RECONSTRUCTED DISCONTINUOUS APPROXIMATION TO STOKES EQUATION IN A SEQUENTIAL LEAST SQUARES FORMULATION*

Ruo Li

CAPT, LMAM and School of Mathematical Sciences, Peking University, Beijing 100871, China Email: rli@math.pku.edu.cn Fanyi Yang¹) School of Mathematics, Sichuan University, Chengdu 610064, China Email: yangfanyi@scu.edu.cn

Abstract

We propose a new least squares finite element method to solve the Stokes problem with two sequential steps. The approximation spaces are constructed by the patch reconstruction with one unknown per element. For the first step, we reconstruct an approximation space consisting of piecewise curl-free polynomials with zero trace. By this space, we minimize a least squares functional to obtain the numerical approximations to the gradient of the velocity and the pressure. In the second step, we minimize another least squares functional to give the solution to the velocity in the reconstructed piecewise divergence-free space. We derive error estimates for all unknowns under both L^2 norms and energy norms. Numerical results in two dimensions and three dimensions verify the convergence rates and demonstrate the great flexibility of our method.

Mathematics subject classification: 65N30.

Key words: Stokes problem, Least squares finite element method, Reconstructed discontinuous approximation, Solenoid and irrotational polynomial bases.

1. Introduction

The Stokes problem, which models a viscous and incompressible fluid flow, is a linearized version of the full Navier-Stokes equation neglecting the nonlinear convective term. Reliable and efficient numerical methods for the Stokes problem have been extensively studied in the references. Among these methods, there were many efforts devoted to develop mixed finite element methods based on the weak formulation of the Stokes problem. A key issue of classical mixed finite element methods is the choice of element types. The pair of finite element spaces are required to satisfy the stability condition, known as the inf-sup condition. We refer the readers to [10, 12, 18] for some examples in classical mixed finite element methods.

The least squares finite element methods for the Stokes problem have been developed in [5– 8,16,20,28,31]. For these methods, least squares principle together with finite element methods can offer the advantage of circumventing the inf-sup condition arising in mixed methods. Bochev and Gunzburger developed a least squares approach based on rewriting the velocity-vorticitypressure formulation as a first-order elliptic system [8]. Cai and his coworkers developed the least squares finite element method based on the L^2 norm residual and C^0 spaces for the Stokes

^{*} Received September 1, 2020 / Revised version received April 13, 2021 / Accepted April 26, 2021 / Published online February 26, 2022 /

¹⁾ Corresponding author

problem, we refer to [5, 14-16] for more details. Liu et al. developed a hybrid least squares finite element method based on continuous finite element spaces. This method attempts to combine the advantages of FOSLS and FOSLL* [28]. The works introduced above are based on conforming finite element spaces and such continuous least squares methods are general techniques in numerical methods. We refer to [9] and the references therein for an overview of least squares finite element methods. Based on discontinuous approximation, the discontinuous least squares finite element methods have also been developed for many problems including the Stokes problem, and we refer to [3, 4, 6, 7, 17, 26] for more details.

In this paper, we propose a new least squares finite element method with the reconstructed discontinuous approximation. The novelty is that we construct three specific approximation spaces which allow us to solve the Stokes problem in two sequential steps. The sequential process is motivated from the idea in [16,27] to define two least-squares-type functionals to approximate unknowns sequentially. The feasibility of this method is based on new approximation spaces which are obtained by solving local least squares problems on each element. In the first step, we reconstruct an approximation space that consists of piecewise irrotational polynomials with zero trace to approximate the gradient of the velocity. This space is an extension of the space proposed in [24], which will also be used in this step to approximate the pressure. The functions in both approximation spaces may be discontinuous across interior faces and we define a least squares functional with the weak imposition of the continuity across the interior faces to seek numerical solutions in approximation to the gradient and the pressure. In the second step, we reconstruct a piecewise divergence-free polynomial space to approximate the velocity, which is also a generalization of the space in [24]. We minimize another least squares functional, together with the numerical gradient obtained in the first step, to solve the numerical solution for the velocity. For the error estimate, we introduce a series of projection operators to derive the convergence rates for all variables with respect to L^2 norms and energy norms. We prove that the convergence orders under energy norms are all optimal and the L^2 errors for all variables can only be proved to be sub-optimal. We conduct a series of numerical examples in two dimensions and three dimensions to confirm our theoretical error estimates. In addition, we observe that the L^2 errors for all unknowns are optimally convergent for approximation spaces of odd orders. Another advantage of our method is the implementation is quite simple. The different types of the reconstruction can be implemented in a uniform way. We present the details to the computer implementation of our method in Appendix 6.

The rest of our paper is organized as follows. In Section 2, we give the notation that will be used in this paper. In Section 3, we introduce the reconstruction operators and the corresponding approximation spaces. The approximation properties of spaces are also presented in this section. In Section 4, we define two least squares functionals for sequentially solving the Stokes problem with reconstructed spaces. We also prove the error estimates under L^2 norms and energy norms. In Section 5, we present a series of numerical results in two dimensions and three dimensions to illustrate the accuracy and the flexibility of our method.

2. Preliminaries

We let Ω be a convex bounded polygonal (polyhedral) domain in $\mathbb{R}^d (d = 2, 3)$ with the boundary $\partial \Omega$. We denote by \mathcal{T}_h a set of polygonal (polyhedral) elements which partition the domain Ω . We denote by \mathcal{E}_h^i the set of interior faces of \mathcal{T}_h and by \mathcal{E}_h^b the set of the faces that are on the boundary $\partial \Omega$. Let $\mathcal{E}_h = \mathcal{E}_h^i \cup \mathcal{E}_h^b$ be the set of all d-1 dimensional faces. Further,

for any element $K \in \mathcal{T}_h$ and any face $e \in \mathcal{E}_h$ we set h_K as the diameter of the element K and h_e as the size of the face e, and the mesh size is denoted by $h = \max_{K \in \mathcal{T}_h} h_K$. It is assumed that the partition \mathcal{T}_h is shape-regular in the sense of that: there exist

- an integer N that is independent of h;
- a number $\sigma > 0$ that is independent of h;
- a compatible sub-decomposition \mathcal{T}_h into triangles (tetrahedrons);

such that

- every polygon (polyhedron) K in \mathcal{T}_h admits a decomposition $\widetilde{\mathcal{T}}_{h|K}$ which has at most N triangles (tetrahedrons);
- for any triangle (tetrahedron) $\widetilde{K} \in \widetilde{\mathcal{T}}_h$, the ratio $h_{\widetilde{K}}/\rho_{\widetilde{K}}$ is bounded by σ where $h_{\widetilde{K}}$ denotes the diameter of \widetilde{K} and $\rho_{\widetilde{K}}$ denotes the radius of the largest disk (ball) inscribed in \widetilde{K} .

The above regularity requirements have several consequences [13, 23]:

- M1 there exists a positive constant σ_v that is independent of h such that $\sigma_v h_K \leq h_e$ for any element $K \in \mathcal{T}_h$ and any face $e \subset \partial K$;
- M2 [trace inequality] there exists a constant C that is independent of h such that

$$\|v\|_{L^{2}(\partial K)}^{2} \leq C\left(h_{K}^{-1}\|v\|_{L^{2}(K)}^{2} + h_{K}\|\nabla v\|_{L^{2}(K)}^{2}\right), \quad \forall v \in H^{1}(K);$$

$$(2.1)$$

M3 [inverse inequality] there exists a constant C that is independent of h such that

$$\|\nabla v\|_{L^{2}(K)} \le Ch_{K}^{-1} \|v\|_{L^{2}(K)}, \quad \forall v \in \mathbb{P}_{k}(K),$$
(2.2)

where $\mathbb{P}_k(\cdot)$ is the polynomial space of degree less than k.

For the sub-decomposition $\widetilde{\mathcal{T}}_h$, we denote by $\widetilde{\mathcal{E}}_h$ the collection of all d-1 dimensional faces in $\widetilde{\mathcal{T}}_h$, and we decompose $\widetilde{\mathcal{E}}_h$ into $\widetilde{\mathcal{E}}_h = \widetilde{\mathcal{E}}_h^i \cup \widetilde{\mathcal{E}}_h^b$ where $\widetilde{\mathcal{E}}_h^i$ and $\widetilde{\mathcal{E}}_h^b$ are the sets of interior faces and boundary faces, respectively. From the regularity assumption, it is clear that $\mathcal{E}_h \subset \widetilde{\mathcal{E}}_h, \mathcal{E}_h^i \subset \widetilde{\mathcal{E}}_h^i$ and $\mathcal{E}_h^b \subset \widetilde{\mathcal{E}}_h^b$.

Then we introduce the trace operators that are associated with weak formulations. Let e be an interior face shared by two adjacent elements K^+ and K^- , and we let \mathbf{n}^+ and \mathbf{n}^- be the unit outward normal on e corresponding to ∂K^+ and ∂K^- , respectively. For the scalar-valued function v and the d dimensional vector-valued function q and $d \times d$ dimensional tensor-valued function τ , we define $v^+ := v|_{e \subset \partial K^+}, v^- := v|_{e \subset \partial K^-}, q^+ := q|_{e \subset \partial K^+}, q^- := q|_{e \subset \partial K^-}, \tau^+ := \tau|_{e \subset \partial K^+}, \tau^- := \tau|_{e \subset \partial K^-}$. The average operator $\{\cdot\}$ on e is defined as

$$\{v\} := \frac{1}{2} \left(v^+ + v^-\right), \quad \{q\} := \frac{1}{2} \left(q^+ + q^-\right), \quad \{\tau\} := \frac{1}{2} \left(\tau^+ + \tau^-\right).$$

The jump operator $\left[\cdot\right]$ on e is defined as

$$\begin{split} \llbracket v \rrbracket &:= v^+ \mathbf{n}^+ + v^- \mathbf{n}^-, \quad \llbracket \mathbf{n} \times \boldsymbol{q} \rrbracket &:= \mathbf{n}^+ \times \boldsymbol{q}^+ + \mathbf{n}^- \times \boldsymbol{q}^-, \\ \llbracket \mathbf{n} \otimes \boldsymbol{q} \rrbracket &:= \mathbf{n}^+ \otimes \boldsymbol{q}^+ + \mathbf{n}^- \otimes \boldsymbol{q}^-, \quad \llbracket \mathbf{n} \otimes \boldsymbol{\tau} \rrbracket &:= \mathbf{n}^+ \otimes \boldsymbol{\tau}^+ + \mathbf{n}^- \otimes \boldsymbol{\tau}^-, \end{split}$$

where \otimes denotes the tensor product between two vectors. For any tensor $\hat{\tau} \in \mathbb{R}^{d \times d}$, if \mathcal{O} is an operator on vector-valued functions, we expend \mathcal{O} to act on $\hat{\tau}$ columnwise. For example, we let $\hat{\tau} = (\hat{\tau}_1, \ldots, \hat{\tau}_d)(\hat{\tau}_i \in \mathbb{R}^d)$, and $\mathbf{n} \otimes \hat{\tau}$ is defined by

$$\mathbf{n}\otimes \widehat{oldsymbol{ au}} := \mathbf{n}\otimes (\widehat{oldsymbol{ au}}_1,\ldots,\widehat{oldsymbol{ au}}_d) = (\mathbf{n}\otimes \widehat{oldsymbol{ au}}_1,\ldots,\mathbf{n}\otimes \widehat{oldsymbol{ au}}_d)^T,$$

and $\nabla \times \hat{\tau}$, $\nabla \cdot \hat{\tau}$ are defined by

$$abla imes \widehat{oldsymbol{ au}} :=
abla imes (\widehat{oldsymbol{ au}}_1, \dots, \widehat{oldsymbol{ au}}_d) = (
abla imes \widehat{oldsymbol{ au}}_1, \dots,
abla imes \widehat{oldsymbol{ au}}_d)^T,$$

 $abla \cdot \widehat{oldsymbol{ au}} :=
abla \cdot (\widehat{oldsymbol{ au}}_1, \dots, \widehat{oldsymbol{ au}}_d) = (
abla \cdot \widehat{oldsymbol{ au}}_1, \dots,
abla \cdot \widehat{oldsymbol{ au}}_d)^T.$

For the boundary face e, the trace operators are defined as

where **n** is the unit outward normal on e.

Hereafter, we let C and C with subscripts denote the generic positive constants that may be different from line to line but independent of the mesh size h. For a bounded domain D, we will use the standard notation and definitions for the Sobolev spaces $L^2(D)$, $L^2(D)^d$, $L^2(D)^{d\times d}$, $H^r(D)$, $H^r(D)^d$, $H^r(D)^{d\times d}$ with r a positive integer, and we will also use their associated inner products, semi-norms and norms. We define the space of divergence-free functions by

$$\mathbf{S}^{r}(D) := \left\{ \boldsymbol{v} \in H^{r}(D)^{d} \mid \nabla \cdot \boldsymbol{v} = 0 \text{ in } D \right\},\$$

and we further define the space of tensor-valued functions by

$$\mathbf{I}^{r}(D) := \left\{ \boldsymbol{\tau} \in H^{r}(D)^{d \times d} \mid \nabla \times \boldsymbol{\tau} = 0, \quad \mathrm{tr}\left(\boldsymbol{\tau}\right) = 0 \text{ in } D \right\},$$

where tr (·) denotes the standard trace operator. For the partition \mathcal{T}_h , we will use the definitions for the broken Sobolev spaces $L^2(\mathcal{T}_h)$, $L^2(\mathcal{T}_h)^d$, $L^2(\mathcal{T}_h)^{d\times d}$, $H^r(\mathcal{T}_h)$, $H^r(\mathcal{T}_h)^d$, $H^r(\mathcal{T}_h)^{d\times d}$ and their corresponding broken semi-norms and norms. Moreover, for any space $X \subset L^2(\Omega)$, we let X/\mathbb{R} consist of the functions in X that have zero mean value on Ω ,

$$X/\mathbb{R} := \left\{ v \in X \mid \int_{\Omega} v \mathrm{d}\boldsymbol{x} = 0 \right\}.$$

The incompressible Stokes problem we are studying in this paper reads: seek the velocity fields u and the pressure p such that

$$-\nu\Delta \boldsymbol{u} + \nabla \boldsymbol{p} = \boldsymbol{f}, \quad \text{in } \Omega,$$

$$\nabla \cdot \boldsymbol{u} = 0, \qquad \text{in } \Omega,$$

$$\boldsymbol{u} = \boldsymbol{g}, \qquad \text{on } \partial\Omega,$$

(2.3)

where f is a give source term and g is the boundary data and ν is the reciprocal of the Reynolds number.

As being declared above, we will propose a new least squares finite element method, together with the reconstructed discontinuous approximation, to solve the problem (2.3). We introduce

a new inter-mediate variable **U** to substitute $\nabla \boldsymbol{u} = (\nabla u^1, \dots, \nabla u^d)$, thus one can reformulate the problem (2.3) into an equivalent first-order system:

$$-\nu \nabla \cdot \mathbf{U} + \nabla p = \boldsymbol{f}, \quad \text{in } \Omega,$$

$$\nabla \boldsymbol{u} - \mathbf{U} = \boldsymbol{0}, \quad \text{in } \Omega,$$

$$\nabla \cdot \boldsymbol{u} = 0, \quad \text{in } \Omega,$$

$$\boldsymbol{u} = \boldsymbol{g}, \quad \text{on } \partial \Omega.$$

$$(2.4)$$

In Section 4, we will go back to the Stokes problem and present the details of the sequential least squares method to the first-order system (2.4) with the reconstructed spaces introduced in Section 3.

3. Reconstructed Approximation Space

In this section, we define three types of reconstruction operators that will be used in numerically solving (2.4). The first one is the reconstruction operator which has been used in [23,25,26] and the other two operators are extensions from the first one. The reconstruction procedure includes two parts. The first part is to construct the element patch and this part is the same for all reconstruction operators. Now we present the details of the construction to the element patch. We begin by assigning a collocation point in each element. For each element K, we specify its barycenter \boldsymbol{x}_K as its corresponding collocation point. Then for each element K we aggregate an element patch S(K) which is a set of elements and consists of K itself and some neighbour elements. Specifically, the element patch S(K) is constructed in a recursive strategy. We first appoint a threshold value #S to control the cardinality of S(K). We let $S_0(K) = \{K\}$ and construct a sequence of element sets $S_t(K)(t \geq 1)$ recursively:

$$S_{t+1}(K) = \bigcup_{\widetilde{K} \in S_t(K)} \bigcup_{\widehat{K} \in \mathcal{N}(\widetilde{K})} \widehat{K}, \qquad t = 0, 1, 2, \dots$$

where $\mathcal{N}(\tilde{K})$ denotes the set of elements face-neighbouring to \tilde{K} . This recursion ends when the depth t satisfies that the cardinality of $S_t(K)$ is greater than #S. We compute the distances between the collocation points of all elements in $S_t(K)$ and the point \boldsymbol{x}_K . We choose the #S smallest distances and gather the corresponding elements to form the element patch S(K). By this recursive strategy, for any element K the cardinality of S(K) is always #S. After constructing the element patch, for element K we denote by $\mathcal{I}(K)$ the set of the collocation points to the elements in S(K):

$$\mathcal{I}(K) := \left\{ \boldsymbol{x}_{\widetilde{K}} \mid \forall \widetilde{K} \in S(K) \right\}.$$

Then we will define three reconstruction operators to approximate the functions in $H^r(\Omega)$, $\mathbf{S}^r(\Omega)$ and $\mathbf{I}^r(\Omega)(r \ge 2)$, respectively.

3.1. Reconstruction for Scalar-valued Functions

We denote by U_h the piecewise constant space associated with \mathcal{T}_h :

$$U_h := \left\{ v_h \in L^2(\Omega) \mid v_h \mid_K \in \mathbb{P}_0(K), \quad \forall K \in \mathcal{T}_h \right\}.$$

We will define a reconstruction operator \mathcal{R}^m from the space U_h onto a subspace of the piecewise polynomial space [23]. Given a piecewise constant function $g \in U_h$, we will seek a polynomial of degree m ($m \ge 1$) on every element. For each element $K \in \mathcal{T}_h$, we solve the following discrete least squares problem to determine a polynomial $\mathcal{R}_K^m g$ of degree m:

$$\mathcal{R}_{K}^{m}g = \operatorname*{arg\,min}_{q \in \mathbb{P}_{m}(S(K))} \sum_{\boldsymbol{x} \in \mathcal{I}(K)} |q(\boldsymbol{x}) - g(\boldsymbol{x})|^{2}$$

s.t. $q(\boldsymbol{x}_{K}) = g(\boldsymbol{x}_{K}).$ (3.1)

We follow [23, Assumption B] to make the assumption to ensure the problem (3.1) has a unique solution.

Assumption 3.1. For any element $K \in \mathcal{T}_h$ and any polynomial $q(\boldsymbol{x}) \in \mathbb{P}_m(S(K))$, we have that

$$q|_{\mathcal{I}(K)} = 0$$
 implies $q|_{S(K)} = 0.$ (3.2)

The Assumption 3.1 is actually a geometrical assumption which rules out the case that the points in $\mathcal{I}(K)$ are lying on an algebraic curve (surface) of degree m and requires the value of #S shall be greater than dim $(\mathbb{P}_m(\cdot))$.

Then the reconstruction operator \mathcal{R}^m is defined in a piecewise manner:

$$(\mathcal{R}^m g)|_K = (\mathcal{R}^m_K g)|_K$$
, for any element $K \in \mathcal{T}_h$.

We note that the polynomial $\mathcal{R}_K^m g$ is linearly dependent on g. Hence, the operator \mathcal{R}^m maps U_h onto a subspace of the piecewise polynomial space, which is denoted by $U_h^m = \mathcal{R}^m U_h$. For any smooth function $q \in C^0(\Omega)$, we define $\tilde{q} \in U_h$ such that

$$\widetilde{q}(\boldsymbol{x}_K) = q(\boldsymbol{x}_K), \quad \forall K \in \mathcal{T}_h,$$

and we extend the operator \mathcal{R}^m to act on the smooth function by defining $\mathcal{R}^m q := \mathcal{R}^m \widetilde{q}$.

In addition, we outline a group of basis functions of U_h^m . For any element K, we define the characteristic function $w_K(\mathbf{x}) \in U_h$ as

$$w_K(\boldsymbol{x}) = \begin{cases} 1, & \boldsymbol{x} \in K, \\ 0, & \text{otherwise,} \end{cases}$$

and we denote λ_K as $\lambda_K = \mathcal{R}^m w_K$.

Lemma 3.1. The functions $\{\lambda_K\}$ ($\forall K \in \mathcal{T}_h$) are linearly independent.

Proof. For any $\lambda_K(\boldsymbol{x})$, the constraint in (3.1) implies that

$$\lambda_K(\boldsymbol{x}_{\widetilde{K}}) = \begin{cases} 1, & \widetilde{K} = K, \\ 0, & \widetilde{K} \neq K. \end{cases}$$
(3.3)

We assume that there exist coefficients $\{a_K\}$ ($\forall K \in \mathcal{T}_h$) such that

$$\sum_{K \in \mathcal{T}_h} a_K \lambda_K(\boldsymbol{x}) = 0, \qquad \forall \boldsymbol{x} \in \Omega.$$
(3.4)

For any element K, let $\boldsymbol{x} = \boldsymbol{x}_K$ in (3.4) and by (3.3) one see that $a_K = 0$, which shows that $\{\lambda_K\}$ are linearly independent. This completes the proof.

Clearly, we have that dim $(\{\lambda_K\}) = \dim(U_h)$. Lemma 3.1 in fact gives that the reconstructed space satisfies dim $(U_h^m) = \dim(U_h)$ and U_h^m is spanned by $\{\lambda_K\}$. Moreover, for the function $g \in U_h$ or $g \in C^0(\Omega)$, one may write $\mathcal{R}^m g$ explicitly:

$$\mathcal{R}^{m}g = \sum_{K \in \mathcal{T}_{h}} g(\boldsymbol{x}_{K})\lambda_{K}(\boldsymbol{x}).$$
(3.5)

Then we give the approximation property of the space U_h^m . For any element K, we define a constant $\Lambda(m, K)$ as

$$\Lambda(m,K) := \max_{q \in \mathbb{P}_m(S(K))} \frac{\max_{\boldsymbol{x} \in S(K)} |q(\boldsymbol{x})|}{\max_{\boldsymbol{x} \in \mathcal{I}(K)} |q(\boldsymbol{x})|}.$$

We let $\Lambda_m := \max_{K \in \mathcal{T}_h} (1 + \Lambda(m, K)(\#S)^{1/2})$ and in [23, 24] the authors proved that under some mild conditions on element patches, the constant Λ_m can be bounded uniformly with respect to the partition. These conditions rely on the size of element patches and the authors also proved that if the number #S is greater than a certain number, the conditions on element patches will be fulfilled, see [23, Lemma 6] and [24, Lemma 3.4]. This certain number is usually too large to be impractical and is not recommended in the implementation. The numerical results demonstrate that our method still has a very good performance even we take the value #S to be far less than that certain number. In Section 5 we list the values of #S for different m that are adopted in the numerical tests.

Then we state the following estimate.

Lemma 3.2. For any element K and any function $g \in C^0(\Omega)$, there holds

$$\|g - \mathcal{R}^m g\|_{L^{\infty}(K)} \leq \Lambda_m \inf_{q \in \mathbb{P}_m(S(K))} \|g - q\|_{L^{\infty}(S(K))}.$$

Proof. We refer to [23, Lemma 3] for the proof.

From Lemma 3.2, it is easy to derive the following approximation properties.

Theorem 3.1. For any element K, there exist constants C such that

$$\begin{aligned} \|g - \mathcal{R}^m g\|_{H^q(K)} &\leq C \Lambda_m h_K^{m+1-q} \|g\|_{H^{m+1}(S(K))}, \qquad q = 0, 1, \\ \|\nabla^q (g - \mathcal{R}^m g)\|_{L^2(\partial K)} &\leq C \Lambda_m h_K^{m+1-q-1/2} \|g\|_{H^{m+1}(S(K))}, \qquad q = 0, 1, \end{aligned}$$
(3.6)

for any $g \in H^{m+1}(\Omega)$.

Proof. We refer to [23, Lemma 4] for the proof.

3.2. Reconstruction for Vector-valued Functions

Here we consider to extend the reconstruction process for vector-valued functions. Precisely, we will introduce a reconstruction operator for functions in the space $\mathbf{S}^{m+1}(\Omega)$. We also start from the piecewise constant space $(U_h)^d$. Given a function $\boldsymbol{g} \in (U_h)^d$ and for any element

 $K \in \mathcal{T}_h$, we solve a polynomial $\widetilde{\mathcal{R}}_K^m \boldsymbol{g}$ of degree m on S(K) by the following discrete least squares problem:

$$\widetilde{\mathcal{R}}_{K}^{m}\boldsymbol{g} = \underset{\boldsymbol{q}\in\mathbb{P}_{m}(S(K))^{d}}{\operatorname{arg\,min}} \sum_{\boldsymbol{x}\in\mathcal{I}(K)} \|\boldsymbol{q}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{x})\|_{l^{d}}^{2},$$
s.t.
$$\begin{cases} \boldsymbol{q}(\boldsymbol{x}_{K}) = \boldsymbol{g}(\boldsymbol{x}_{K}), \\ \nabla \cdot \boldsymbol{q} = 0, \end{cases}$$
(3.7)

where

$$\|\boldsymbol{v}\|_{l^d}^2 := v_1^2 + \ldots + v_d^2, \qquad \forall \boldsymbol{v} = (v_1, \ldots, v_d)^T \in \mathbb{R}^d.$$

Based on Assumption 3.1, it is trivial to check the uniqueness and the existence of the solution to the problem (3.7). Then the reconstruction operator $\widetilde{\mathcal{R}}^m$ is piecewise defined as

$$(\widetilde{\mathcal{R}}^m \boldsymbol{g})|_K = (\widetilde{\mathcal{R}}_K^m \boldsymbol{g})|_K$$
, for any element $K \in \mathcal{T}_h$

It should be noted that $\widetilde{\mathcal{R}}_{K}^{m} \boldsymbol{g}$ still has a linear dependence on \boldsymbol{g} . Therefore, the operator $\widetilde{\mathcal{R}}^{m}$ maps the space $(U_{h})^{d}$ into the piecewise divergence-free polynomial space of degree m, and we denote by $\mathbf{S}_{h}^{m} = \widetilde{\mathcal{R}}^{m}(U_{h})^{d}$. We also extend the operator $\widetilde{\mathcal{R}}^{m}$ to act on the smooth function as the operator \mathcal{R}^{m} . For the function $\boldsymbol{g}(\boldsymbol{x}) \in C^{0}(\Omega)^{d}$, we define a piecewise constant function $\widetilde{\boldsymbol{g}}(\boldsymbol{x}) \in (U_{h})^{d}$ such that

$$\widetilde{\boldsymbol{g}}(\boldsymbol{x}_K) = \boldsymbol{g}(\boldsymbol{x}_K), \quad \forall K \in \mathcal{T}_h,$$

and we define $\widetilde{\mathcal{R}}^m \boldsymbol{g} := \widetilde{\mathcal{R}}^m \widetilde{\boldsymbol{g}}$. We also present a group of basis functions to the space \mathbf{S}_h^m . For any element K, we define an indicator function $\widetilde{\boldsymbol{w}}_K^i(\boldsymbol{x}) \in (U_h)^d$, which reads

$$\widetilde{\boldsymbol{w}}_{K}^{i}(\boldsymbol{x}) = egin{cases} \boldsymbol{e}_{i}, & \boldsymbol{x} \in K, \ 0, & ext{otherwise}, \end{cases}$$

where \boldsymbol{e}_i is a $d \times 1$ unit vector whose *i*-th entry is 1. Then we define $\widetilde{\boldsymbol{\lambda}}_K^i$ as $\widetilde{\boldsymbol{\lambda}}_K^i = \widetilde{\mathcal{R}}_K^m \widetilde{\boldsymbol{w}}_K^i$ and we have the following lemma.

Lemma 3.3. The functions $\{\lambda_K^i\}$ ($\forall K \in \mathcal{T}_h, 1 \leq i \leq d$) are linearly independent.

Proof. The proof results from the constraint in (3.7) and is similar to the proof of Lemma 3.1. $\hfill \Box$

Analogously, we conclude that the space \mathbf{S}_h^m is spanned by $\{\widetilde{\boldsymbol{\lambda}}_K^i\}$ and for the function $\boldsymbol{g} = (g^1, \ldots, g^d) \in (U_h)^d$ or $\boldsymbol{g} = (g^1, \ldots, g^d) \in C^0(\Omega)^d$, we can write $\widetilde{\mathcal{R}}^m \boldsymbol{g}$ as

$$\widetilde{\mathcal{R}}^m \boldsymbol{g} = \sum_{K \in \mathcal{T}_h} \sum_{i=1}^d g^i(\boldsymbol{x}_K) \widetilde{\boldsymbol{\lambda}}_K^i(\boldsymbol{x}).$$

Further, we give the approximation property of the operator $\widetilde{\mathcal{R}}^m$.

Lemma 3.4. For any element K and any function $\boldsymbol{g} \in C^0(\Omega)^d$, there holds

$$\|\boldsymbol{g} - \widetilde{\mathcal{R}}^{m} \boldsymbol{g}\|_{L^{\infty}(K)} \leq \sqrt{d} \Lambda_{m} \inf_{\boldsymbol{q} \in \mathbb{P}_{m}(S(K))^{d} \cap \mathbf{S}^{0}(S(K))} \|\boldsymbol{g} - \boldsymbol{q}\|_{L^{\infty}(S(K))}.$$
(3.8)

Proof. For any divergence-free polynomial $\boldsymbol{q} \in \mathbb{P}_m(S(K))^d \cap \mathbf{S}^0(S(K))$, we clearly have that $\widetilde{\mathcal{R}}_K^m \boldsymbol{q} = \boldsymbol{q}$ from the definition of the least squares problem (3.7). We deduce that

$$\begin{aligned} \|\boldsymbol{g} - \widetilde{\mathcal{R}}^{m} \boldsymbol{g}\|_{L^{\infty}(K)} &\leq \|\boldsymbol{g} - \boldsymbol{q}\|_{L^{\infty}(K)} + \|\widetilde{\mathcal{R}}^{m}(\boldsymbol{q} - \boldsymbol{g})\|_{L^{\infty}(K)} \\ &\leq \|\boldsymbol{g} - \boldsymbol{q}\|_{L^{\infty}(K)} + \Lambda(m, K) \max_{\boldsymbol{x} \in \mathcal{I}(K)} |\widetilde{\mathcal{R}}^{m}(\boldsymbol{q} - \boldsymbol{g})| \\ &\leq \|\boldsymbol{g} - \boldsymbol{q}\|_{L^{\infty}(K)} + \Lambda(m, K) \sqrt{d\#S} \max_{\boldsymbol{x} \in \mathcal{I}(K)} |\boldsymbol{q} - \boldsymbol{g}| \\ &\leq \sqrt{d} \left(1 + \Lambda(m, K) \sqrt{\#S}\right) \|\boldsymbol{g} - \boldsymbol{q}\|_{L^{\infty}(S(K))}, \end{aligned}$$

which gives us the estimate (3.8) and completes the proof.

Theorem 3.2. For any element K, there exist constants C such that

$$\|\boldsymbol{g} - \mathcal{\tilde{R}}^{m} \boldsymbol{g}\|_{H^{q}(K)} \leq C\Lambda_{m} h_{K}^{m+1-q} \|\boldsymbol{g}\|_{H^{m+1}(S(K))}, \qquad q = 0, 1, \\ \|\nabla^{q} (\boldsymbol{g} - \mathcal{\tilde{R}}^{m} \boldsymbol{g})\|_{L^{2}(\partial K)} \leq C\Lambda_{m} h_{K}^{m+1-q-1/2} \|\boldsymbol{g}\|_{H^{m+1}(S(K))}, \qquad q = 0, 1,$$
(3.9)

for any $\boldsymbol{g} \in \mathbf{S}^{m+1}(\Omega)$.

Proof. By [2, Theorem 4.1] and [23, Assumption A], there exists an approximation polynomial $\tilde{q} \in \mathbb{P}_m(S(K))^d \cap \mathbf{S}^0(S(K))$ such that

$$\|\boldsymbol{g} - \widetilde{\boldsymbol{q}}\|_{L^{\infty}(S(K))} \le Ch_{K}^{m+1-d/2} \|\boldsymbol{g}\|_{H^{m+1}(S(K))}$$

By Lemma 3.4, we obtain that

$$\begin{aligned} \|\boldsymbol{g} - \widetilde{\mathcal{R}}^{m}\boldsymbol{g}\|_{L^{2}(K)} &\leq Ch_{K}^{d/2}\|\boldsymbol{g} - \widetilde{\mathcal{R}}^{m}\boldsymbol{g}\|_{L^{\infty}(K)} \leq Ch_{K}^{d/2}\Lambda_{m}\|\boldsymbol{g} - \widetilde{\boldsymbol{q}}\|_{L^{\infty}(S(K))} \\ &\leq C\Lambda_{m}h_{K}^{m+1}\|\boldsymbol{g}\|_{H^{m+1}(S(K))}, \end{aligned}$$

and together with the inverse inequality M2, we have that

$$\begin{aligned} \|\boldsymbol{g} - \mathcal{R}^{m}\boldsymbol{g}\|_{H^{1}(K)} &\leq \|\boldsymbol{g} - \widetilde{\boldsymbol{q}}\|_{H^{1}(K)} + \|\widetilde{\boldsymbol{q}} - \mathcal{R}^{m}\boldsymbol{g}\|_{H^{1}(K)} \\ &\leq \|\boldsymbol{g} - \widetilde{\boldsymbol{q}}\|_{H^{1}(K)} + Ch_{K}^{-1}\|\widetilde{\boldsymbol{q}} - \widetilde{\mathcal{R}}^{m}\boldsymbol{g}\|_{L^{2}(K)} \\ &\leq C\Lambda_{m}h_{K}^{m+1}\|\boldsymbol{g}\|_{H^{m+1}(S(K))}. \end{aligned}$$

Applying the trace inequality M2 gives the trace estimate in (3.9).

3.3. Reconstruction for Tensor-valued Functions

In this subsection, we consider the reconstruction for the tensor-valued functions in the space $\mathbf{I}^{m+1}(\Omega)$. Again, we start from a piecewise constant space. Since the functions in $\mathbf{I}^{m+1}(\Omega)$ have zero trace, we define the space \mathbf{U}_h consisting of piecewise constant functions with zero trace, which reads

$$\mathbf{U}_h := \left\{ \boldsymbol{v}_h \in (U_h)^{d \times d} \mid \operatorname{tr} (\boldsymbol{v}_h) = 0 \right\}.$$

For the function $\boldsymbol{g} \in \mathbf{U}_h$, we will seek a polynomial $\widehat{\mathcal{R}}_K^m \boldsymbol{g}$ of degree m on S(K) by solving the following problem:

$$\widehat{\mathcal{R}}_{K}^{m}\boldsymbol{g} = \operatorname*{arg\,min}_{\boldsymbol{q}\in\mathbb{P}_{m}(S(K))^{d\times d}} \sum_{\boldsymbol{x}\in\mathcal{I}(K)} \|\boldsymbol{q}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{x})\|_{l^{d\times d}}^{2},$$
s.t.
$$\begin{cases} \boldsymbol{q}(\boldsymbol{x}_{K}) = \boldsymbol{g}(\boldsymbol{x}_{K}), \\ \nabla \times \boldsymbol{q} = 0, \\ \operatorname{tr}(\boldsymbol{q}) = 0, \end{cases}$$
(3.10)

where

$$\|oldsymbol{ au}\|_{l^d imes l^d}^2 := \|oldsymbol{ au}_1\|_{l^d}^2 + \ldots + \|oldsymbol{ au}_d\|_{l^d}^2, \qquad orall oldsymbol{ au} = (oldsymbol{ au}_1, \ldots, oldsymbol{ au}_d) \in \mathbb{R}^{d imes d}.$$

Similarly, we have that the problem (3.10) has a unique solution by Assumption 3.1. The global reconstruction operator $\widehat{\mathcal{R}}^m$ is piecewise defined by

$$(\widehat{\mathcal{R}}^m \boldsymbol{g})|_K = (\widehat{\mathcal{R}}^m_K \boldsymbol{g})|_K$$
, for any element $K \in \mathcal{T}_h$

The solution $\widehat{\mathcal{R}}_K^m g$ is linearly dependent on g and we denote by \mathbf{I}_h^m the image of the operator $\widehat{\mathcal{R}}^m$. Then we still extend the operator $\widehat{\mathcal{R}}^m$ to act on the smooth functions. For the function $g(\mathbf{x}) \in C^0(\Omega)^{d \times d}$ with zero trace, we let $\widehat{g}(\mathbf{x}) \in \mathbf{U}_h$ such that

$$\widehat{\boldsymbol{g}}(\boldsymbol{x}_K) = \boldsymbol{g}(\boldsymbol{x}_K), \quad \forall K \in \mathcal{T}_h,$$

and again we define $\widehat{\mathcal{R}}^m \boldsymbol{g} := \widehat{\mathcal{R}}^m \widehat{\boldsymbol{g}}$. Here we give a group of basis functions to the space \mathbf{I}_h^m . We define a group of characteristic functions $\widehat{\boldsymbol{w}}_K^{i,j}(\boldsymbol{x}) \in \mathbf{U}_h$ but we shall consider the zero trace condition of functions in \mathbf{U}_h . Actually this condition implies that there are only $d^2 - 1$ characteristic functions on every element. For any element K, we define $\widehat{\boldsymbol{w}}_K^{i,j}(\boldsymbol{x})$ as

$$\widehat{\boldsymbol{w}}_{K}^{i,j}(\boldsymbol{x}) = \begin{cases} \boldsymbol{e}_{i,j}, & \boldsymbol{x} \in K, \\ 0, & \text{otherwise,} \end{cases} \quad 1 \leq i \neq j \leq d, \end{cases}$$

where $e_{i,j}$ is the $d \times d$ matrix whose (i, j) entry is 1 and the other entries are 0. For $1 \le i < d$, we define $\widehat{\boldsymbol{w}}_{K}^{i,i}(\boldsymbol{x})$ as

$$\widehat{oldsymbol{w}}_{K}^{i,i}(oldsymbol{x}) = egin{cases} \widehat{oldsymbol{e}}_{i,i}, & oldsymbol{x} \in K, \ 0, & ext{otherwise}, \end{cases}$$

where $\hat{e}_{i,i}$ is the $d \times d$ matrix whose (i,i) entry is 1, (d,d) entry is -1 and other entries are 0.

We define $\widehat{\lambda}_{K}^{i,j} = \widehat{\mathcal{R}}^{m} \widehat{\boldsymbol{w}}_{K}^{i,j}$ and we state that the functions $\{\widehat{\boldsymbol{\lambda}}_{K}^{i,j}\}$ are a group of basis functions to the space \mathbf{I}_{h}^{m} .

Lemma 3.5. The functions $\{\widehat{\lambda}_{K}^{i,j}\}$ ($\forall K \in \mathcal{T}_{h}, 1 \leq i, j \leq d, i+j < 2d$) are linearly independent.

Proof. The proof of Lemma 3.5 is analogous to the proof of Lemmas 3.3 and 3.1.

Clearly, we can know that $\mathbf{I}_{h}^{m} = \operatorname{span}\{\widehat{\lambda}_{K}^{i,j}\}\$ and for any function $\boldsymbol{g} = (g^{i,j}(\boldsymbol{x}))_{d \times d} \in \mathbf{U}_{h}$ or $\boldsymbol{g} = (g^{i,j}(\boldsymbol{x}))_{d \times d} \in C^{0}(\Omega)^{d \times d}$ with tr $(\boldsymbol{g}) = 0$, $\widehat{\mathcal{R}}^{m}\boldsymbol{g}$ can be expressed as

$$\widehat{\mathcal{R}}^{m}\boldsymbol{g} = \sum_{K \in \mathcal{T}_{h}} \sum_{1 \leq i, j \leq d \text{ and } i+j < 2d} g^{i,j}(\boldsymbol{x}_{K}) \widehat{\boldsymbol{\lambda}}_{K}^{i,j}(\boldsymbol{x}).$$
(3.11)

Moreover, we give the approximation property of the space \mathbf{I}_{h}^{m} .

Lemma 3.6. For any element K and any function $\boldsymbol{g} \in C^0(\Omega)^{d \times d}$ with tr $(\boldsymbol{g}) = 0$, there holds

$$\|\boldsymbol{g} - \widehat{\mathcal{R}}^{m}\boldsymbol{g}\|_{L^{\infty}(K)} \leq d\Lambda_{m} \inf_{\boldsymbol{q} \in \mathbb{P}_{m}(S(K))^{d \times d} \cap \mathbf{I}^{0}(S(K))} \|\boldsymbol{g} - \boldsymbol{q}\|_{L^{\infty}(S(K))}.$$
(3.12)

Proof. According to the problem (3.1), we have that $\widehat{\mathcal{R}}^m q = q$ for any $q \in \mathbb{P}_m(S(K))^{d \times d} \cap \mathbf{I}^0(S(K))$. We obtain that

$$\begin{split} \|\boldsymbol{g} - \widehat{\mathcal{R}}^{m}\boldsymbol{g}\|_{L^{\infty}(K)} &\leq \|\boldsymbol{g} - \boldsymbol{q}\|_{L^{\infty}(K)} + \|\widehat{\mathcal{R}}^{m}(\boldsymbol{q} - \boldsymbol{g})\|_{L^{\infty}(K)} \\ &\leq \|\boldsymbol{g} - \boldsymbol{q}\|_{L^{\infty}(K)} + \Lambda(m, K) \max_{\boldsymbol{x} \in \mathcal{I}(K)} |\widetilde{\mathcal{R}}^{m}(\boldsymbol{q} - \boldsymbol{g})| \\ &\leq \|\boldsymbol{g} - \boldsymbol{q}\|_{L^{\infty}(K)} + d\Lambda(m, K) \sqrt{\#S} \max_{\boldsymbol{x} \in \mathcal{I}(K)} |\boldsymbol{q} - \boldsymbol{g}| \\ &\leq d \left(1 + \Lambda(m, K) \sqrt{\#S} \right) \|\boldsymbol{g} - \boldsymbol{q}\|_{L^{\infty}(S(K))}, \end{split}$$

which completes the proof.

Theorem 3.3. For any element K, there exist constants C such that

$$\|\boldsymbol{g} - \widehat{\mathcal{R}}^{m} \boldsymbol{g}\|_{H^{q}(K)} \leq C\Lambda_{m} h_{K}^{m+1-q} \|\boldsymbol{g}\|_{H^{m+1}(S(K))}, \qquad q = 0, 1, \|\nabla^{q} (\boldsymbol{g} - \widehat{\mathcal{R}}^{m} \boldsymbol{g}\|_{L^{2}(\partial K)} \leq C\Lambda_{m} h_{K}^{m+1-q-1/2} \|\boldsymbol{g}\|_{H^{m+1}(S(K))}, \qquad q = 0, 1,$$
(3.13)

for any $\boldsymbol{g} \in \mathbf{I}^{m+1}(\Omega)$.

Proof. Since $\nabla \times \boldsymbol{g} = 0$ and tr $(\boldsymbol{g}) = 0$, there exists a function $\hat{\boldsymbol{g}} \in \mathbf{S}^{m+2}(S(K))$ such that $\boldsymbol{g} = \nabla \hat{\boldsymbol{g}}$ [18, Lemma 2.1]. By [2, Theorem 4.1], there exists a polynomial $\hat{\boldsymbol{q}} \in \mathbb{P}_{m+1}(S(K))^d \cap \mathbf{S}^0(S(K))$ such that

$$\|\boldsymbol{g} - \nabla \widehat{\boldsymbol{q}}\|_{L^{\infty}(S(K))} = \|\nabla (\widehat{\boldsymbol{g}} - \widehat{\boldsymbol{q}})\|_{L^{\infty}(S(K))} \le Ch_{K}^{m+1-d/2} \|\widehat{\boldsymbol{g}}\|_{H^{m+2}(S(K))}$$
$$= Ch_{K}^{m+1-d/2} \|\boldsymbol{g}\|_{H^{m+1}(S(K))}.$$

Clearly, $\nabla \widehat{q} \in \mathbb{P}_m(S(K))^{d \times d} \cap \mathbf{I}^0(S(K))$. By Lemma 3.6 and the approximation estimate of \widehat{q} , we deduce that

$$\begin{aligned} \|\boldsymbol{g} - \widehat{\mathcal{R}}^{m} \boldsymbol{g}\|_{L^{2}(K)} &\leq C h_{K}^{d/2} \|\boldsymbol{g} - \widehat{\mathcal{R}}^{m} \boldsymbol{g}\|_{L^{\infty}(K)} \leq C h_{K}^{d/2} \Lambda_{m} \|\boldsymbol{g} - \nabla \widehat{\boldsymbol{q}}\|_{L^{\infty}(S(K))} \\ &\leq C \Lambda_{m} h_{K}^{m+1} \|\boldsymbol{g}\|_{H^{m+1}(S(K))}. \end{aligned}$$

Together with the inverse inequality M3, we have

$$\begin{aligned} \|\boldsymbol{g} - \mathcal{R}^{m}\boldsymbol{g}\|_{H^{1}(K)} &\leq \|\boldsymbol{g} - \nabla \widehat{\boldsymbol{q}}\|_{H^{1}(K)} + \|\nabla \widehat{\boldsymbol{q}} - \mathcal{R}^{m}\boldsymbol{g}\|_{H^{1}(K)} \\ &\leq \|\boldsymbol{g} - \nabla \widehat{\boldsymbol{q}}\|_{H^{1}(K)} + Ch_{K}^{-1}\|\nabla \widehat{\boldsymbol{q}} - \widehat{\mathcal{R}}^{m}\boldsymbol{g}\|_{L^{2}(K)} \\ &\leq \|\boldsymbol{g} - \nabla \widehat{\boldsymbol{q}}\|_{H^{1}(K)} + Ch_{K}^{-1}\left(\|\boldsymbol{g} - \nabla \widehat{\boldsymbol{q}}\|_{L^{2}(K)} + \|\boldsymbol{g} - \widehat{\mathcal{R}}^{m}\boldsymbol{g}\|_{L^{2}(K)}\right) \\ &\leq C\Lambda_{m}h_{K}^{m}\|\boldsymbol{g}\|_{H^{m+1}(S(K))}.\end{aligned}$$

The trace estimate of (3.13) follows from the trace inequality M2.

We have established three types of reconstruction operators and their corresponding approximation spaces and the approximation results. In Appendix 6, we present some details of the computer implementation to reconstructed spaces.

Remark 3.1. The computational cost of constructing the approximation spaces mainly has two parts. The first part is the cost to construct element patches. In numerical tests, we choose the accuracy m = 1, 2, 3 and in this case at most 6 recursive steps on each element are enough

49

to construct the element patch. The computational cost of this step is very low, see Appendix 6. For the second step, as we illustrate in Appendix 6, we only need to store three matrices with sizes $\#S \times \#S$, $d\#S \times d\#S$ and $d^2\#S \times d^2\#S$, respectively. This part is also very cheap. We note that in our implementation, the computational time to construct approximation spaces is only 5% of the computational time to use the numerical integration to assemble the total stiff matrix.

4. Sequential Least Squares Method for Stokes Problem

In this section, we propose a sequential least squares finite element method to solve the Stokes problem based on the first-order system (2.4). We are motivated by the idea in [16,27] to decouple the system (2.4) into two steps. The first first-order system is defined to seek the numerical approximations to the gradient **U** and the pressure p, which reads

$$-\nu\nabla\cdot\mathbf{U} + \nabla p = \boldsymbol{f}, \qquad \text{in } \Omega, \\ \mathbf{n} \times \mathbf{U} = \mathbf{n} \times \nabla \boldsymbol{g}, \qquad \text{on } \partial\Omega.$$
 (4.1)

The first equation in (4.1) is the same as the first equation in (2.4) and the boundary condition in (2.4) provides the tangential trace $\mathbf{n} \times \mathbf{U}$ on the boundary $\partial \Omega$. Then, we define a least squares functional $J_h^{\mathbf{p}}(\cdot, \cdot)$ for numerically solving the system (4.1), which reads

$$J_{h}^{\mathbf{p}}(\mathbf{V}_{h},q_{h}) := \sum_{K\in\mathcal{T}_{h}} \|-\nu\nabla\cdot\mathbf{V}_{h}+\nabla q_{h}-\boldsymbol{f}\|_{L^{2}(K)}^{2} + \sum_{e\in\mathcal{E}_{h}^{i}} \left(\frac{\eta}{h_{e}}\|[q_{h}]\|_{L^{2}(e)}^{2} + \frac{\eta}{h_{e}}\|[\mathbf{n}\otimes\mathbf{V}_{h}]\|_{L^{2}(e)}^{2}\right) + \sum_{e\in\mathcal{E}_{h}^{b}} \frac{\eta}{h_{e}}\|\mathbf{n}\times\mathbf{V}_{h}-\mathbf{n}\times\nabla\boldsymbol{g}\|_{L^{2}(e)}^{2}, \quad \forall (\mathbf{V}_{h},q_{h})\in H^{1}(\mathcal{T}_{h})^{d\times d}\times H^{1}(\mathcal{T}_{h}), \quad (4.2)$$

where η is a positive parameter and will be specified later on. In (4.2), the trace terms are used to weakly impose the continuity condition as well as the boundary condition. We minimize the functional (4.2) over the spaces $\mathbf{I}_{h}^{m} \times \widetilde{U}_{h}^{m}$ to give approximations to \mathbf{U} and p and here the space \widetilde{U}_{h}^{m} is defined by U_{h}^{m}/\mathbb{R} .

Specifically, the minimization problem is defined as to find $\mathbf{U}_h \in \mathbf{I}_h^m$ and $p_h \in \widetilde{U}_h^m$ such that

$$(\mathbf{U}_h, p_h) = \underset{(\mathbf{V}_h, q_h) \in \mathbf{I}_h^m \times \widetilde{U}_h^m}{\arg\min} J_h^{\mathbf{p}}(\mathbf{V}_h, q_h).$$
(4.3)

We write the Euler-Lagrange equation to solve the problem (4.3) and the corresponding discrete variational problem reads: find $(\mathbf{U}_h, p_h) \in \mathbf{I}_h^m \times \widetilde{U}_h^m$ such that

$$a_{h}^{\mathbf{p}}(\mathbf{U}_{h}, p_{h}; \mathbf{V}_{h}, q_{h}) = l_{h}^{\mathbf{p}}(\mathbf{V}_{h}, q_{h}), \quad \forall (\mathbf{V}_{h}, q_{h}) \in \mathbf{I}_{h}^{m} \times \widetilde{U}_{h}^{m},$$
(4.4)

where the bilinear form $a_h^{\mathbf{p}}(\cdot; \cdot)$ is

$$a_{h}^{\mathbf{p}}(\mathbf{U}_{h}, p_{h}; \mathbf{V}_{h}, q_{h}) = \sum_{K \in \mathcal{T}_{h}} \int_{K} (-\nu \nabla \cdot \mathbf{U}_{h} + \nabla p_{h}) (-\nu \nabla \cdot \mathbf{V}_{h} + \nabla q_{h}) d\mathbf{x}$$
$$+ \sum_{e \in \mathcal{E}_{h}^{i}} \int_{e} \frac{\eta}{h_{e}} [p_{h}] \cdot [q_{h}] d\mathbf{s} + \sum_{e \in \mathcal{E}_{h}^{i}} \int_{e} \frac{\eta}{h_{e}} [\mathbf{n} \otimes \mathbf{U}_{h}] : [\mathbf{n} \otimes \mathbf{V}_{h}] d\mathbf{s}$$
$$+ \sum_{e \in \mathcal{E}_{h}^{b}} \int_{e} \frac{\eta}{h_{e}} (\mathbf{n} \times \mathbf{U}_{h}) \cdot (\mathbf{n} \times \mathbf{V}_{h}) d\mathbf{s},$$
(4.5)

and the linear form $l_h^{\mathbf{p}}(\cdot)$ is

$$l_h^{\mathbf{p}}(\mathbf{V}_h, q_h) = \sum_{K \in \mathcal{T}_h} \int_K \boldsymbol{f} \left(-\nu \nabla \cdot \mathbf{V}_h + \nabla q_h \right) \mathrm{d}\boldsymbol{x} + \sum_{e \in \mathcal{E}_h^b} \int_e \frac{\eta}{h_e} (\mathbf{n} \times \mathbf{V}_h) \cdot (\mathbf{n} \times \nabla \boldsymbol{g}) \mathrm{d}\boldsymbol{s}.$$

Then we will focus on the error estimate to the problem (4.4). To do this, we will require some projection results which play a key role in the analysis. We define V_h^m and \tilde{V}_h^m as piecewise polynomial spaces with respect to the partition \mathcal{T}_h and the sub-decomposition $\tilde{\mathcal{T}}_h$,

$$V_h^m := \{ v_h \in L^2(\Omega) \mid v_h|_K \in \mathbb{P}_m(K), \quad \forall K \in \mathcal{T}_h \}, \widetilde{V}_h^m := \{ v_h \in L^2(\Omega) \mid v_h|_{\widetilde{K}} \in \mathbb{P}_m(\widetilde{K}), \quad \forall \widetilde{K} \in \widetilde{\mathcal{T}}_h \},$$

and clearly we have that $V_h^m \subset \widetilde{V}_h^m$. Then we state following lemmas.

Lemma 4.1. For any $v_h \in V_h^m$, there exists a function $\widetilde{v}_h \in \widetilde{V}_h^m$ such that

$$v_{h} = \widetilde{v}_{h}, \qquad \text{in any } K \in \mathcal{T}_{h},$$

$$[\widetilde{v}_{h}] = 0, \qquad \text{on any } \widetilde{e} \in \widetilde{\mathcal{E}}_{h} \setminus \mathcal{E}_{h},$$

$$\sum_{\widetilde{e} \in w(e)} h_{\widetilde{e}}^{\beta} \|[\widetilde{v}_{h}]\|_{L^{2}(\widetilde{e})}^{2} \leq C h_{e}^{\beta} \|[v_{h}]\|_{L^{2}(e)}^{2}, \quad \text{on any } e \in \mathcal{E}_{h} \text{ and } \beta = -1, 1,$$

$$(4.6)$$

where $w(e) = \{ \widetilde{e} \in \widetilde{\mathcal{E}}_h \mid \widetilde{e} \subset e \}.$

Proof. The fact $V_h^m \subset \widetilde{V}_h^m$ directly implies that there exists a polynomial $\widetilde{v}_h \in \widetilde{V}_h^m$ satisfying the equalities in (4.6) and

$$\sum_{\widetilde{e} \in w(e)} \| [\![\widetilde{v}_h]\!] \|_{L^2(\widetilde{e})}^2 = \| [\![v_h]\!] \|_{L^2(e)}^2$$

Hence,

$$\sum_{\widetilde{e}\in w(e)} h_{\widetilde{e}}^{\beta} \| [\widetilde{v}_{h}] \|_{L^{2}(\widetilde{e})}^{2} = \sum_{\widetilde{e}\in w(e)} h_{e}^{\beta} \left(\frac{h_{\widetilde{e}}}{h_{e}}\right)^{\beta} \| [\widetilde{v}_{h}] \|_{L^{2}(\widetilde{e})}^{2}$$
$$\leq Ch_{e}^{\beta} \sum_{\widetilde{e}\in w(e)} \| [\widetilde{v}_{h}] \|_{L^{2}(\widetilde{e})}^{2} = Ch_{e}^{\beta} \| [v_{h}] \|_{L^{2}(e)}^{2},$$

where the last inequality follows from the mesh regularity assumption. This completes the proof. $\hfill \Box$

Lemma 4.2. For any
$$v_h \in V_h^m$$
, there exists a function $\widetilde{v}_h \in \widetilde{V}_h^m \cap H^1(\Omega)$ such that

$$\sum_{K \in \mathcal{T}_h} \|\nabla^{\alpha}(v_h - \widetilde{v}_h)\|_{L^2(K)}^2 \leq C \sum_{e \in \mathcal{E}_h^i} h_e^{1-2\alpha} \|[v_h]\|_{L^2(e)}^2, \quad \alpha = 0, 1.$$
(4.7)

Proof. By Lemma 4.1, there exists a piecewise polynomial $\hat{v}_h \in \widetilde{V}_h^m$ satisfying the estimate (4.6). By [21, Theorem 2.1], there exists a function $\tilde{v}_h \in \widetilde{V}_h^m \cap H^1(\Omega)$ such that

$$\sum_{\widetilde{K}\in\widetilde{\mathcal{T}}_h} \|\nabla^{\alpha}(\widehat{v}_h-\widetilde{v}_h)\|_{L^2(\widetilde{K})}^2 \leq C \sum_{\widetilde{e}\in\widetilde{\mathcal{E}}_h^i} h_{\widetilde{e}}^{1-2\alpha} \|[\widehat{v}_h]\|_{L^2(\widetilde{e})}^2, \quad \alpha=0,1$$

Combining (4.6), we have that

$$\sum_{K\in\mathcal{T}_h} \|\nabla^{\alpha}(v_h - \widetilde{v}_h)\|_{L^2(K)}^2 = \sum_{\widetilde{K}\in\widetilde{\mathcal{T}}_h} \|\nabla^{\alpha}(\widehat{v}_h - \widetilde{v}_h)\|_{L^2(\widetilde{K})}^2 \le C \sum_{\widetilde{e}\in\widetilde{\mathcal{E}}_h^i} h_{\widetilde{e}}^{1-2\alpha} \|[\widehat{v}_h]\|_{L^2(\widetilde{e})}^2$$
$$\le C \sum_{e\in\mathcal{E}_h^i} h_e^{1-2\alpha} \|[v_h]\|_{L^2(e)}^2,$$

which gives the inequality (4.7) and this completes the proof.

Lemma 4.3. For any $\boldsymbol{q}_h \in (V_h^m)^d$, there exists a function $\widetilde{\boldsymbol{q}}_h \in (\widetilde{V}_h^m)^d \cap H^1(\Omega)^d$ such that

$$\sum_{\widetilde{K}\in\widetilde{\mathcal{T}}_{h}} \|\nabla^{\alpha}(\boldsymbol{q}_{h}-\widetilde{\boldsymbol{q}}_{h})\|_{L^{2}(\widetilde{K})}^{2}$$

$$\leq C\left(\sum_{\widetilde{e}\in\widetilde{\mathcal{E}}_{h}^{i}} h_{\widetilde{e}}^{1-2\alpha} \|[\widetilde{\mathbf{n}}\otimes\boldsymbol{q}_{h}]\|_{L^{2}(\widetilde{e})}^{2} + \sum_{\widetilde{e}\in\widetilde{\mathcal{E}}_{h}^{b}} h_{\widetilde{e}}^{1-2\alpha} \|\widetilde{\mathbf{n}}\times\boldsymbol{q}_{h}\|_{L^{2}(\widetilde{e})}^{2}\right), \quad \alpha = 0, 1,$$

$$(4.8)$$

and the tangential trace $\mathbf{n} \times \widetilde{\mathbf{q}}_h$ vanishes on the boundary $\partial \Omega$.

Proof. Again by [21, Theorem 2.1], there exists a piecewise polynomial $\hat{q}_h \in (\widetilde{V}_h^m)^d \cap H^1(\Omega)^d$ such that

$$\sum_{\widetilde{K}\in\widetilde{\mathcal{T}}_{h}} \|\nabla^{\alpha}(\boldsymbol{q}_{h}-\widehat{\boldsymbol{q}}_{h})\|_{L^{2}(\widetilde{K})}^{2} \leq C \sum_{\widetilde{e}\in\widetilde{\mathcal{E}}_{h}^{i}} h_{\widetilde{e}}^{1-2\alpha} \|[\widetilde{\mathbf{n}}\otimes\boldsymbol{q}_{h}]\|_{L^{2}(\widetilde{e})}^{2}.$$
(4.9)

We will construct a new piecewise polynomial $\tilde{\boldsymbol{q}}_h \in (\tilde{V}_h^m)^d \cap H^1(\Omega)^d$ based on $\hat{\boldsymbol{q}}_h$, which satisfies the inequality (4.8) and its tangential trace vanishes on the boundary.

We denote by $\mathcal{N} = \{\boldsymbol{\nu}_0, \boldsymbol{\nu}_1, \dots, \boldsymbol{\nu}_n\}$ the Lagrange points with respect to the partition \mathcal{T}_h and we let $\{\phi_{\boldsymbol{\nu}_0}, \phi_{\boldsymbol{\nu}_1}, \dots, \phi_{\boldsymbol{\nu}_n}\}$ be the corresponding basis functions such that $\phi_{\boldsymbol{\nu}_i}(\boldsymbol{\nu}_j) = \delta_{ij}$. Then we divide the set \mathcal{N} into three disjoint subsets (see Fig. 4.1):

$$\mathcal{N}_{i} := \{ \boldsymbol{\nu} \in \mathcal{N} \mid \boldsymbol{\nu} \text{ is interior to the domain } \Omega \},$$

$$\mathcal{N}_{v} := \{ \boldsymbol{\nu} \in \mathcal{N} \mid \boldsymbol{\nu} \text{ is shared by two different sides of the boundary } \partial \Omega \},$$

$$\mathcal{N}_{b} := \mathcal{N} \setminus (\mathcal{N}_{i} \cup \mathcal{N}_{v}).$$

(4.10)

We note that the points in \mathcal{N}_b are interior to one slide of the boundary $\partial\Omega$, and particularly in two dimensions the points in \mathcal{N}_v are vertices of the boundary $\partial\Omega$.



Fig. 4.1. Examples of Lagrange nodes in two dimensions (left) / in three dimensions (right).

By $\{\phi_{\boldsymbol{\nu}_i}\}$, the function $\widehat{\boldsymbol{q}}_h = (\widehat{q}_h^1, \dots, \widehat{q}_h^d)$ can be expanded as $\widehat{q}_h^j = \sum_{\boldsymbol{\nu} \in \mathcal{N}} \alpha_{\boldsymbol{\nu}}^j \phi_{\boldsymbol{\nu}} (1 \le j \le d)$.

Then we construct a new group of coefficients $\{\beta_{\boldsymbol{\nu}}^j\}$ by

$$\beta_{\boldsymbol{\nu}}^{j} := \begin{cases} \alpha_{\boldsymbol{\nu}}^{j}, & \text{for } \boldsymbol{\nu} \in \mathcal{N}_{i}, \\ \widetilde{\beta}_{\boldsymbol{\nu}}^{j}, & \text{for } \boldsymbol{\nu} \in \mathcal{N}_{b}, \\ 0, & \text{for } \boldsymbol{\nu} \in \mathcal{N}_{v}. \end{cases}$$
(4.11)

For $\boldsymbol{\nu} \in \mathcal{N}$, we denote $\boldsymbol{\alpha}_{\boldsymbol{\nu}}$ and $\boldsymbol{\beta}_{\boldsymbol{\nu}}$ as $\boldsymbol{\alpha}_{\boldsymbol{\nu}} = (\alpha_{\boldsymbol{\nu}}^1, \dots, \alpha_{\boldsymbol{\nu}}^d)^T$ and $\boldsymbol{\beta}_{\boldsymbol{\nu}} = (\beta_{\boldsymbol{\nu}}^1, \dots, \beta_{\boldsymbol{\nu}}^d)^T$, respectively. Then we determine the values of $\tilde{\beta}_{\boldsymbol{\nu}}^j$. By the definition (4.10), for any $\boldsymbol{\nu} \in \mathcal{N}_b$ there exists a boundary face $\tilde{e} \in \tilde{\mathcal{E}}_h^b$ that includes the point $\boldsymbol{\nu}$, and we let its corresponding coefficients satisfy that

$$\widetilde{\mathbf{n}} \times \boldsymbol{\beta}_{\boldsymbol{\nu}} = 0, \quad \widetilde{\mathbf{n}} \cdot \boldsymbol{\beta}_{\boldsymbol{\nu}} = \widetilde{\mathbf{n}} \cdot \boldsymbol{\alpha}_{\boldsymbol{\nu}},$$

where $\widetilde{\mathbf{n}}$ is the unit outward normal with respect to the boundary face \widetilde{e} . We construct a new piecewise polynomial $\widetilde{\mathbf{q}}_h = (\widetilde{q}_h^1, \ldots, \widetilde{q}_h^d)^T$ where $\widetilde{q}_h^j = \sum_{\boldsymbol{\nu} \in \mathcal{N}} \beta_{\boldsymbol{\nu}}^j \phi_{\boldsymbol{\nu}} (1 \leq j \leq d)$. It is trivial to check that the trace $\mathbf{n} \times \widetilde{\mathbf{q}}_h$ vanishes on the boundary $\partial\Omega$. Then we will estimate the term $\|\nabla(\widehat{\mathbf{q}}_h - \widetilde{\mathbf{q}}_h)\|_{L^2(\Omega)}$. Since $\widehat{\mathbf{q}}_h$ and $\widetilde{\mathbf{q}}_h$ have the same value on the points in \mathcal{N}_i , one can see that

$$\|\nabla^{\alpha}(\widehat{\boldsymbol{q}}_{h}-\widetilde{\boldsymbol{q}}_{h})\|_{L^{2}(\Omega)}^{2} \leq C \sum_{\boldsymbol{\nu}\in\mathcal{N}_{b}\cup\mathcal{N}_{v}} |\boldsymbol{\alpha}_{\boldsymbol{\nu}}-\boldsymbol{\beta}_{\boldsymbol{\nu}}|^{2} \|\nabla^{\alpha}\phi_{\boldsymbol{\nu}}\|_{L^{2}(\Omega)}^{2}.$$

We first consider the points in the set \mathcal{N}_b . Again for any $\boldsymbol{\nu} \in \mathcal{N}_b$, we have that there exists a face $\tilde{e} \in \tilde{\mathcal{E}}_h^b$ such that $\boldsymbol{\nu} \in \tilde{e}$, and by the scaling argument and the shape regularity of the partition $\tilde{\mathcal{T}}_h$, there holds $\|\nabla^{\alpha}\phi_{\boldsymbol{\nu}}\|_{L^2(\Omega)}^2 \leq Ch_{\tilde{e}}^{d-2\alpha}$. Combining with (4.11) and the inverse estimate, we obtain that

$$\begin{split} &\sum_{\boldsymbol{\nu}\in\mathcal{N}_{b}}|\boldsymbol{\alpha}_{\boldsymbol{\nu}}-\boldsymbol{\beta}_{\boldsymbol{\nu}}|^{2}\|\nabla^{\alpha}\phi_{\boldsymbol{\nu}}\|_{L^{2}(\Omega)}^{2}\\ &=\sum_{\boldsymbol{\nu}\in\mathcal{N}_{b}}\|\nabla^{\alpha}\phi_{\boldsymbol{\nu}}\|_{L^{2}(\Omega)}^{2}\left(|\mathbf{\widetilde{n}}\times(\boldsymbol{\alpha}_{\boldsymbol{\nu}}-\boldsymbol{\beta}_{\boldsymbol{\nu}})|^{2}+|\mathbf{\widetilde{n}}\cdot(\boldsymbol{\alpha}_{\boldsymbol{\nu}}-\boldsymbol{\beta}_{\boldsymbol{\nu}})|^{2}\right)\\ &\leq C\sum_{\boldsymbol{\nu}\in\mathcal{N}_{b}}h_{\overline{e}}^{d-2\alpha}|\mathbf{\widetilde{n}}\times\boldsymbol{\alpha}_{\boldsymbol{\nu}}|^{2}=C\sum_{\boldsymbol{\nu}\in\mathcal{N}_{b}}h_{\overline{e}}^{d-2\alpha}|\mathbf{\widetilde{n}}\times\widehat{\boldsymbol{q}}_{h}(\boldsymbol{\nu})|^{2}\\ &\leq C\sum_{\boldsymbol{\nu}\in\mathcal{N}_{b}}h_{\overline{e}}^{d-2\alpha}\|\mathbf{\widetilde{n}}\times\widehat{\boldsymbol{q}}_{h}\|_{L^{\infty}(\widetilde{e})}^{2}\leq C\sum_{\boldsymbol{\nu}\in\mathcal{N}_{b}}h_{\overline{e}}^{1-2\alpha}\|\mathbf{\widetilde{n}}\times\widehat{\boldsymbol{q}}_{h}\|_{L^{2}(\widetilde{e})}^{2}.\end{split}$$

Then we move on to the points in \mathcal{N}_v . By definition (4.10), for every $\boldsymbol{\nu} \in \mathcal{N}_v$ there exist two adjacent faces $\tilde{e}_1 \in \tilde{\mathcal{E}}_h^b$ and $\tilde{e}_2 \in \tilde{\mathcal{E}}_h^b$ such that $\boldsymbol{\nu} \in \tilde{e}_1 \cap \tilde{e}_2$. We note that \tilde{e}_1 and \tilde{e}_2 are not parallel and are included in two different sides of $\partial\Omega$. We let $\tilde{\mathbf{n}}_1$ and $\tilde{\mathbf{n}}_2$ be the unit outward normals corresponding to \tilde{e}_1 and \tilde{e}_2 , respectively, and clearly we have $\tilde{\mathbf{n}}_1 \neq \tilde{\mathbf{n}}_2$. This fact implies that there exists a positive constant C such that

$$|\boldsymbol{v}|^2 \le C\left(|\widetilde{\mathbf{n}}_1 \times \boldsymbol{v}|^2 + |\widetilde{\mathbf{n}}_2 \times \boldsymbol{v}|^2\right), \quad \text{for } \forall \boldsymbol{v} \in \mathbb{R}^d.$$
(4.12)

It should be noted that the constant C only depends on the angle of $\tilde{\mathbf{n}}_1$ and $\tilde{\mathbf{n}}_2$ and this angle

only depends on the boundary $\partial \Omega$. Then we derive that

$$\begin{split} &\sum_{\boldsymbol{\nu}\in\mathcal{N}_{v}}|\boldsymbol{\alpha}_{\boldsymbol{\nu}}-\boldsymbol{\beta}_{\boldsymbol{\nu}}|^{2}\|\nabla^{\alpha}\phi_{\boldsymbol{\nu}}\|_{L^{2}(\Omega)}^{2}=\sum_{\boldsymbol{\nu}\in\mathcal{N}_{v}}\|\nabla^{\alpha}\phi_{\boldsymbol{\nu}}\|_{L^{2}(\Omega)}^{2}|\boldsymbol{\alpha}_{\boldsymbol{\nu}}|^{2}\\ &\leq C\sum_{\boldsymbol{\nu}\in\mathcal{N}_{v}}\|\nabla^{\alpha}\phi_{\boldsymbol{\nu}}\|_{L^{2}(\Omega)}^{2}\left(|\mathbf{\widetilde{n}}_{1}\times\boldsymbol{\alpha}_{\boldsymbol{\nu}}|^{2}+|\mathbf{\widetilde{n}}_{2}\times\boldsymbol{\alpha}_{\boldsymbol{\nu}}|^{2}\right)\\ &\leq C\sum_{\boldsymbol{\nu}\in\mathcal{N}_{v}}\left(h_{\tilde{e}_{1}}^{d-2\alpha}|\mathbf{\widetilde{n}}_{1}\times\hat{\boldsymbol{q}}_{h}(\boldsymbol{\nu})|^{2}+h_{\tilde{e}_{2}}^{d-2\alpha}|\mathbf{\widetilde{n}}_{2}\times\hat{\boldsymbol{q}}_{h}(\boldsymbol{\nu})|^{2}\right)\\ &\leq C\sum_{\boldsymbol{\nu}\in\mathcal{N}_{v}}\left(h_{e_{1}}^{d-2\alpha}\|\mathbf{\widetilde{n}}_{1}\times\hat{\boldsymbol{q}}_{h}\|_{L^{\infty}(\tilde{e}_{1})}^{2}+h_{\tilde{e}_{2}}^{d-2\alpha}\|\mathbf{\widetilde{n}}_{2}\times\hat{\boldsymbol{q}}_{h}\|_{L^{\infty}(\tilde{e}_{2})}^{2}\right)\\ &\leq C\sum_{\boldsymbol{\nu}\in\mathcal{N}_{v}}\left(h_{\tilde{e}_{1}}^{1-2\alpha}\|\mathbf{\widetilde{n}}_{1}\times\hat{\boldsymbol{q}}_{h}\|_{L^{2}(\tilde{e}_{1})}^{2}+h_{\tilde{e}_{2}}^{1-2\alpha}\|\mathbf{\widetilde{n}}_{1}\times\hat{\boldsymbol{q}}_{h}\|_{L^{2}(\tilde{e}_{2})}^{2}\right)\\ &\leq C\sum_{\tilde{e}\in\tilde{\mathcal{E}}_{h}^{\tilde{b}}}h^{1-2\alpha}\|\mathbf{\widetilde{n}}\times\hat{\boldsymbol{q}}_{h}\|_{L^{2}(\tilde{e})}^{2}. \end{split}$$

Thus, we arrive at

$$\|\nabla^{\alpha}(\widehat{\boldsymbol{q}}_{h}-\widetilde{\boldsymbol{q}}_{h})\|_{L^{2}(\Omega)}^{2} \leq C \sum_{e \in \widetilde{\mathcal{E}}_{h}^{b}} h_{e}^{1-2\alpha} \|\mathbf{n} \times \widehat{\boldsymbol{q}}_{h}\|_{L^{2}(e)}^{2}.$$

We finally present that the error $\|\nabla^{\alpha}(\boldsymbol{q}_{h} - \widetilde{\boldsymbol{q}}_{h})\|_{L^{2}(\widetilde{\mathcal{T}}_{h})}$ satisfies the estimate (4.8). We have that

$$\begin{split} \|\nabla^{\alpha}(\boldsymbol{q}_{h}-\widetilde{\boldsymbol{q}}_{h})\|_{L^{2}(\widetilde{\mathcal{T}}_{h})}^{2} &\leq C\left(\|\nabla^{\alpha}(\boldsymbol{q}_{h}-\widehat{\boldsymbol{q}}_{h})\|_{L^{2}(\widetilde{\mathcal{T}}_{h})}^{2} + \|\nabla^{\alpha}(\widetilde{\boldsymbol{q}}_{h}-\widehat{\boldsymbol{q}}_{h})\|_{L^{2}(\Omega)}^{2}\right) \\ &\leq C\left(\sum_{\widetilde{e}\in\widetilde{\mathcal{E}}_{h}^{i}}h_{\widetilde{e}}^{1-2\alpha}\|[\mathbf{n}\otimes\boldsymbol{q}_{h}]\|_{L^{2}(\widetilde{e})}^{2} + \sum_{\widetilde{e}\in\widetilde{\mathcal{E}}_{h}^{b}}h_{\widetilde{e}}^{1-2\alpha}\|\mathbf{n}\times\widehat{\boldsymbol{q}}_{h}\|_{L^{2}(\widetilde{e})}^{2}\right), \end{split}$$

and together with [21, Theorem 2.1] and the trace inequality, we deduce that

$$\begin{split} \sum_{e\in\tilde{\mathcal{E}}_{h}^{b}} h_{\tilde{e}}^{1-2\alpha} \|\widetilde{\mathbf{n}}\times\widehat{\mathbf{q}}_{h}\|_{L^{2}(\tilde{e})}^{2} &\leq C \sum_{e\in\tilde{\mathcal{E}}_{h}^{b}} \left(h_{\tilde{e}}^{1-2\alpha}\|\widetilde{\mathbf{n}}\times\mathbf{q}_{h}\|_{L^{2}(\tilde{e})}^{2} + h_{\tilde{e}}^{1-2\alpha}\|\widetilde{\mathbf{n}}\times(\mathbf{q}_{h}-\widehat{\mathbf{q}}_{h})\|_{L^{2}(\tilde{e})}^{2}\right) \\ &\leq C \left(\sum_{\tilde{e}\in\tilde{\mathcal{E}}_{h}^{b}} h_{\tilde{e}}^{1-2\alpha}\|\widetilde{\mathbf{n}}\times\mathbf{q}_{h}\|_{L^{2}(\tilde{e})}^{2} + \sum_{K\in\tilde{\mathcal{T}}_{h}} h_{K}^{-2\alpha}\|\mathbf{q}_{h}-\widehat{\mathbf{q}}_{h}\|_{L^{2}(K)}^{2}\right) \\ &\leq C \left(\sum_{\tilde{e}\in\tilde{\mathcal{E}}_{h}^{i}} h_{\tilde{e}}^{1-2\alpha}\|\widetilde{\mathbf{n}}\otimes\mathbf{q}_{h}\|_{L^{2}(\tilde{e})}^{2} + \sum_{\tilde{e}\in\tilde{\mathcal{E}}_{h}^{b}} h_{\tilde{e}}^{1-2\alpha}\|\widetilde{\mathbf{n}}\times\mathbf{q}_{h}\|_{L^{2}(\tilde{e})}^{2}\right). \end{split}$$

Combining all inequalities yields the estimate (4.8) and completes the proof.

Lemma 4.4. For any $\mathbf{V}_h \in (V_h^m)^{d \times d}$, there exists a function $\widetilde{\mathbf{V}}_h \in (\widetilde{V}_h^m)^{d \times d} \cap H^1(\Omega)^{d \times d}$ such that

$$\sum_{K\in\mathcal{T}_{h}} \|\nabla^{\alpha}(\mathbf{V}_{h} - \widetilde{\mathbf{V}}_{h})\|_{L^{2}(K)}^{2}$$

$$\leq C\left(\sum_{e\in\mathcal{E}_{h}^{i}} h_{e}^{1-2\alpha} \|[\mathbf{n}\otimes\mathbf{V}_{h}]\|_{L^{2}(e)}^{2} + \sum_{e\in\mathcal{E}_{h}^{b}} h_{e}^{1-2\alpha} \|\mathbf{n}\times\mathbf{V}_{h}\|_{L^{2}(e)}^{2}\right), \qquad (4.13)$$

for $\alpha = 0, 1$, and the tangential trace $\mathbf{n} \times \widetilde{\mathbf{V}}_h$ vanishes on the boundary $\partial \Omega$.

Proof. We columnwise expand \mathbf{V}_h as $\mathbf{V}_h = (\boldsymbol{v}_h^1, \dots, \boldsymbol{v}_h^d)$ where $\boldsymbol{v}_h^i \in (V_h^m)^d (1 \le i \le d)$. By Lemma 4.1, there exist a piecewise polynomial function $\widehat{\boldsymbol{v}}_h^i \in (\widetilde{V}_h^m)^d$ such that

$$\begin{aligned} \boldsymbol{v}_{h}^{i} &= \widehat{\boldsymbol{v}}_{h}^{i}, & \text{in any } \widetilde{K} \in \widetilde{\mathcal{T}}_{h}, \\ [\widetilde{\mathbf{n}} \otimes \widehat{\boldsymbol{v}}_{h}^{i}] &= 0, & \text{on any } \widetilde{e} \in \widetilde{\mathcal{E}}_{h} \setminus \mathcal{E}_{h}, \\ \sum_{\widetilde{e} \in w(e)} h_{\widetilde{e}}^{1-2\alpha} \| [\widetilde{\mathbf{n}} \otimes \widehat{\boldsymbol{v}}_{h}^{i}] \|_{L^{2}(\widetilde{e})}^{2} \leq C h_{e}^{1-2\alpha} \| [\mathbf{n} \otimes \boldsymbol{v}_{h}^{i}] \|_{L^{2}(e)}^{2}, & \text{on any } e \in \mathcal{E}_{h}^{i}, \\ \sum_{\widetilde{e} \in w(e)} h_{\widetilde{e}}^{1-2\alpha} \| [\widetilde{\mathbf{n}} \times \widehat{\boldsymbol{v}}_{h}^{i}] \|_{L^{2}(\widetilde{e})}^{2} \leq C h_{e}^{1-2\alpha} \| [\mathbf{n} \times \boldsymbol{v}_{h}^{i}] \|_{L^{2}(e)}^{2}, & \text{on any } e \in \mathcal{E}_{h}^{b}. \end{aligned}$$

$$(4.14)$$

By Lemma 4.3, for every $\hat{\boldsymbol{v}}_h^i$ there exists a piecewise polynomial function $\tilde{\boldsymbol{v}}_h^i \in (\tilde{V}_h^m)^d \cap H^1(\Omega)^d$ satisfying the estimate (4.8) and the tangential trace of $\tilde{\boldsymbol{v}}_h^i$ equals to 0 on the boundary. We define $\tilde{\mathbf{V}}_h$ as $\tilde{\mathbf{V}}_h = (\tilde{\boldsymbol{v}}_h^1, \dots, \tilde{\boldsymbol{v}}_h^d)$. By (4.14) and the estimate (4.8), it can be seen that for the polynomial $\tilde{\mathbf{V}}_h$ the estimate (4.13) holds true and its tangential trace vanishes on $\partial\Omega$. This completes the proof.

Next, we focus on the continuity and the coercivity of the bilinear form $a_h^{\mathbf{p}}(\cdot; \cdot)$. We begin by defining the following two energy norms $\|\cdot\|_{\mathbf{U}}$ and $\|\cdot\|_{\mathbf{p}}$:

$$\|\mathbf{V}_{h}\|_{\mathbf{U}}^{2} := \sum_{K \in \mathcal{T}_{h}} \|\nabla \cdot \mathbf{V}_{h}\|_{L^{2}(K)}^{2} + \sum_{e \in \mathcal{E}_{h}^{i}} h_{e}^{-1} \|[\mathbf{n} \otimes \mathbf{V}_{h}]\|_{L^{2}(e)}^{2} + \sum_{e \in \mathcal{E}_{h}^{b}} h_{e}^{-1} \|\mathbf{n} \times \mathbf{V}_{h}\|_{L^{2}(e)}^{2},$$

for any $\mathbf{V}_h \in \mathbf{I}_h^m + \mathbf{I}^1(\Omega)$, and

$$\|q_h\|_{\mathbf{p}}^2 := \sum_{K \in \mathcal{T}_h} \|\nabla q_h\|_{L^2(K)}^2 + \sum_{e \in \mathcal{E}_h^i} h_e^{-1} \|[q_h]\|_{L^2(e)},$$

and any $q_h \in \widetilde{U}_h^m + H^1(\Omega)/\mathbb{R}$. We have the following estimates which show that $\|\cdot\|_{\mathbf{U}}$ and $\|\cdot\|_{\mathbf{p}}$ are actually norms on their corresponding spaces.

Lemma 4.5. There exist a constant C such that

$$\|\mathbf{V}_h\|_{L^2(\Omega)} \le C \|\mathbf{V}_h\|_{\mathbf{U}}, \quad \forall \mathbf{V}_h \in \mathbf{I}_h^m + \mathbf{I}^1(\Omega).$$
(4.15)

Proof. we refer to [27, Lemma 4.1] for the proof.

Lemma 4.6. There exist a constant C such that

$$||q_h||_{L^2(\Omega)} \le C ||q_h||_{\mathbf{p}}, \quad \forall q_h \in U + H^1(\Omega)/R.$$
 (4.16)

Proof. We refer to [11] for the proof.

Now we are ready to state that the bilinear form $a_h^{\mathbf{p}}(\cdot; \cdot)$ is bounded and coercive with respect to energy norms $\|\cdot\|_{\mathbf{U}}$ and $\|\cdot\|_{\mathbf{p}}$ for any positive η .

Lemma 4.7. For the bilinear form $a_h^{\mathbf{p}}(\cdot; \cdot)$ with any $\eta > 0$, there exists a positive constant C such that

$$|a_h^{\mathbf{p}}(\mathbf{U}_h, p_h; \mathbf{V}_h, q_h)| \le C \left(\|\mathbf{U}_h\|_{\mathbf{U}}^2 + \|p_h\|_{\mathbf{p}}^2 \right)^{1/2} \left(\|\mathbf{V}_h\|_{\mathbf{U}}^2 + \|q_h\|_{\mathbf{p}}^2 \right)^{1/2},$$
(4.17)

for any $\mathbf{U}_h, \mathbf{V}_h \in \mathbf{I}_h^m + \mathbf{I}^1(\Omega)$ and any $p_h, q_h \in \widetilde{U}_h^m + H^1(\Omega)/\mathbb{R}$.

Proof. Without loss of generality, we set $\nu = 1$. Obviously we have that

$$\int_{K} \| - \nabla \cdot \mathbf{U}_{h} + \nabla p_{h} \|^{2} \mathrm{d}\boldsymbol{x} \leq C \left(\int_{K} \| \nabla \cdot \mathbf{U}_{h} \|^{2} \mathrm{d}\boldsymbol{x} + \int_{K} \| \nabla p_{h} \|^{2} \mathrm{d}\boldsymbol{x} \right), \quad \forall K \in \mathcal{T}_{h},$$

and applying the Cauchy-Schwarz inequality to (4.5) directly gives the estimate (4.17), which completes the proof. $\hfill \Box$

Lemma 4.8. For the bilinear form $a_h^{\mathbf{p}}(\cdot; \cdot)$ with any $\eta > 0$, there exists a positive constant C such that

$$a_{h}^{\mathbf{p}}(\mathbf{U}_{h}, p_{h}; \mathbf{U}_{h}, p_{h}) \ge C\left(\|\mathbf{U}_{h}\|_{\mathbf{U}}^{2} + \|p_{h}\|_{\mathbf{p}}^{2}\right), \qquad (4.18)$$

for any $\mathbf{U}_h \in \mathbf{I}_h^m$ and any $p_h \in \widetilde{U}_h^m$.

Proof. It is also sufficient to prove for the case $\nu = 1$. For any $\mathbf{U}_h \in \mathbf{I}_h^m$, Lemma 4.4 implies that there exists a function $\mathbf{V}_h \in H^1(\Omega)^{d \times d}$ such that $\mathbf{n} \times \mathbf{V}_h = 0$ on $\partial\Omega$ and the estimate (4.13) holds. For any $p_h \in \widetilde{U}_h^m$, there exists a function $q_h \in H^1(\Omega)$ satisfying the estimate (4.7) by Lemma 4.2.

Here we prove for the three-dimensional case. Since $\mathbf{n} \times \mathbf{V}_h = 0$ on $\partial \Omega$ and the domain Ω is assumed to be a bounded convex polygon (polyhedron), we have the following Helmholtz decomposition [16]:

$$\mathbf{V}_h = \nabla \boldsymbol{q}^T + \nabla \times \boldsymbol{\Psi},$$

where $\boldsymbol{q} \in H^1_0(\Omega)^d \cap H^2(\Omega)^d$ is the solution of

$$\Delta \boldsymbol{q} = \nabla \cdot \mathbf{V}_h, \quad \text{in } \Omega, \quad \boldsymbol{q} = 0, \quad \text{on } \partial \Omega.$$
(4.19)

Since $q_h \in H^1(\Omega)$, the regularity of generalized Stokes problem [16,22] provides that

$$\|\Delta \boldsymbol{q}\|_{L^{2}(\Omega)}^{2} + \|\nabla q_{h}\|_{L^{2}(\Omega)}^{2} \leq C\left(\|-\Delta \boldsymbol{q} + \nabla q_{h}\|_{L^{2}(\Omega)}^{2} + \|\nabla \cdot \boldsymbol{q}\|_{L^{2}(\Omega)}^{2}\right).$$
(4.20)

Together with [16, Lemma 3.2] and the auxiliary problem (4.19), we obtain

$$\|\nabla \cdot \mathbf{V}_{h}\|_{L^{2}(\Omega)}^{2} + \|\nabla q_{h}\|_{L^{2}(\Omega)}^{2}$$

$$\leq C \left(\|-\nabla \cdot \mathbf{V}_{h} + \nabla q_{h}\|_{L^{2}(\Omega)}^{2} + \|\operatorname{tr}(\mathbf{V}_{h})\|_{H^{1}(\Omega)}^{2} + \|\nabla \times \mathbf{V}_{h}\|_{L^{2}(\Omega)}^{2}\right).$$
 (4.21)

We note that the inequality (4.21) also holds in two dimensions and the proof is similar.

Then we are ready to establish the coercivity (4.18) and we first take the parameter $\eta = 1$. We have that

$$\begin{aligned} \|\mathbf{U}_{h}\|_{\mathbf{U}}^{2} + \|p_{h}\|_{\mathbf{P}}^{2} &= \sum_{K \in \mathcal{T}_{h}} \|\nabla \cdot \mathbf{U}_{h}\|_{L^{2}(K)}^{2} + \sum_{K \in \mathcal{T}_{h}} \|\nabla p_{h}\|_{L^{2}(K)}^{2} \\ &+ \sum_{e \in \mathcal{E}_{h}^{i}} h_{e}^{-1} \|[\mathbf{n} \otimes \mathbf{U}_{h}]\|_{L^{2}(e)}^{2} + \sum_{e \in \mathcal{E}_{h}^{b}} h_{e}^{-1} \|\mathbf{n} \times \mathbf{V}_{h}\|_{L^{2}(e)}^{2} + \sum_{e \in \mathcal{E}_{h}^{i}} h_{e}^{-1} \|[q_{h}]\|_{L^{2}(e)}, \end{aligned}$$

and

$$\begin{aligned} \|\nabla \cdot \mathbf{U}_{h}\|_{L^{2}(\mathcal{T}_{h})}^{2} + \|\nabla p_{h}\|_{L^{2}(\mathcal{T}_{h})}^{2} \leq C \left(\|\nabla \cdot \mathbf{V}_{h}\|_{L^{2}(\Omega)}^{2} + \|\nabla q_{h}\|_{L^{2}(\Omega)}^{2}\right) \\ + C \left(\|\nabla \cdot (\mathbf{U}_{h} - \mathbf{V}_{h})\|_{L^{2}(\mathcal{T}_{h})}^{2} + \|\nabla (p_{h} - q_{h})\|_{L^{2}(\mathcal{T}_{h})}^{2}\right). \end{aligned}$$

From (4.21) and the above two inequalities, we arrive at

$$\|\mathbf{U}_{h}\|_{\mathbf{U}}^{2} + \|p_{h}\|_{\mathbf{p}}^{2} \le C \left(a_{h}^{\mathbf{p}}(\mathbf{U}_{h}, p_{h}; \mathbf{U}_{h}, p_{h}) + \|\nabla \cdot \mathbf{V}_{h}\|_{L^{2}(\Omega)}^{2} + \|\nabla q_{h}\|_{L^{2}(\Omega)}^{2} \right).$$

Thus, it suffices to show that the right hand side of (4.21) can be bounded by $a_h^{\mathbf{p}}(\mathbf{U}_h, p_h; \mathbf{U}_h, p_h)$. We further deduce that

$$\begin{aligned} \| -\nabla \cdot \mathbf{V}_h + \nabla q_h \|_{L^2(\Omega)}^2 &\leq C \left(\| -\nabla \cdot \mathbf{U}_h + \nabla p_h \|_{L^2(\mathcal{T}_h)}^2 \right) \\ &+ C \left(\| \nabla \cdot (\mathbf{U}_h - \mathbf{V}_h) \|_{L^2(\mathcal{T}_h)}^2 + \| \nabla (p_h - q_h) \|_{L^2(\mathcal{T}_h)}^2 \right), \end{aligned}$$

and since $\mathbf{U}_h \in \mathbf{I}_h^m$, we have

$$\begin{aligned} \|\mathrm{tr}\,(\mathbf{V}_{h})\|_{H^{1}(\Omega)}^{2} + \|\nabla \times \mathbf{V}_{h}\|_{L^{2}(\Omega)}^{2} &= \|\mathrm{tr}\,(\mathbf{V}_{h} - \mathbf{U}_{h})\|_{H^{1}(\mathcal{T}_{h})}^{2} + \|\nabla \times (\mathbf{V}_{h} - \mathbf{U}_{h})\|_{L^{2}(\mathcal{T}_{h})}^{2} \\ &\leq C\|(\mathbf{V}_{h} - \mathbf{U}_{h})\|_{H^{1}(\mathcal{T}_{h})}^{2}. \end{aligned}$$

Combining all inequalities above and the estimates (4.8) and (4.7), we conclude that

$$\|\mathbf{U}_h\|_{\mathbf{U}}^2 + \|p_h\|_{\mathbf{p}}^2 \le Ca_h^{\mathbf{p}}(\mathbf{U}_h, p_h; \mathbf{U}_h, p_h).$$

By a scaling argument, we can obtain that for any positive parameter η the coercivity (4.18) holds true, which completes the proof.

We have established the boundedness and coercivity of the bilinear form $a_h^{\mathbf{p}}(\cdot;\cdot)$, which implies that there exists a unique solution to the discrete problem (4.4). We state the error estimate to the numerical approximations obtained by (4.4).

Theorem 4.1. Let $(\mathbf{U}, p) \in \mathbf{I}^{m+1}(\Omega) \times H^{m+1}(\Omega)/\mathbb{R}$ be the solution to the problem (4.1) and let $(\mathbf{U}_h, p_h) \in \mathbf{I}_h^m \times \widetilde{U}_h^m$ be the solution to the problem (4.4), there exists a constant C such that $\|\mathbf{U} - \mathbf{U}_h\|_{\mathbf{U}} + \|p - p_h\|_{\mathbf{p}} \leq Ch^m \left(\|\mathbf{U}\|_{H^{m+1}(\Omega)} + \|p\|_{H^{m+1}(\Omega)}\right).$ (4.22)

Proof. For the exact solution (\mathbf{U}, p) , the jump term vanishes on interior faces, that is

$$\llbracket \mathbf{n} \times \mathbf{U} \rrbracket = 0, \quad \llbracket p \rrbracket = 0, \quad \text{on any } e \in \mathcal{E}_h^i.$$

Hence, for any $(\mathbf{V}_h, q_h) \in \mathbf{I}_h^m \times \widetilde{U}_h^m$ we have that

$$J_h^{\mathbf{p}}(\mathbf{V}_h, q_h) = a_h^{\mathbf{p}}(\mathbf{U} - \mathbf{V}_h, p - q_h; \mathbf{U} - \mathbf{V}_h, p - q_h).$$

We let $\mathbf{V}_h = \widehat{\mathcal{R}}^m \mathbf{U}$ and $q_h = \mathcal{R}^m p$, and together with the coercivity (4.18) and the boundedness (4.17), we obtain that

$$\begin{aligned} \|\mathbf{U} - \mathbf{U}_{h}\|_{\mathbf{U}} + \|p - p_{h}\|_{\mathbf{p}} &\leq Ca_{h}^{\mathbf{p}}(\mathbf{U} - \mathbf{U}_{h}, p - p_{h}; \mathbf{U} - \mathbf{U}_{h}, p - p_{h})^{1/2} = CJ_{h}^{\mathbf{p}}(\mathbf{U}_{h}, p_{h})^{1/2} \\ &\leq CJ_{h}^{\mathbf{p}}(\mathbf{V}_{h}, q_{h})^{1/2} \leq Ca_{h}^{\mathbf{p}}(\mathbf{U} - \mathbf{V}_{h}, p - q_{h}; \mathbf{U} - \mathbf{V}_{h}, p - q_{h})^{1/2} \\ &\leq C\left(\|\mathbf{U} - \mathbf{V}_{h}\|_{\mathbf{U}} + \|p - q_{h}\|_{\mathbf{p}}\right). \end{aligned}$$

Applying the approximation estimates (3.13) and (3.6) and the trace estimate, it is trivial to obtain

$$\|\mathbf{U} - \mathbf{V}_h\|_{\mathbf{U}} \le Ch^m \|\mathbf{U}\|_{H^{m+1}(\Omega)}, \quad \|p - q_h\|_{\mathbf{p}} \le Ch^m \|p\|_{H^{m+1}(\Omega)},$$

we sthe proof

which completes the proof.

The error estimate of the numerical approximation \mathbf{U}_h to the gradient with solving the minimization problem (4.3) has been established. Now let us consider another first-order system to solve the velocity \boldsymbol{u} :

$$\nabla \boldsymbol{u} - \boldsymbol{\mathrm{U}} = \boldsymbol{0}, \quad \text{in } \Omega,$$

$$\boldsymbol{u} = \boldsymbol{q}, \quad \text{on } \partial \Omega.$$
 (4.23)

We define the least squares functional $J_h^{\mathbf{u}}(\cdot)$ for solving (4.23):

$$J_{h}^{\mathbf{u}}(\boldsymbol{v}_{h}) := \sum_{K \in \mathcal{T}_{h}} \|\nabla \boldsymbol{v}_{h} - \mathbf{U}_{h}\|_{L^{2}(K)}^{2} + \sum_{e \in \mathcal{E}_{h}^{i}} \frac{\mu}{h_{e}} \|[\mathbf{n} \otimes \boldsymbol{v}_{h}]\|_{L^{2}(e)}^{2} + \sum_{e \in \mathcal{E}_{h}^{b}} \frac{\mu}{h_{e}} \|\boldsymbol{v}_{h} - \boldsymbol{g}\|_{L^{2}(e)}^{2}, \quad (4.24)$$

where μ is a positive parameter. It is noticeable that in (4.24) the first term contains the numerical approximation \mathbf{U}_h since the exact gradient is unavailable to us. We minimize the functional $J_h^{\mathbf{u}}(\cdot)$ in the piecewise divergence-free polynomial space \mathbf{S}_h^m to seek a numerical solution. The piecewise divergence-free property provides a local mass conservation, which is very desirable in solving the incompressible fluid flow problem [19, 30]. The minimization problem is given by

$$\boldsymbol{u}_h = \operatorname*{arg\,min}_{\boldsymbol{v}_h \in \mathbf{S}_h^m} J_h^{\mathbf{u}}(\boldsymbol{v}_h). \tag{4.25}$$

We also write its Euler-Lagrange equation to solve (4.25). Thus, the corresponding discrete variational problem is defined as to find $u_h \in \mathbf{S}_h^m$ such that

$$a_h^{\mathbf{u}}(\boldsymbol{u}_h, \boldsymbol{v}_h) = l_h^{\mathbf{u}}(\boldsymbol{v}_h), \quad \forall \boldsymbol{v}_h \in \mathbf{S}_h^m,$$
(4.26)

where the bilinear form $a_h^{\mathbf{u}}(\cdot, \cdot)$ is

$$a_h^{\mathbf{u}}(\boldsymbol{u}_h, \boldsymbol{v}_h) = \sum_{K \in \mathcal{T}_h} \int_K \nabla \boldsymbol{u}_h : \nabla \boldsymbol{v}_h \mathrm{d}\boldsymbol{x} + \sum_{e \in \mathcal{E}_h^i} \int_e \frac{\mu}{h_e} [\mathbf{n} \otimes \boldsymbol{u}_h] : [\mathbf{n} \otimes \boldsymbol{v}_h] \mathrm{d}\boldsymbol{s} + \sum_{e \in \mathcal{E}_h^b} \int_e \frac{\mu}{h_e} \boldsymbol{u}_h \cdot \boldsymbol{v}_h \mathrm{d}\boldsymbol{s},$$

and the linear form $l_h^{\mathbf{u}}(\cdot)$ is

$$l_h^{\mathbf{u}}(\boldsymbol{v}_h) = \sum_{K\in\mathcal{T}_h} \int_K
abla \boldsymbol{v}_h : \mathbf{U}_h \mathrm{d}\boldsymbol{x} + \sum_{e\in\mathcal{E}_h^b} \int_e rac{\mu}{h_e} \boldsymbol{v}_h \cdot \boldsymbol{g} \mathrm{d}\boldsymbol{s},$$

Then we derive the error estimate of the numerical solution to (4.24). We introduce an energy norm $\|\cdot\|_{\mathbf{u}}$ which is defined as

$$\|\boldsymbol{v}_h\|_{\mathbf{u}}^2 := \sum_{K \in \mathcal{T}_h} \|\nabla \boldsymbol{v}_h\|_{L^2(K)}^2 + \sum_{e \in \mathcal{E}_h^i} h_e^{-1} \|[\mathbf{n} \otimes \boldsymbol{v}_h]\|_{L^2(e)}^2 + \sum_{e \in \mathcal{E}_h^b} h_e^{-1} \|\boldsymbol{v}_h\|_{L^2(e)}^2,$$

for any $\boldsymbol{v}_h \in H^1(\Omega)^d + \mathbf{S}_h^m$. We state the following lemma to give a bound for the norm $\|\cdot\|_{\mathbf{u}}$.

Lemma 4.9. There exists a positive constant C such that

$$\|\boldsymbol{v}_h\|_{L^2(\Omega)} \le C \|\boldsymbol{v}_h\|_{\mathbf{u}},$$
 (4.27)

for any $\boldsymbol{v}_h \in H^1(\Omega)^d + \mathbf{S}_h^m$.

Proof. We refer to [1, 11] for the proof.

By the definition of the bilinear form
$$a_h^{\mathbf{u}}(\cdot, \cdot)$$
, it is easy to find that for any $\mu > 0$, there exist constants C such that

$$\begin{aligned} |a_h^{\mathbf{u}}(\boldsymbol{u}_h, \boldsymbol{v}_h)| &\leq C \|\boldsymbol{u}_h\|_{\mathbf{u}} \|\boldsymbol{v}_h\|_{\mathbf{u}}, \quad \forall \boldsymbol{u}_h, \boldsymbol{v}_h \in \mathbf{S}_h^m + H^1(\Omega)^d, \\ |a_h^{\mathbf{u}}(\boldsymbol{v}_h, \boldsymbol{v}_h)| &\geq C \|\boldsymbol{v}_h\|_{\mathbf{u}}^2, \qquad \forall \boldsymbol{u}_h \in \mathbf{S}_h^m, \end{aligned}$$

which implies there exists a unique solution to the problem (4.26). Finally, we present the error estimate of the numerical solution in approximation to the velocity u.

Theorem 4.2. Let $\mathbf{u} \in \mathbf{S}^{m+1}(\Omega)$ be the solution to (2.4) and let $\mathbf{u}_h \in \mathbf{S}_h^m$ be the solution to (4.26), there exists a positive constant C such that

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\mathbf{u}} \le C\left(h^m \|\boldsymbol{u}\|_{H^{m+1}(\Omega)} + \|\mathbf{U} - \mathbf{U}_h\|_{L^2(\Omega)}\right),\tag{4.28}$$

where \mathbf{U}_h is the numerical solution in (4.3).

Proof. For smooth \boldsymbol{u} , the trace $[\mathbf{n} \otimes \boldsymbol{u}] = 0$ on any interior faces. We let $\boldsymbol{v}_h = \widetilde{\mathcal{R}}^m \boldsymbol{u}$ be the interpolant of \boldsymbol{u} , and we obtain that

$$\begin{split} \|\boldsymbol{u} - \boldsymbol{u}_{h}\|_{\mathbf{u}}^{2} &= \sum_{K \in \mathcal{T}_{h}} \|\nabla \boldsymbol{u} - \nabla \boldsymbol{u}_{h}\|_{L^{2}(K)}^{2} + \sum_{e \in \mathcal{E}_{h}^{i}} h_{e}^{-1} \|[\mathbf{n} \otimes (\boldsymbol{u} - \boldsymbol{u}_{h})]\|_{L^{2}(e)}^{2} + \sum_{e \in \mathcal{E}_{h}^{b}} h_{e}^{-1} \|\boldsymbol{u} - \boldsymbol{u}_{h}\|_{L^{2}(e)}^{2} \\ &\leq C \left(\sum_{K \in \mathcal{T}_{h}} \|\mathbf{U}_{h} - \nabla \boldsymbol{u}_{h}\|_{L^{2}(K)}^{2} + \sum_{K \in \mathcal{T}_{h}} \|\mathbf{U} - \mathbf{U}_{h}\|_{L^{2}(K)}^{2} \right) \\ &+ \sum_{e \in \mathcal{E}_{h}^{i}} h_{e}^{-1} \|[\mathbf{n} \otimes (\boldsymbol{u} - \boldsymbol{u}_{h})]\|_{L^{2}(e)}^{2} + \sum_{e \in \mathcal{E}_{h}^{b}} h_{e}^{-1} \|\boldsymbol{u} - \boldsymbol{u}_{h}\|_{L^{2}(e)}^{2} \\ &\leq C \left(J_{h}^{\mathbf{u}}(\boldsymbol{u}_{h}) + \|\mathbf{U} - \mathbf{U}_{h}\|_{L^{2}(\Omega)}^{2} \right) \leq C \left(J_{h}^{\mathbf{u}}(\boldsymbol{v}_{h}) + \|\mathbf{U} - \mathbf{U}_{h}\|_{L^{2}(\Omega)}^{2} \right) \\ &\leq C \left(\|\boldsymbol{u} - \boldsymbol{v}_{h}\|_{\mathbf{u}} + \|\mathbf{U} - \mathbf{U}_{h}\|_{L^{2}(\Omega)}^{2} \right)^{2} \\ &\leq C \left(h^{m} \|\boldsymbol{u}\|_{H^{m+1}(\Omega)} + \|\mathbf{U} - \mathbf{U}_{h}\|_{L^{2}(\Omega)}^{2} \right)^{2}. \end{split}$$

The last inequality follows from the approximation property (3.9) and the trace estimate M2, which completes the proof.

5. Numerical Results

In this section, we present a series of numerical experiments to demonstrate the accuracy of our method in both two dimensions and three dimensions. We take the accuracy order as $1 \leq m \leq 3$ and for different m we list the values #S that are used in numerical experiments in Table 5.1. For all test problems, the Reynolds number Re and the parameters η and μ are chosen to be 1.

d = 2	m	1	2	3
u – 2	#S	5	10	15
d-3	m	1	2	3
u = 0	#S	8	18	36

Table 5.1: #S for $1 \le m \le 3$.



Fig. 5.1. The triangular meshes with mesh size h = 1/10 (left) / h = 1/20 (right) for Example 5.1.

5.1. 2D Example

Example 5.1. We solve the Stokes problem (2.3) on the domain $\Omega = (0, 1)^2$ with the analytical solution

$$\boldsymbol{u}(x,y) = \begin{bmatrix} \sin(2\pi x)\cos(2\pi y) \\ -\cos(2\pi x)\sin(2\pi y) \end{bmatrix}, \quad p(x,y) = x^2 + y^2 - \frac{2}{3},$$

to show the convergence rates of our method. The source term f and the Dirichlet data g are taken accordingly. We employ a series of triangular meshes with mesh size $h = 1/10, 1/20, \ldots, 1/160$, see Fig. 5.1. The numerical results are displayed in Fig. 5.2 and Fig. 5.3. For the first part (4.1), we plot the numerical error under energy norm $\|\mathbf{U} - \mathbf{U}_h\|_{\mathbf{U}} + \|p - p_h\|_{\mathbf{P}}$ in Fig. 5.2, which approaches zero at the speed $O(h^m)$. The convergence rates are consistent with the theoretical analysis (4.22). For L^2 error, we also plot the errors $\|\mathbf{U} - \mathbf{U}_h\|_{L^2(\Omega)}$ and $\|p - p_h\|_{L^2(\Omega)}$ in Fig. 5.2, and we numerically detect the odd/even situation. For odd m, the L^2 errors seem to converge to zero at the optimal speed, and for even m, the errors have a sub-optimal convergence rate. For the second system (4.23), we show the numerical results in



Fig. 5.2. The convergence rates of $\|\mathbf{U} - \mathbf{U}_h\|_{\mathbf{U}} + \|p - p_h\|_{\mathbf{p}}$ (left) / $\|\mathbf{U} - \mathbf{U}_h\|_{L^2(\Omega)}$ (middle) / $\|p - p_h\|_{L^2(\Omega)}$ (right) for Example 5.1.



Fig. 5.3. The convergence rates of $\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\mathbf{u}}$ (left) / $\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega)}$ (right) for Example 5.1.

Fig. 5.3. Clearly, the error under energy norm $\|\cdot\|_{\mathbf{u}}$ converges to zero with the rate $O(h^m)$ as the mesh size approaches 0. For the L^2 norm, we also observe the optimal rate and sub-optimal rate for odd m and even m, respectively. We note that the convergence orders under all error measurements are in perfect agreement with our theoretical error estimates.

Example 5.2. Here we solve the Stokes problem to show the great flexibility of the proposed method. The exact solution and the computational domain are taken the same as in Example 1 but in the example we use a series of polygonal meshes with 250, 1000, 4000, 16000 elements. These meshes consist of very general elements and are generated by PolyMesher [32], see Fig. 5.4. The numerical errors under all error measurements for both two systems (4.1) and (4.23) are shown in Fig. 5.5 and Fig. 5.6, respectively. Again we observe the optimal convergence orders for all energy norms. For L^2 norm, the odd/even situation is still detected. On such polygonal meshes, the numerically computed orders agree with our error estimates.



Fig. 5.4. The polygonal meshes with 250 elements (left) / 1000 elements (right) for Example 5.2.



Fig. 5.5. The convergence rates of $\|\mathbf{U} - \mathbf{U}_h\|_{\mathbf{U}} + \|p - p_h\|_{\mathbf{p}}$ (left) / $\|\mathbf{U} - \mathbf{U}_h\|_{L^2(\Omega)}$ (middle) / $\|p - p_h\|_{L^2(\Omega)}$ (right) for Example 5.2.

Example 5.3. In this example, we test the modified lid-driven cavity problem [29] to investigate the performance of our method dealing with the problem with low regularities. We consider the unit square domain $\Omega = (0, 1)^2$, which is subjected to a horizontal flow on the boundary y = 1 with the velocity $\boldsymbol{u}(x, y) = (4x(1-x), 0)^T$. The condition of remaining boundaries is no-slip boundary condition. The source term \boldsymbol{f} is selected to be $(0, 0)^T$. The velocity field on the upper boundary involves singularity in the upper right and left corners, but the restraints are not as strong as for the well-known standard lid-driven cavity problem [29]. We solve this



Fig. 5.6. The convergence rates of $\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\mathbf{u}}$ (left) / $\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega)}$ (right) for Example 5.2.



Fig. 5.7. The convergence rates of $\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\mathbf{u}}$ (left) / $\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega)}$ (right) for Example 5.3.



Fig. 5.8. Velocity vectors (left) and the streamline of the flow (right) with the accuracy m = 2 for Example 5.3.

problem on the triangular partition with h = 1/10, h = 1/20, h = 1/40 and h = 1/80, see Fig. 5.1. Since the analytical solution is unknown and we take the numerical solution which is obtained with the mesh size h = 1/320 and the accuracy m = 3 as the exact solution. The numerical errors in approximation to the velocity are presented in Fig. 5.7. The convergence



Fig. 5.9. Velocity vectors (left) and the streamline of the flow (right) with the accuracy m = 3 for Example 5.3.

rates under energy norms and L^2 norms are detected to be O(h) for all accuracy $1 \le m \le 3$. A possible explanation of such convergence rates can be traced back to the low regularity of this problem. Figure 5.8 and Figure 5.9 present the numerical results obtained on the mesh level h = 1/80 with the accuracy m = 2 and m = 3, respectively. We can observe the main vortex in the center of the domain and two small vortices in the bottom left and right corners.

5.2. 3D Example

Example 5.4. In this example, we consider the Stokes problem in three dimensions. We solve the problem on the domain $\Omega = (0, 1)^3$ and we take a series of tetrahedral meshes with the resolution h = 1/4, h = 1/8 and h = 1/16 (see Fig. 5.10). We choose the analytical solution \boldsymbol{u} and p as

$$\boldsymbol{u}(x,y,z) = \begin{bmatrix} 1 - e^x \cos(2\pi y) \\ \frac{1}{2\pi} e^x \sin(2\pi y) \\ 0 \end{bmatrix}, \quad p(x,y,z) = x^2 + y^2 - \frac{2}{3}$$

The numerical errors in approximation to the gradient and the pressure are collected in Table 5.2, and the numerical errors in solving the second first-order system are gathered in Table 5.3. We also depict the velocity field and the contour of $|u_h|$ in Fig. 5.11 and the numerical solution in this figure is obtained on the mesh level h = 1/16 with the accuracy m = 3. Here,

Table 5.2: The numerical results of $\|\mathbf{U} - \mathbf{U}_h\|_{\mathbf{U}} + \|p - p_h\|_{\mathbf{p}}$, $\|\mathbf{U} - \mathbf{U}_h\|_{L^2(\Omega)}$ and $\|p - p_h\|_{L^2(\Omega)}$ for Example 5.4.

m	h	$\ \mathbf{U}-\mathbf{U}_h\ _{\mathbf{U}}+\ p-p_h\ _{\mathbf{p}}$	order	$\ \mathbf{U}-\mathbf{U}_h\ _{L^2(\Omega)}$	order	$\ p-p_h\ _{L^2(\Omega)}$	order
	1/4	4.162e + 1	-	2.022e-0	-	3.421e-0	-
1	1/8	2.446e+1	0.77	9.998e-1	1.02	1.743e-0	0.97
	1/16	1.206e + 1	1.02	3.653e-1	1.46	7.143e-1	1.29
	1/4	1.476e + 1	-	6.388e-1	-	1.053e-0	-
2	1/8	4.213e-0	1.81	1.284e-1	2.31	3.425e-1	1.62
	1/16	1.167e-0	1.92	3.128e-2	2.03	7.752e-2	2.13
	1/4	4.125e-0	-	1.557e-1	-	1.409e-1	-
3	1/8	4.913e-1	3.06	1.125e-2	3.79	1.010e-2	3.80
	1/16	5.431e-2	3.13	7.507e-4	3.91	6.931e-4	3.86



Fig. 5.10. The tetrahedral meshes with mesh size h = 1/8 (left) / mesh size h = 1/16 (right) for three-dimensional examples.



Fig. 5.11. The velocity field (left) / the contour of $|u_h|$ (right) for Example 5.4.

$\mid m$	h	$\ \ oldsymbol{u} - oldsymbol{u}_h \ _{\mathbf{u}}$	order	$\ oldsymbol{u}-oldsymbol{u}_h\ _{L^2(\Omega)}$	order
	1/4	3.495e-0	-	2.865e-1	-
1	1/8	2.075e-0	0.75	1.278e-1	1.16
	1/16	1.062e-0	0.97	5.081e-2	1.33
	1/4	2.395e-0	-	1.509e-1	-
2	1/8	5.402e-1	2.15	3.259e-2	2.21
	1/16	1.367e-1	1.98	7.653e-3	2.09
	1/4	4.640e-1	-	2.307e-2	-
3	1/8	5.382e-2	3.10	1.747e-3	3.72
	1/16	6.649e-3	3.02	1.255e-4	3.81

Table 5.3: The numerical results of $\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\mathbf{u}}$ and $\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega)}$ for Example 5.4.

we still observe the odd/even situation. For odd m, the errors under L^2 norm seem to converge to zero optimally as the mesh size tends to zero. For even m, the convergence rates for all variables under L^2 norm are numerically detected to be sub-optimal. Again we note that all computed convergence orders are consistent with theoretical analysis.

Example 5.5. In the last example, we solve another three-dimensional test problem. The domain and the meshes are selected the same as the previous example. For this test, the exact



Fig. 5.12. The velocity field (left) / the contour of $|u_h|$ (right) for Example 5.5.

Table 5.4: The numerical results of $\|\mathbf{U} - \mathbf{U}_h\|_{\mathbf{U}} + \|p - p_h\|_{\mathbf{p}}$, $\|\mathbf{U} - \mathbf{U}_h\|_{L^2(\Omega)}$ and $\|p - p_h\|_{L^2(\Omega)}$ for Example 5.5.

m	h	$\ \mathbf{U} - \mathbf{U}_h\ _{\mathbf{U}} + \ p - p_h\ _{\mathbf{p}}$	order	$\ \mathbf{U}-\mathbf{U}_h\ _{L^2(\Omega)}$	order	$\ p-p_h\ _{L^2(\Omega)}$	order
	1/4	7.312e-0	-	3.516e-1	-	1.932e-1	-
1	1/8	3.691e-0	0.98	1.101e-1	1.67	9.032e-2	1.10
	1/16	1.831e-0	1.01	3.105e-2	1.82	3.132e-2	1.53
	1/4	1.946e-0	-	5.619e-2	-	7.300e-2	-
2	1/8	5.082e-1	1.93	1.163e-2	2.27	1.921e-2	1.92
	1/16	1.287e-1	1.98	2.646e-3	2.12	4.932e-3	1.98
	1/4	7.028e-1	-	2.878e-2	-	2.609e-2	-
3	1/8	7.992e-2	3.13	1.881e-3	3.93	1.848e-3	3.82
	1/16	9.250e-4	3.11	1.181e-4	4.01	1.093e-4	4.08

Table 5.5: The numerical results of $\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\mathbf{u}}$ and $\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega)}$ for Example 5.5.

m	h	$\ oldsymbol{u}-oldsymbol{u}_h\ _{\mathbf{u}}$	order	$\ oldsymbol{u}-oldsymbol{u}_h\ _{L^2(\Omega)}$	order
	1/4	1.217e-0	-	6.262e-2	-
1	1/8	6.123e-1	0.99	2.155e-2	1.53
	1/16	3.041e-1	1.00	6.256e-3	1.78
	1/4	3.311e-1	-	1.839e-2	-
2	1/8	8.045e-2	2.04	3.552e-3	2.37
	1/16	1.975e-2	2.02	7.529e-4	2.23
	1/4	1.157e-1	-	6.098e-3	-
3	1/8	1.336e-2	3.11	3.874e-4	3.97
	1/16	1.561e-3	3.09	2.357e-5	4.03

solution is

$$\boldsymbol{u}(x,y,z) = \begin{bmatrix} \sin(\pi x)\cos(\pi y)e^{-2z} \\ \cos(\pi x)\sin(\pi y)e^{-2z} \\ \pi\cos(\pi x)\cos(\pi y)e^{-2z} \end{bmatrix}, \quad p(x,y,z) = x^2 + y^2 + z^2 - 1,$$

and the source term and the boundary data are taken from u and p. We list the numerical errors in approximation to the gradient of the velocity and pressure in Table 5.4 and the numerical errors of the velocity are given in Table 5.5. It can be clearly seen that the convergence rates for all unknowns are optimal under energy norms, and the old/even situation of L^2 errors is also observed. In addition, we plot the numerical solution with the mesh resolution h = 1/16

and the accuracy m = 3 to show the velocity field and the contour of $|u_h|$ in Fig. 5.12.

6. Conclusion

We constructed three types of approximation spaces by patch reconstructions. These reconstructed discontinuous spaces allow us to numerically solve the Stokes problem in two sequential steps. In the three spaces, the gradient of velocity, the velocity and the pressure are approximated, respectively. We first employed a reconstructed space that consists of piecewise curl-free polynomials with zero trace to approximate the gradient of the velocity and the pressure. Then we obtained the approximation to the velocity in the reconstructed piecewise divergence-free space. The convergence rates for all unknowns under L^2 norms and energy norms are derived. We presented a series of numerical tests in two and three dimensions to verify the error estimates and illustrate the great flexibility of the method we proposed. In addition, the computer program is able to handle approximation spaces of any high order and the elements with various geometry in a uniform manner.

Appendix

In Appendix, we present the detailed computer implementation of constructing the approximation spaces introduced in Section 3. The construction contains two steps that are constructing element patches and solving local least squares problems on every element. We first give the recursive algorithm to the construction of the element patch, see Alg. A.1. From Alg. A.1, it can be seen that for each element K, the construction to the corresponding element patch involves two parts: enlarging the set $S_t(K)$ and sorting distances between points in $\mathcal{I}(K)$ and \mathbf{x}_K . The two parts require at most $O((\#S)^2)$ computational cost, which gives us that the computation complexity to Alg. A.1 is $O(N_e(\#S)^2)$, where N_e is the number of all elements. Thus, Alg. A.1 is very fast since #S is usually much smaller than N_e .

Algorithm A.1. Construction to Element Patch						
Input: a partition \mathcal{T}_h and a threshold $\#S$;						
Output: the element patch of each element in \mathcal{T}_h ;						
1: for every $K \in \mathcal{T}_h$ do						
2: initialize $t = 0, S_t(K) = \{K\};$						
3: while the cardinality of $S_t(K) < \#S$ do						
4: set $S_{t+1}(K) = S_t(K);$						
5: for every $\widetilde{K} \in S_t(K)$ do						
6: add all adjacent face-neighbouring elements of \widetilde{K} to $S_{t+1}(K)$;						
7: end for						
8: let $t = t + 1$;						
9: end while						
10: collect collocation points of all elements in $S_t(K)$ in $\mathcal{I}(K)$;						
11: sort the distances between points in $\mathcal{I}(K)$ and \boldsymbol{x}_K ;						
12: select the $\#S$ smallest values and collect the corresponding elements to form						
S(K);						
13: end for						

Then we will explain how to solve the least squares problems (3.1), (3.7) and (3.10). The key of solving least squares problems is to construct a group of polynomial bases that satisfy the constraints in (3.7) and (3.10). For (3.7), we shall construct the bases of the divergence-free polynomial space, that is $\mathbb{P}_k(D)^d \cap \mathbf{S}^0(D)$, here D is a bounded domain and k is a positive integer. In two dimensions, we can directly take the curl of the natural polynomial bases

$$1, x, y, x^2, xy, y^2, x^3, x^2y, xy^2, y^3, \dots$$

to obtain the bases of divergence-free polynomials, as illustrated in [2]. For example, if the linear accuracy is considered we have that

$$\mathbb{P}_1(D)^2 \cap \mathbf{S}^0(D) = \operatorname{span}\left\{ \begin{bmatrix} 1\\ 0 \end{bmatrix}, \begin{bmatrix} 0\\ 1 \end{bmatrix}, \begin{bmatrix} 0\\ x \end{bmatrix}, \begin{bmatrix} y\\ 0 \end{bmatrix}, \begin{bmatrix} x\\ -y \end{bmatrix} \right\}.$$

For the second-order case, it is easy to find that

$$\mathbb{P}_{2}(D)^{2} \cap \mathbf{S}^{0}(D)$$

$$= \operatorname{span}\left\{ \begin{bmatrix} 1\\ 0 \end{bmatrix}, \begin{bmatrix} 0\\ 1 \end{bmatrix}, \begin{bmatrix} 0\\ x \end{bmatrix}, \begin{bmatrix} y\\ 0 \end{bmatrix}, \begin{bmatrix} x\\ -y \end{bmatrix}, \begin{bmatrix} 0\\ x^{2} \end{bmatrix}, \begin{bmatrix} x^{2}\\ -2xy \end{bmatrix}, \begin{bmatrix} -2xy\\ y^{2} \end{bmatrix}, \begin{bmatrix} y^{2}\\ 0 \end{bmatrix} \right\}.$$

To get a group of divergence-free polynomials bases in three dimensions is a bit more complicated and we outline a method which is easy to implement. We construct two groups of polynomials $\widetilde{\mathbf{S}}_{k}^{1}(D)$ and $\widetilde{\mathbf{S}}_{k}^{2}(D)$ whose union actually forms a group of bases. The first group $\widetilde{\mathbf{S}}_{k}^{1}(D)$ consists of the vector-valued polynomials which only have one nonzero entry. Specifically speaking, $\widetilde{\mathbf{S}}_{k}^{1}(D)$ has three types of polynomials, that is $\widetilde{\mathbf{S}}_{k}^{1}(D) = \mathbf{Q}_{k}^{1}(D) \cup \mathbf{Q}_{k}^{2}(D) \cup \mathbf{Q}_{k}^{3}(D)$, where

$$\mathbf{Q}_{k}^{1}(D) := \left\{ \boldsymbol{p} = (p_{1}, 0, 0)^{T} \in \mathbb{P}_{k}(D)^{3} \mid p_{1}(y, z) \in \mathbb{P}_{k}(y, z) \right\},
\mathbf{Q}_{k}^{2}(D) := \left\{ \boldsymbol{p} = (0, p_{2}, 0)^{T} \in \mathbb{P}_{k}(D)^{3} \mid p_{2}(x, z) \in \mathbb{P}_{k}(x, z) \right\},
\mathbf{Q}_{k}^{3}(D) := \left\{ \boldsymbol{p} = (0, 0, p_{3})^{T} \in \mathbb{P}_{k}(D)^{3} \mid p_{3}(y, z) \in \mathbb{P}_{k}(x, y) \right\},$$
(A.1)

and $\mathbb{P}_k(a, b)$ denotes the polynomial space of degree k based on the coordinate (a, b),

$$\mathbb{P}_{k}(a,b) = \operatorname{span}\left\{1, a, b, a^{2}, ab, b^{2}, \dots, a^{k}, a^{k-1}b, \dots, ab^{k-1}, b^{k}\right\}.$$

Then $\widetilde{\mathbf{S}}_k^2(D)$ has two types of polynomials, that is $\widetilde{\mathbf{S}}_k^2(D) = \mathbf{R}_k^1(D) \cup \mathbf{R}_k^2(D)$, where

$$\mathbf{R}_{k}^{1}(D) := \left\{ \boldsymbol{p} = (p_{1}, p_{2}, 0)^{T} \in \mathbb{P}_{k}(D)^{3} \mid p_{1} = x^{t}q, \ p_{2} = -\int q(y, z)dy, q(y, z) \in \mathbb{P}_{k-t}(y, z), \ 1 \le t \le k \right\},$$

$$\mathbf{R}_{k}^{2}(D) := \left\{ \boldsymbol{p} = (p_{1}, 0, p_{3})^{T} \in \mathbb{P}_{k}(D)^{3} \mid p_{1} = x^{t}q, \ p_{3} = -\int q(y, z)dz, q(y, z) \in \mathbb{P}_{k-t}(y, z), \ 1 \le t \le k \right\}.$$
(A.2)

It is trivial to verify that the polynomials in $\widetilde{\mathbf{S}}_{k}^{1}(D)$ and $\widetilde{\mathbf{S}}_{k}^{2}(D)$ are divergence-free, and we state the following lemma.

Lemma A.1. The divergence-free polynomial space $\mathbb{P}_k(D)^3 \cap \mathbf{S}^0(D)$ satisfies that $\mathbb{P}_k(D)^3 \cap \mathbf{S}^0(D) = \widetilde{\mathbf{S}}_k^1(D) \cup \widetilde{\mathbf{S}}_k^2(D)$.

Proof. We let $q \in \mathbb{P}_k(D)^3$ such that $q = \widetilde{\mathbf{S}}_k^1(D) \cap \widetilde{\mathbf{S}}_k^2(D)$. By the definition (A.1), the first entry of q only depends on y and z. From (A.2), the first entry of q must rely on x, which

gives $\boldsymbol{q} = \boldsymbol{0}$. Hence, we have that $\widetilde{\mathbf{S}}_{k}^{1}(D) \cap \widetilde{\mathbf{S}}_{k}^{2}(D) = \{\boldsymbol{0}\}$. From (A.1) and (A.2), we can know that $\dim(\widetilde{\mathbf{S}}_{k}^{1}(D)) = 3(k+2)(k+1)/2$ and $\dim(\widetilde{\mathbf{S}}_{k}^{2}(D)) = (k+2)(k+1)k/3$. By [2], we have that $\dim(\mathbb{P}_{k}(D)^{3} \cap \mathbf{S}^{0}(D)) = 3C_{k+3}^{3} - C_{k+2}^{3}$, which exactly implies $\dim(\mathbb{P}_{k}(D)^{3} \cap \mathbf{S}^{0}(D)) = \dim(\widetilde{\mathbf{S}}_{k}^{1}(D)) + \dim(\widetilde{\mathbf{S}}_{k}^{2}(D))$. This fact gives us that $\mathbb{P}_{k}(D)^{3} \cap \mathbf{S}^{0}(D) = \widetilde{\mathbf{S}}_{k}^{1}(D) \cup \widetilde{\mathbf{S}}_{k}^{2}(D)$ and completes the proof.

Further, we give an example of the linear accuracy. In this case, we can obtain that

$$\widetilde{\mathbf{S}}_{1}^{1}(D) = \operatorname{span}\left\{ \begin{bmatrix} 1\\0\\0 \end{bmatrix}, \begin{bmatrix} 0\\1\\0 \end{bmatrix}, \begin{bmatrix} 0\\0\\1 \end{bmatrix}, \begin{bmatrix} y\\0\\0 \end{bmatrix}, \begin{bmatrix} z\\0\\0 \end{bmatrix}, \begin{bmatrix} 0\\z\\0 \end{bmatrix}, \begin{bmatrix} 0\\z\\0 \end{bmatrix}, \begin{bmatrix} 0\\z\\0 \end{bmatrix}, \begin{bmatrix} 0\\0\\z \end{bmatrix}, \begin{bmatrix} 0\\0\\z \end{bmatrix}, \begin{bmatrix} 0\\0\\z \end{bmatrix} \right\},$$

and

$$\widetilde{\mathbf{S}}_{2}^{2}(D) = \operatorname{span}\left\{ \begin{bmatrix} x \\ -y \\ 0 \end{bmatrix}, \begin{bmatrix} x \\ 0 \\ -z \end{bmatrix} \right\}.$$

Hence, $\mathbb{P}_1(D)^3 \cap \mathbf{S}^0(D) = \widetilde{\mathbf{S}}_1^1(D) \cup \widetilde{\mathbf{S}}_1^2(D).$

Then we consider to solve the problem (3.10), which requires us to construct the polynomial space consists of the curl-free polynomials with zero trace, that is $\mathbb{P}_k(D)^{d \times d} \cap \mathbf{I}^0(D)$. Actually after obtaining the bases of the divergence-free polynomial space, it is easy to get the bases of the polynomial space $\mathbb{P}_k(D)^{d \times d} \cap \mathbf{I}^0(D)$. We can take the gradient of the divergence-free polynomial bases to get those bases. Again we take k = 1 for an example and we can obtain that

$$\mathbb{P}_{1}(D)^{d \times d} \cap \mathbf{I}^{0}(D)$$

=span $\left\{ \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & x \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} x & -y \\ 0 & -x \end{bmatrix}, \begin{bmatrix} -y & 0 \\ -x & y \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ y & 0 \end{bmatrix} \right\}.$

In the rest of Appendix, we present the details of the computer implementation to the reconstructed space. Let us construct the space \mathbf{I}_{h}^{1} in two dimensions as an illustration. We consider the element K_{0} and we let its element patch $S(K_{0})$ formed by its face-neighbouring elements, see Fig. A.1.



Fig. A.1. K_0 and its neighbours (left) / collocation points (right).

For a tensor-valued function $\boldsymbol{g} = (g_{00}, g_{01}; g_{10}, g_{11})^T \in \mathbf{U}_h(g_{00} = -g_{11})$, the least squares problem (3.7) on $S(K_0)$ takes the form

$$\widehat{\mathcal{R}}_{K_0}^1 \boldsymbol{g} = \min_{\boldsymbol{q} \in \mathbb{P}_m(S(K_0))^{2\times} \cap \mathbf{I}^0(S(K_0))} \sum_{\boldsymbol{x} \in \mathcal{I}(K_0)} \|\boldsymbol{q}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{x})\|_{l^2 \times l^2}^2, \quad \text{s.t. } \boldsymbol{q}(\boldsymbol{x}_{K_0}) = \boldsymbol{g}(\boldsymbol{x}_{K_0}).$$
(A.3)

From the bases of the polynomial space $\mathbb{P}_1(S(K_0))^2 \cap \mathbf{I}^0(S(K))$, the polynomial $p(\mathbf{x})$ in (A.3) has the form

$$p(\mathbf{x}) = a_0 \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + a_1 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + a_2 \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + a_3 \begin{bmatrix} 0 & x \\ 0 & 0 \end{bmatrix} + a_4 \begin{bmatrix} x & -y \\ 0 & -x \end{bmatrix} + a_5 \begin{bmatrix} -y & 0 \\ -x & y \end{bmatrix} + a_6 \begin{bmatrix} 0 & 0 \\ y & 0 \end{bmatrix}.$$

By the constraint in (A.3), we can know the values of a_0 , a_1 and a_2 and we rewrite the polynomial p(x) as

$$\boldsymbol{p}(\boldsymbol{x}) = \begin{bmatrix} g_{00}(\boldsymbol{x}_{K_0}) & g_{01}(\boldsymbol{x}_{K_0}) \\ g_{10}(\boldsymbol{x}_{K_0}) & g_{11}(\boldsymbol{x}_{K_0}) \end{bmatrix} + a_3 \begin{bmatrix} 0 & x - x_{K_0} \\ 0 & 0 \end{bmatrix} + a_4 \begin{bmatrix} x - x_{K_0} & -y + y_{K_0} \\ 0 & -x + x_{K_0} \end{bmatrix} \\ + a_5 \begin{bmatrix} -y + y_{K_0} & 0 \\ -x + x_{K_0} & y - y_{K_0} \end{bmatrix} + a_6 \begin{bmatrix} 0 & 0 \\ y - y_{K_0} & 0 \end{bmatrix},$$

where $\boldsymbol{x}_{K_i} = (x_{K_i}, y_{K_i}) (0 \le i \le 3)$. Thus the problem (A.3) is equivalent to

$$\arg \min_{a_{3},a_{4},a_{5},a_{6}\in\mathbb{R}} \sum_{i=1}^{5} \left\| a_{3} \begin{bmatrix} 0 & x_{K_{i}} - x_{K_{0}} \\ 0 & 0 \end{bmatrix} + a_{4} \begin{bmatrix} x_{K_{i}} - x_{K_{0}} & -y_{K_{i}} + y_{K_{0}} \\ 0 & -x_{K_{i}} + x_{K_{0}} \end{bmatrix} + a_{5} \begin{bmatrix} -y_{K_{i}} + y_{K_{0}} & 0 \\ -x_{K_{i}} + x_{K_{0}} & y_{K_{i}} - y_{K_{0}} \end{bmatrix} + a_{6} \begin{bmatrix} 0 & 0 \\ y_{K_{i}} - y_{K_{0}} & 0 \end{bmatrix} \\
- \begin{bmatrix} g_{00}(\boldsymbol{x}_{K_{i}}) - g_{00}(\boldsymbol{x}_{K_{0}}) & g_{01}(\boldsymbol{x}_{K_{i}}) - g_{01}(\boldsymbol{x}_{K_{0}}) \\ g_{10}(\boldsymbol{x}_{K_{i}}) - g_{10}(\boldsymbol{x}_{K_{0}}) & g_{11}(\boldsymbol{x}_{K_{i}}) - g_{11}(\boldsymbol{x}_{K_{0}}) \end{bmatrix} \right\|_{l^{2}\times l^{2}}^{2}. \tag{A.4}$$

The solution to (A.4) reads

$$\begin{bmatrix} a_3\\ a_4\\ a_5\\ a_6 \end{bmatrix} = (A^T A)^{-1} A^T \begin{bmatrix} 2(g_{00}(\boldsymbol{x}_{K_1}) - g_{00}(\boldsymbol{x}_{K_0}))\\ g_{01}(\boldsymbol{x}_{K_1}) - g_{01}(\boldsymbol{x}_{K_0})\\ g_{10}(\boldsymbol{x}_{K_1}) - g_{10}(\boldsymbol{x}_{K_0})\\ \cdots\\ 2(g_{00}(\boldsymbol{x}_{K_3}) - g_{00}(\boldsymbol{x}_{K_0}))\\ g_{01}(\boldsymbol{x}_{K_3}) - g_{01}(\boldsymbol{x}_{K_0})\\ g_{10}(\boldsymbol{x}_{K_3}) - g_{10}(\boldsymbol{x}_{K_0}) \end{bmatrix},$$

where

$$A = \begin{bmatrix} 0 & 2(x_{K_1} - x_{K_0}) & -2(y_{K_1} - y_{K_0}) & 0 \\ x_{K_1} - x_{K_0} & -y_{K_1} + y_{K_0} & 0 & 0 \\ 0 & 0 & -x_{K_1} + x_{K_0} & y_{K_1} - y_{K_0} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 2(x_{K_3} - x_{K_0}) & -2(y_{K_3} - y_{K_0}) & 0 \\ x_{K_3} - x_{K_0} & -y_{K_3} + y_{K_0} & 0 & 0 \\ 0 & 0 & -x_{K_3} + x_{K_0} & y_{K_3} - y_{K_0} \end{bmatrix}.$$

By rearrangement, we can obtain the solution to (A.3), which takes the form

$$\begin{bmatrix} a_0\\a_1\\a_2\\a_3\\a_4\\a_5\\a_6 \end{bmatrix} = \begin{bmatrix} I_{3\times3} & 0\\-MI_{9\times3} & M \end{bmatrix} \begin{bmatrix} g_{00}(\boldsymbol{x}_{K_0})\\g_{01}(\boldsymbol{x}_{K_0})\\g_{10}(\boldsymbol{x}_{K_0})\\\vdots\\g_{00}(\boldsymbol{x}_{K_3})\\g_{01}(\boldsymbol{x}_{K_3})\\g_{10}(\boldsymbol{x}_{K_3}) \end{bmatrix}, \quad M = (A^T A)^{-1} A^T \begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0\\0 & I_{2\times2} & 0 & 0 & 0 & 0\\0 & 0 & 0 & I_{2\times2} & 0 & 0\\0 & 0 & 0 & 0 & 0 & I_{2\times2} \end{bmatrix},$$

where $I_{2\times 2}$ and $I_{3\times 3}$ are 2×2 identity matrix and 3×3 identity matrix and $I_{9\times 3} = (I_{3\times 3}, I_{3\times 3}, I_{3\times 3})^T$. We note that the collocation points in $\mathcal{I}(K)$ totally determine the matrix M, and by the expansion (3.11), the coefficient matrix

$$\begin{bmatrix} I_{3\times3} & 0\\ -MI_{9\times3} & M \end{bmatrix}$$
(A.5)

actually contains all information of the basis functions $\widehat{\lambda}_{K_i}^{j,k}$ $(0 \le i \le 3, 1 \le j, k \le 2, j + k < 4)$ on the element K_0 . Then we can use the coefficient matrix (A.5) on each element to represent the reconstructed space \mathbf{I}_h^1 . For the spaces U_h^1 and \mathbf{S}_h^1 , their constructions are very similar. In addition, such a computer implementation can be easily adapted to three dimensions and the case when higher-order accuracy is considered.

Acknowledgements. This research was supported by the Science Challenge Project (No. TZ2016002) and the National Natural Science Foundation in China (No.11971041 and 11421101).

References

- D.N. Arnold, An interior penalty finite element method with discontinuous elements, SIAM J. Numer. Anal., 19:4 (1982), 742–760.
- [2] G.A. Baker, W.N. Jureidini, and O.A. Karakashian, Piecewise solenoidal vector fields and the Stokes problem, SIAM J. Numer. Anal., 27:6 (1990), 1466–1485.
- [3] R. Bensow and M.G. Larson, Discontinuous least-squares finite element method for the div-curl problem, *Numer. Math.*, 101:4 (2005), 601–617.
- [4] R.E. Bensow and M. G. Larson, Discontinuous/continuous least-squares finite element methods for elliptic problems, *Math. Models Methods Appl. Sci.*, 15:6 (2005), 825–842.
- [5] F. Bertrand, Z. Cai, and E.Y. Park, Least-squares methods for elasticity and Stokes equations with weakly imposed symmetry, *Comput. Methods Appl. Math.*, 19:3 (2019), 415–430.
- [6] P. Bochev, James Lai, and Luke Olson, A locally conservative, discontinuous least-squares finite element method for the Stokes equations, *Internat. J. Numer. Methods Fluids*, 68:6 (2012), 782– 804.
- [7] P. Bochev, James Lai, and Luke Olson, A non-conforming least-squares finite element method for incompressible fluid flow problems, *Internat. J. Numer. Methods Fluids*, **72**:3 (2013), 375–402.
- [8] P.B. Bochev and M.D. Gunzburger, Analysis of least squares finite element methods for the Stokes equations, *Math. Comp.*, 63:208 (1994), 479–506.
- [9] P.B. Bochev and M.D. Gunzburger, *Least-squares finite element methods*, Applied Mathematical Sciences, vol. 166, Springer, New York, 2009.
- [10] D. Boffi, F. Brezzi, and M. Fortin, Mixed Finite Element Methods and Applications, Springer Series in Computational Mathematics, vol. 44, Springer, Heidelberg, 2013.
- [11] S.C. Brenner, Poincaré-Friedrichs inequalities for piecewise H¹ functions, SIAM J. Numer. Anal., 41:1 (2003), 306–324.
- [12] S.C. Brenner and L. R. Scott, The Mathematical Theory of Finite Element Methods, third ed., Texts in Applied Mathematics, vol. 15, Springer, New York, 2008.
- [13] F. Brezzi, A. Buffa, and K. Lipnikov, Mimetic finite differences for elliptic problems, M2AN Math. Model. Numer. Anal., 43:2 (2009), 277–295.
- [14] Z. Cai, Least squares for the perturbed Stokes equations and the Reissner-Mindlin plate, SIAM J. Numer. Anal., 38:5 (2000), 1561–1581.
- [15] Z. Cai, C.O. Lee, T.A. Manteuffel, and S. F. McCormick, First-order system least squares for the Stokes and linear elasticity equations: further results, *SIAM J. Sci. Comput.*, **21**:5 (2000), 1728–1739, Iterative methods for solving systems of algebraic equations (Copper Mountain, CO, 1998).

- [16] Z. Cai, T.A. Manteuffel, and S.F. McCormick, First-order system least squares for the Stokes equations, with application to linear elasticity, SIAM J. Numer. Anal., 34:5 (1997), 1727–1741.
- [17] M.I. Gerritsma and M.M.J. Proot, Analysis of a discontinuous least squares spectral element method, Proceedings of the Fifth International Conference on Spectral and High Order Methods (ICOSAHOM-01) (Uppsala), 17 (2002), 297–306.
- [18] V. Girault and P.A. Raviart, *Finite element methods for Navier-Stokes equations*, Springer Series in Computational Mathematics, vol. 5, Springer-Verlag, Berlin, 1986, Theory and algorithms.
- [19] J.J. Heys, E. Lee, T.A. Manteuffel, and S.F. McCormick, On mass-conserving least-squares methods, SIAM J. Sci. Comput., 28:5 (2006), 1675–1693.
- [20] B.N. Jiang and C.L. Chang, Least-squares finite elements for the Stokes problem, Comput. Methods Appl. Mech. Engrg., 78:3 (1990), 297–311.
- [21] O.A. Karakashian and F. Pascal, Convergence of adaptive discontinuous Galerkin approximations of second-order elliptic problems, SIAM J. Numer. Anal., 45:2 (2007), 641–665.
- [22] R.B. Kellogg and J.E. Osborn, A regularity result for the Stokes problem in a convex polygon, J. Functional Analysis, 21:4 (1976), 397–431.
- [23] R. Li, P. Ming, Z. Sun, and Z. Yang, An arbitrary-order discontinuous Galerkin method with one unknown per element, J. Sci. Comput., 80:1 (2019), 268–288.
- [24] R. Li, P. Ming, and F. Tang, An efficient high order heterogeneous multiscale method for elliptic problems, *Multiscale Model. Simul.*, 10:1 (2012), 259–283.
- [25] R. Li, Z. Sun, and F. Yang, Solving eigenvalue problems in a discontinuous approximate space by patch reconstruction, SIAM J. Sci. Comput., 41:5 (2019), A3381–A3400.
- [26] R. Li and F. Yang, A least squares method for linear elasticity using a patch reconstructed space, Comput. Methods Appl. Mech. Engrg., 363:1 (2020), 112902.
- [27] R. Li and F. Yang, A sequential least squares method for Poisson equation using a patch reconstructed space, SIAM J. Numer. Anal., 58:1 (2020), 353–374.
- [28] K. Liu, T.A. Manteuffel, S.F. McCormick, J.W. Ruge, and L. Tang, Hybrid first-order system least squares finite element methods with application to Stokes equations, *SIAM J. Numer. Anal.*, 51:4 (2013), 2214–2237.
- [29] C. Nisters and A. Schwarz, Efficient stress-velocity least-squares finite element formulations for the incompressible Navier-Stokes equations, *Comput. Methods Appl. Mech. Engrg.*, **341** (2018), 333–359.
- [30] M.M.J. Proot and M.I. Gerritsma, Mass- and momentum conservation of the least-squares spectral element method for the Stokes problem, J. Sci. Comput., 27:1–3 (2006), 389–401.
- [31] B.C. Shin and P. Hessari, Least-squares spectral method for velocity-vorticity-pressure form of the Stokes equations, Numer. Methods Partial Differential Equations, 32:2 (2016), 661–680.
- [32] C. Talischi, G.H. Paulino, A. Pereira, and I.F.M. Menezes, PolyMesher: a general-purpose mesh generator for polygonal elements written in Matlab, *Struct. Multidiscip. Optim.*, 45:3 (2012), 309–328.