3/2}} + \|p\|_{H^{1/2}} + \iota (\|\mathbf{u}\|_{H^{5/2}} + \|p\|_{H^{3/2}}) \right) \leq Ch^{1/2} \|\mathbf{f}\|_{L^2},$$

where we have used (2.29) and (2.30) in the last step.

Using the trace inequality (4.1), we bound the consistency error functional as

$$\begin{aligned} \iota^2 \left| \sum_{e \in \mathcal{E}_h} \int_e (\partial_n \boldsymbol{\sigma} \mathbf{n}) \cdot \llbracket \partial_n \mathbf{w} \rrbracket d\sigma(\mathbf{x}) \right| &\leq C \iota^2 h^{1/2} (\|\mathbf{u}\|_{H^2} + \|p\|_{H^1})^{1/2} (\|\mathbf{u}\|_{H^3} + \|p\|_{H^2})^{1/2} \|\nabla_h^2 \mathbf{w}\|_{L^2} \\ &\leq Ch^{1/2} \|\mathbf{f}\|_{L^2} \|\nabla \mathbf{w}\|_{\iota, h}, \end{aligned}$$

where we have used (2.28) in the last step.

Combining these inequalities, the discrete inf-sup condition (4.12) and the triangle inequalities gives (4.14).  $\square$

**Corollary 4.6.** *There exists  $C$  independent of  $\iota, \lambda$  and  $h$  such that*

$$(4.15) \quad \|(\mathbf{u}_0 - \mathbf{u}_h, p_0 - p_h)\|_{\iota, h} \leq C(\iota^{1/2} + h^{1/2}) \|\mathbf{f}\|_{L^2},$$

where  $\mathbf{u}_0$  is the solution of (2.8), and  $p_0 = \lambda \operatorname{div} \mathbf{u}_0$ .

*Proof.* A combination of Theorem 2.9, Theorem 4.5, and the triangle inequality gives (4.15).  $\square$

## 5. NUMERICAL EXAMPLES

In this part, we report the numerical performance for the proposed element of the lowest-order, i.e.,  $r = 2$ . We test the accuracy and robustness of the element pair for the nearly incompressible materials. All examples are carried out on the nonuniform mesh. We are interested in the case when the Poisson's ratio  $\nu$  is close to 0.5 and we report the relative errors  $\|\nabla(\mathbf{u} - \mathbf{u}_h)\|_{\iota, h} / \|\nabla \mathbf{u}\|_{\iota}$  and the rates of convergence.

We let  $\Omega = (0, 1)^2$ , and set Young's modulus  $E = 1$ . The Lamé constants are determined by

$$\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)}, \quad \mu = \frac{E}{2(1+\nu)}.$$

We set  $\nu = 0.3$  for the ordinary cases, hence  $\lambda = 0.5769$ , and  $\mu = 0.3846$ , and we set  $\nu = 0.4999$  for the nearly incompressible materials, hence  $\lambda = 1.6664e3$ , and  $\mu = 0.3334$ .

**5.1. The first example.** We test the performance of the element pair by solving a completely incompressible problem, which means  $\operatorname{div} \mathbf{u} = 0$ . Let  $\mathbf{u} = (u_1, u_2)$  with

$$u_1 = -\sin^3(\pi x) \sin(2\pi y) \sin(\pi y), \quad u_2 = \sin(2\pi x) \sin(\pi x) \sin^3(\pi y).$$

Therefore  $\operatorname{div} \mathbf{u} = 0$ , and  $\mathbf{f}$  is independent of  $\lambda$ .

In view of Table 2, the optimal rates of convergence are observed with the completely incompressible media, which is consistent with the error bound (4.13).

TABLE 2. Relative errors and convergence rates for the 1st example

$\iota \backslash h$	1/8	1/16	1/32	1/64
$\nu = 0.3, \lambda = 0.5769, \mu = 0.3846$				
1e+00	2.592e-01	1.333e-01	6.519e-02	3.246e-02
rate		0.96	1.03	1.01
1e-06	4.252e-02	1.159e-02	2.784e-03	6.918e-04
rate		1.88	2.06	2.01
$\nu = 0.4999, \lambda = 1.6664e3, \mu = 0.3334$				
1e+00	2.592e-01	1.333e-01	6.519e-02	3.246e-02
rate		0.96	1.03	1.01
1e-06	4.252e-02	1.159e-02	2.784e-03	6.918e-04
rate		1.88	2.06	2.01

5.2. **The second example.** This example is motivated by [Wih06], which admits a singular solution. The exact solution  $\mathbf{u} = (u_1, u_2)$  expressed in the polar coordinates as

$$u_1 = u_\rho(\rho, \theta) \cos \theta - u_\theta(\rho, \theta) \sin \theta, \quad u_2 = u_\rho(\rho, \theta) \sin \theta + u_\theta(\rho, \theta) \cos \theta,$$

where

$$u_\rho = \frac{1}{2\mu} \rho^\alpha \left( -(\alpha + 1) \cos((\alpha + 1)\theta) + (C_2 - (\alpha + 1))C_1 \cos((\alpha - 1)\theta) \right),$$

$$u_\theta = \frac{1}{2\mu} \rho^\alpha \left( (\alpha + 1) \sin((\alpha + 1)\theta) + (C_2 + \alpha - 1)C_1 \sin((\alpha - 1)\theta) \right),$$

and  $\alpha = 1.5, \omega = 3\pi/4$ ,

$$C_1 = -\frac{\cos((\alpha + 1)\omega)}{\cos((\alpha - 1)\omega)} \quad \text{and} \quad C_2 = \frac{2(\lambda + 2\mu)}{\lambda + \mu}.$$

It may be verified that  $\mathbf{u} \in [H^{5/2-\varepsilon}(\Omega)]^2$  for a small number  $\varepsilon > 0$ . A direct calculation gives that  $\mathbf{f} \equiv \mathbf{0}$ , while it is nearly incompressible because

$$\operatorname{div} \mathbf{u} = -\frac{3(1 + \sqrt{2})}{\lambda + \mu} \rho^{1/2} \cos(\theta/2).$$

TABLE 3. Relative errors and convergence rates for the 2nd example

$\iota \backslash h$	1/8	1/16	1/32	1/64
$\nu = 0.3, \lambda = 0.5769, \mu = 0.3846$				
1e+00	1.062e-01	7.554e-02	5.355e-02	3.792e-02
rate		0.49	0.50	0.50
1e-06	2.809e-03	1.001e-03	3.549e-04	1.257e-04
rate		1.49	1.50	1.50
$\nu = 0.4999, \lambda = 1.6664e3, \mu = 0.3334$				
1e+00	1.149e-01	8.200e-02	5.824e-02	4.135e-02
rate		0.49	0.49	0.49
1e-06	4.399e-03	1.567e-03	5.558e-04	1.968e-04
rate		1.49	1.50	1.50

It follows from Table 3 that the rates of convergence are sub-optimal. It is reasonable because the solution  $\mathbf{u}$  is singular, which is similar to the results in [Wih06]. The element pair is robust for the nearly incompressible materials.

**5.3. The third example.** In the last example, we test a problem with strong boundary layer effects. Such effects have been frequently observed in the stain elasticity model [EGH<sup>+</sup>02, LMS17, LM19, LMW21]. It is shown that the numerical solution converges to the solution of (2.8) when  $\iota \ll h$ .

When  $\iota \rightarrow 0$ , the boundary value problem (1.1) reduces to (2.8). Let  $\mathbf{u}_0 = (u_1^0, u_2^0)$  with

$$u_1^0 = -\sin^2(\pi x) \sin(2\pi y), \quad u_2^0 = \sin(2\pi x) \sin^2(\pi y)$$

be the solution of problem (2.8). The source term  $\mathbf{f}$  is computed from (2.8). A direct calculation gives that  $\operatorname{div} \mathbf{u}_0 = 0$ , and  $\mathbf{f}$  is independent of  $\lambda$ . The exact solution  $\mathbf{u}$  for (1.1) is unknown, while it has strong boundary layer effects. In this case, we take  $\iota \ll h$ , and report the relative error  $\|\nabla(\mathbf{u}_0 - \mathbf{u}_h)\|_{\iota, h} / \|\nabla \mathbf{u}_0\|_{\iota}$ .

TABLE 4. Relative errors and convergence rates for the 3rd example

$\iota \backslash h$	1/8	1/16	1/32	1/64
$\nu = 0.3, \lambda = 0.5769, \mu = 0.3846$				
1e-04	1.311e-01	8.966e-02	6.299e-02	4.476e-02
rate		0.55	0.51	0.49
1e-06	1.311e-01	8.960e-02	6.283e-02	4.432e-02
rate		0.55	0.51	0.50
$\nu = 0.4999, \lambda = 1.6664e3, \mu = 0.3334$				
1e-04	1.312e-01	8.968e-02	6.300e-02	4.476e-02
rate		0.55	0.51	0.49
1e-06	1.312e-01	8.963e-02	6.284e-02	4.432e-02
rate		0.55	0.51	0.50

It follows from Table 4 that the rate of convergence for the element pair changes to 1/2 because of the boundary layer effects, which is consistent with the theoretical result. The element is still robust when the solution has strong boundary layer effects in the nearly incompressible limit.

#### REFERENCES

- [AA02a] E. Amanatidou and N. Aravas, *Mixed finite element formulations of strain-gradient elasticity problems*, Comput. Methods Appl. Mech. Engrg. **191** (2002), 1723–1751.
- [AA02b] H. Askes and E. C. Aifantis, *Numerical modeling of size effects with gradient elasticity-Formulation, meshless discretization and examples*, Int. J. Fract. **117** (2002), 347–358.
- [AA92] S. B. Altan and E. C. Aifantis, *On the structure of the mode III crack-tip in gradient elasticity*, Scripta Metal. Mater. **26** (1992), 319–324.
- [ABdVL<sup>+</sup>13] F. Auricchio, L. Beirão da Veiga, C. Lovadina, A. Reali, R. L. Taylor, and P. Wriggers, *Approximation of incompressible large deformation elastic problems: some unresolved issues*, Comput. Mech. **52** (2013), 1153–1167.
- [ABF84] D. N. Arnold, F. Brezzi, and M. Fortin, *A stable finite element for the Stokes equations*, Calcolo **21** (1984), 337–344.
- [ADN64] S. Agmon, L. Douglis, and A. Nirenberg, *Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions II*, Comm. Pure Appl. Math. **17** (1964), 35–92.
- [AF03] R. A. Adams and J. J. F. Fournier, *Sobolev Spaces*, 2nd ed., Academic Press, 2003.
- [Aif99] E. C. Aifantis, *Strain gradient interpretation of size effects*, Int. J. Fract. **95** (1999), 299–314.
- [BB03] C. Bacuta and J. H. Bramble, *Regularity estimates for solutions of the equations of linear elasticity in convex plane polygonal domains*, Z. Angew. Math. Phys. **54** (2003), 874–878.
- [BBF13] D. Boffi, F. Brezzi, and M. Fortin, *Mixed Finite Element Methods and Applications*, 2nd ed., Springer-Verlag, Berlin Heidelberg, 2013.
- [BFS65] F. A. Bogner, R. L. Fox, and L. A. Schmit, *The generation of interelement compatible stiffness and mass matrices by the use of interpolation formulas*, Proceedings of the Conferences on the Matrix Methods in Structural Mechanics, 1965.

- [BG85] C. Bernardi and R. Geneviève, *Analysis of some finite elements for the stokes problems*, Math. Comp. **44** (1985), no. 169, 71–79.
- [BGNS11] S. C. Brenner, T. Gudi, M. Neilan, and L.-Y. Sung,  *$C^0$  penalty methods for the fully nonlinear Monge-Ampère equation*, Math. Comp. **80** (2011), 1979–1995.
- [BM05] D. Braess and P. B. Ming, *A finite element method for nearly incompressible elasticity problems*, Math. Comp. **74** (2005), 25–52.
- [BN11] S. C. Brenner and M. Neilan, *A  $C^0$  interior penalty method for a fourth order elliptic singular perturbation problem*, SIAM J. Numer. Anal. **49** (2011), 869–892.
- [Bra96] D. Braess, *Stability of saddle-point problems with penalty*, RAIRO Anal. Numér. **30** (1996), 731–742.
- [BS08] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, 3rd ed., Springer Science + Buiness Media LLC, 2008.
- [BS92a] I. Babuška and M. Suri, *Locking effects in the finite element approximation of elasticity problems*, Numer. Math. **62** (1992), 439–463.
- [BS92b] S. C. Brenner and L.-Y. Sung, *Linear finite element methods for planar linear elasticity*, Math. Comp. **59** (1992), 321–338.
- [Cia78] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.
- [Clé75] Ph. Clément, *Approximation by finite element functions using local regularization*, RAIRO Anal. Numér. **9** (1975), 77–84.
- [CM10] M. Costabel and A. McIntosh, *On Bogovskii and regularized Poincaré integral operators for de Rham complexes on Lipschitz domains*, Math. Z. **265** (2010), no. 2, 297–320.
- [CR72] P. G. Ciarlet and P.-A. Raviart, *General Lagrange and Hermite interpolation in  $R^n$  with applications to finite element methods*, Arch. Rational Mech. Anal. **46** (1972), 177–199.
- [DM13] R. Danchin and P. B. Mucha, *Divergence*, Discrete Contin. Dyn. Syst. Ser. S **6** (2013), 1163–1172.
- [DX14] C. F. Dunkl and Y. Xu, *Orthogonal Polynomials of Several Variables*, 2nd ed., Encyclopedia of Mathematics and its Applications, Cambridge University Press, 2014.
- [EGH<sup>+</sup>02] G. Engel, K. Garikipati, T. J. R. Hughes, M. G. Larsson, L. Mazzei, and R. L. Taylor, *Continuous/discontinuous finite element approximations of fourth-order elliptic problems in structural and continuum mechanics with applications to thin beams and plates, and strain gradient elasticity*, Comput. Methods Appl. Mech. Engrg. **191** (2002), 3669–3750.
- [FH97] N. A. Fleck and J. W. Hutchinson, *Strain gradient plasticity*, Advances in Applied Mechanics, vol. 33, Academic Press, 1997, pp. 295–361.
- [Fis11] P. Fischer,  *$C^1$  Continuous Methods in Computational Gradient Elasticity*, 2011. Der Technischen Fakultät der Universität Erlangen-Nürnberg.
- [FKM<sup>+</sup>11] P. Fischer, M. Klassen, J. Mergheim, P. Steinmann, and R. Müller, *Isogeometric analysis of 2D gradient elasticity*, Comput. Mech. **47** (2011), 325–334.
- [FMS10] P. Fischer, J. Mergheim, and P. Steinmann, *On the  $C^1$  continuous discretization of non-linear gradient elasticity: a comparison of NEM and FEM based on Bernstein-Bézier patches*, Int. J. Numer. Meth. Engrg. **82** (2010), 1282–1307.
- [FN14] X. B. Feng and M. Neilan, *Convergence of a fourth-order singular perturbation of the  $n$ -dimensional radially symmetric Monge-Ampère equation*, Appl. Anal. **93** (2014), 1626–1646.
- [For77] M. Fortin, *An analysis of the convergence of mixed finite element methods*, RAIRO Anal. Numér. **11** (1977), 341–354.
- [Gal11] G. P. Galdi, *An Introduction to the Mathematical Theory of the Navier-Stokes Equations: Steady-State Problems*, Springer Science+Business Media, LLC, 2011.
- [GLN12] J. Guzmán, D. Leykekhman, and M. Neilan, *A family of non-conforming elements and the analysis of Nitsche’s method for a singularly perturbed fourth order problem*, Calcolo **49** (2012), 95–125.
- [Her65] L. R. Hermann, *Elasticity equations for incompressible and nearly incompressible materials by a variational theorem*, AIAA, J. **3** (1965), 1896–1900.
- [HH69] I. Hlaváček and M. Hlaváček, *On the existence and uniqueness of solution and some variational principles in linear theories of elasticity with couple-stresses. II: Mindlin’s elasticity with microstructure and the first strain-gradient theory*, Aplikace Matematiky **14** (1969), 411–427.
- [Joh72] F. John, *The transition from thin plate to membrane in the case of a plate under uniform tension*, Continuum mechanics and related problems of analysis, Izdat. “Nauka”, Moscow, 1972, pp. 193–201.
- [KN20] S. Khakalo and J. Niiranen, *Anisotropic strain gradient thermoelasticity for cellular structures: Plate models, homogenization and isogeometric analysis*, J. Mech. Phys. Solids **134** (2020), 103728.

- [Koi64] W. T. Koiter, *Couple-stresses in the theory of elasticity I*, Nederl. Akad. Wetensch. Proc. Ser. B **67** (1964), 17–29.
- [Kor08] A. Korn, *Solution générale du problème d'équilibre dans la théorie de l'élasticité dans le cas où les efforts sont donnés à la surface*, Ann. Fac. Sci. Toulouse Sci. Math. Sci. Phys. (2) **10** (1908), 165–269.
- [Kor09] ———, *Über einige ungleichungen, welche in der Theorie der elastischen und elektrischen Schwingungen eine Rolle spielen*, Bull. Intern. Cracov. Akad. Umiejetnosci (Classe Sci. Math. Nat.) (1909), 706–724.
- [LM19] Y. L. Liao and P. B. Ming, *A family of nonconforming rectangular elements for strain gradient elasticity*, Adv. Appl. Math. Mech. **11** (2019), no. 6, 1263–1286.
- [LMS17] H. L. Li, P. B. Ming, and Z.-C. Shi, *Two robust nonconforming  $H^2$ -elements for linear strain gradient elasticity*, Numer. Math. **137** (2017), 691–711.
- [LMW21] H. L. Li, P. B. Ming, and H. Y. Wang,  *$H^2$ -Korn's inequality and the nonconforming elements for the strain gradient elastic model*, J. Sci. Comput. **88** (2021), 78–100. <https://doi.org/10.1007/s10915-021-01597-7>.
- [LW17] Ph. Laurencot and Ch. Walker, *Some singular equations modeling MEMS*, Bull. Amer. Math. Soc. **54** (2017), 437–479.
- [Min64] R. D. Mindlin, *Micro-structure in linear elasticity*, Arch. Rational Mech. Anal. **16** (1964), 51–78.
- [MSW13] K.-A. Mardal, J. Schöberl, and R. Winther, *A uniformly stable Fortin operator for the Taylor-Hood element*, Numer. Math. **123** (2013), 537–551.
- [MTW02] K. A. Mardal, X.-C. Tai, and R. Winther, *A robust finite element method for Darcy-Stokes flow*, SIAM J. Numer. Anal. **40** (2002), 1605–1631.
- [NKBN16] J. Niiranen, S. Khakalo, V. Balabanov, and A. H. Niemi, *Variational formulation and isogeometric analysis for fourth-order boundary value problems of gradient-elastic bar and plane strain/stress problems*, Comput. Methods Appl. Mech. Engrg. **308** (2016), 182–211.
- [NTW01] T. K. Nilssen, X.-C. Tai, and R. Winther, *A robust nonconforming  $H^2$ -element*, Math. Comp. **70** (2001), 489–505.
- [PB15] V. Phueng and P. M. Baiz, *Mixed finite element formulations for strain-gradient elasticity problems using the FEniCS environment*, Finite Elem. Anal. Des. **96** (2015), 23–40.
- [PZV09] S.-A. Papanicolopoulos, A. Zervos, and I. Vardoulakis, *A three-dimensional  $C^1$  finite element for gradient elasticity*, Int. J. Numer. Meth. Engrg. **135** (2009), 1396–1415.
- [RVdVG14] S. Rudraraju, A. Van der Ven, and K. Garikipati, *Three-dimensional isogeometric solutions to general boundary value problems of Toupin's gradient elasticity theory at finite strains*, Comput. Methods Appl. Mech. Engrg. **278** (2014), 705–728.
- [Sch76] Z. Schuss, *Singular perturbations and the transition from thin plate to membrane*, Proc. Amer. Math. Soc. **58** (1976), 139–147.
- [Sem92] B. Semper, *Conforming finite element approximations for a fourth-order singular perturbation problem*, SIAM J. Numer. Anal. **29** (1992), 1043–1058.
- [Sem94] ———, *Locking in finite-element approximations to long thin extensible beams*, IMA J. Numer. Anal. **14** (1994), 97–109.
- [SKF99] J. Y. Shu, W. E. King, and N. A. Fleck, *Finite elements for materials with strain gradient effects*, Internat. J. Numer. Meth. Engrg., **44** (1999), 373–391.
- [STP85] J. Simo, R. L. Taylor, and K. Pister, *Variational and projection methods for the volume constraint in finite deformation elasto-plasticity*, Comput. Methods Appl. Mech. Engrg. **51** (1985), 177–208.
- [Sze75] G. Szegő, *Orthogonal Polynomials*, 4th ed., AMS, Providence, Rhode Island, 1975. AMS Coll. Publ. Vol. XXIII.
- [TH73] C. Taylor and P. Hood, *A numerical solution of the Navier–Stokes equations using the finite element technique*, Comput. & Fluids **1** (1973), 73–100.
- [Tia21] S. D. Tian, *New Nonconforming Finite Element Methods for Fourth Order Elliptic Problems*, 2021. Ph.D. Thesis, Peking University.
- [Vog83] M. Vogelius, *An analysis of the  $p$ -version of the finite element method for nearly incompressible materials uniformly valid, optimal error estimates*, Numer. Math. **41** (1983), 39–53.
- [Wei06] Y. G. Wei, *A new finite element method for strain gradient theories and applications to fracture analyses*, European Journal of Mechanics A/Solids **25** (2006), 897–913.
- [Wih06] T. P. Wihler, *Locking-free adaptive discontinuous Galerkin FEM for linear elasticity problems*, Math. Comp. **75** (2006), 1087–1102.
- [WZZ12] M. Wang, P. H. Zu, and S. Zhang, *High accuracy nonconforming finite elements for fourth order problems*, Sci. China Math. **55** (2012), no. 10, 2183–2192.

[ZMKZ12] L. Zybell, U. Mühlich, M. Kuna, and Z. L. Zhang, *A three-dimensional finite element for gradient elasticity based on mixed-type formulation*, *Comput. Mater. Sci.* **52** (2012), 268–273.

[ZPV09] A. Zervos, S.-A. Papanicolopoulos, and I. Vardoulakis, *Two finite element discretizations for gradient elasticity*, *J. Eng. Mech.-ASCE* **135** (2009), 203–213.

LSEC, INSTITUTE OF COMPUTATIONAL MATHEMATICS AND SCIENTIFIC/ENGINEERING COMPUTING, AMSS, CHINESE ACADEMY OF SCIENCES, No. 55, EAST ROAD ZHONG-GUAN-CUN, BEIJING 100190, CHINA

SCHOOL OF MATHEMATICAL SCIENCES, UNIVERSITY OF CHINESE ACADEMY OF SCIENCES, BEIJING 100049, CHINA

*Email address:* liaoyulei@lsec.cc.ac.cn, mpb@lsec.cc.ac.cn

LABORATORY OF COMPUTATIONAL PHYSICS, INSTITUTE OF APPLIED PHYSICS AND COMPUTATIONAL MATHEMATICS, BEIJING 100088, CHINA

*Email address:* xu\_yun@iapcm.ac.cn