GENERAL DEGREE DIVERGENCE-FREE FINITE ELEMENT METHODS FOR THE STOKES PROBLEM ON SMOOTH DOMAINS

REBECCA DURST AND MICHAEL NEILAN

Abstract. In this paper, we construct and analyze divergence-free finite element methods for the Stokes problem on smooth domains. The discrete spaces are based on the Scott-Vogelius finite element pair of arbitrary polynomial degree greater than two. By combining the Piola transform with the classical isoparametric framework, and with a judicious choice of degrees of freedom, we prove that the method converges with optimal order in the energy norm. We also show that the discrete velocity error converges with optimal order in the L^2-norm. Numerical experiments are presented, which support the theoretical results.

1. Introduction

Divergence-free methods for the Stokes problem have grown in popularity due to the various advantages they present. This includes pressure-robustness, which allows the errors of the pressure and velocity to be decoupled so that the scheme is well-suited to systems in which the pressure term in the Stokes problem is dominant (i.e. systems with a large pressure gradient or small viscosity). Other advantages include mass-conservation and parameter robustness. Consequently, these methods have become an active area of research (see, e.g., [17, 9, 6, 8, 10]). However, most work on these methods is focused on polyhedral domains. The extension to smooth domains (with optimal-order convergence) is non-trivial and only recently have various approaches been proposed [15, 13, 14].

In this paper, we propose an arbitrary degree, divergence-free, isoparametric finite element scheme based on the Scott-Vogelius pair [17]. On polygonal domains, this approach approximates the velocity with continuous, piecewise polynomials of degree k, and approximates the pressure with discontinuous polynomials of degree $(k-1)$. It is well known that the stability of this pair depends on both the triangulation and the polynomial degree k. We will work on Clough-Tocher splits which yield a stable element pair provided $k \geq 2$. This is a commonly used method allowing greater flexibility with respect to polynomial degree.

In our approach, we combine this Scott-Vogelius pair with an isoparametric paradigm. To do so, we apply k-degree polynomial diffeomorphisms to define the curvilinear triangulation and the finite element spaces. While this approach is classical for isoparametric elements (see [4, 16]), its extension to divergence-free methods is non-standard and a direct application of this approach fails to lead to divergence-free and pressure-robust schemes. In particular, using classical isoparametric Lagrange finite element spaces for velocity approximations disrupts the divergence-free and pressure-robust properties of the scheme. Instead, we employ the divergence-preserving Piola transform in the definition of the discrete velocity space. This transform is defined on the macro (unrefined) triangulation, and we treat the resulting finite element spaces as macro elements defined on the unrefined triangulation.

The primary challenge in this approach lies in the fact that the Piola transform pollutes the continuity of functions in the Lagrange finite element space. More specifically, when the functions in the discrete velocity space are defined by the Piola transform, only normal continuity across interior edges is guaranteed. Thus, the resulting space is only $H(div)$-conforming. Nonetheless, the spaces are

The second author was supported in part by the NSF, grant DMS-2309425.

All data generated or analysed during this study are included in this article.

designed to have weak continuity properties that are leveraged to ensure consistency and stability so that no additional terms in the bilinear form (e.g., penalty terms) are required in the method.

Consequently, one of the main contributions of this paper is to design a finite element space that combines the Lagrange finite element space with the Piola transform and possesses sufficient weak continuity properties across interior edges. We achieve such a space via a judicious choice of edge degrees of freedom; specifically, these are taken as the Gauss-Lobatto points of interior edges. This construction allows us to derive a general estimate of the jumps of discrete velocity functions across interior edges (cf. Lemma 4.7).

This work is an extension of [15] where the lowest-order case $k = 2$ was considered. As expected, some of the results in [15] extend to the general case, such as scaling arguments and inf-sup stability. However, the weak continuity properties of the discrete velocity space is subtle, and a naive extension of [15] to arbitrary polynomial degree does not necessarily lead to an optimal-order convergent method. Another contribution is L^2 error estimates. Again, this requires new estimates of the discrete velocity functions across interior edges.

The organization of the paper is as follows. In the next section, we introduce notation, state the properties of the polynomial diffeomorphism, describe the domain discretization, and introduce the Piola transform. We also establish some necessary preliminary results that are later used in the convergence analysis. In Section 3, we define the local finite element spaces and the degrees of freedom and introduce the global spaces in Section 4. Also in Section 4, we discuss the weak continuity properties of the function spaces and show that the method is inf-sup stable. In Section 5, we introduce the finite element method and derive optimal-order H^1 and L^2 error estimates for the velocity and pressure solutions, respectively. Then, in Section 6, we prove optimal-order convergence in L^2 for the discrete velocity solution, and in Section 7, we provide numerical experiments to verify our theoretical results. Finally, some auxiliary results are proved in Appendices A and B.

2. Preliminaries

Let $\Omega \subset \mathbb{R}^2$ be an open, bounded, and sufficiently smooth domain with boundary $\partial\Omega$. We then construct a mesh following the divergence-free isoparametric method outlined in [15].

2.1. **Isoparametric framework.** We begin with a shape regular, affine (simplicial) triangulation $\tilde{\mathcal{T}}_h$, with sufficiently small mesh size $h = \max_{\tilde{T}\in\tilde{\mathcal{T}}_h} \operatorname{diam}(\tilde{T})$. Furthermore, we assume that the boundary vertices lie on $\partial\Omega$, that $\tilde{\Omega}_h := \operatorname{int}\big(\cup_{\tilde{T}\in\tilde{\mathcal{T}}_h} \bar{\tilde{T}}\big)$ is an $\mathcal{O}(h^2)$ polygonal approximation of Ω, and each $\tilde{T} \in \tilde{\mathcal{T}}_h$ has at most two boundary vertices.

Next, we let $G : \tilde{\Omega}_h \to \Omega$ be a bijective map between the domain and the mesh with $\|G\|_{W^{1,\infty}(\tilde{\Omega}_h)} \le C$. Here and throughout the paper, C denotes a generic positive constant that is independent of any mesh parameter and may take on different values at each occurrence. We define G such that $G|_{\tilde{T}}(x) = x$ at all vertices of $\tilde{T}$. Furthermore, we assume that G is the identity map on interior edges, i.e., edges containing at most one vertex on the boundary.

From here, we define a mesh with curved boundaries following a standard isoparametric framework (see e.g. [4, 2, 12, 5]). In particular, we define G_h to be the piecewise polynomial nodal interpolant of G of degree $\le k$ ($k \ge 2$), with $\|G_h\|_{W^{1,\infty}(\tilde{T})} \le C$ and $\|G_h^{-1}\|_{W^{1,\infty}(\tilde{T})} \le C$ for all $\tilde{T} \in \tilde{\mathcal{T}}_h$. Then, the isoparametric triangulation and computational domain are given by

$$\mathcal{T}_h := \{G_h(\tilde{T}) : \ \tilde{T} \in \tilde{\mathcal{T}}_h\}, \quad \Omega_h := \operatorname{int}\big(\cup_{T\in\mathcal{T}_h} \bar{T}\big).$$

In particular, Ω_h is an $O(h^{k+1})$ approximation to Ω. We denote by $\|\cdot\|_{H_h^m(\Omega_h)}$ the piecewise norm with respect to $\mathcal{T}_h$, i.e.,

$$\|q\|_{H_h^m(\Omega_h)}^2 = \sum_{T\in\mathcal{T}_h} \|q\|_{H^m(T)}^2.$$

We also denote by ∇_h the piecewise gradient operator with respect to $\mathcal{T}_h$, so that $\nabla_h q|_T = \nabla(q|_T)$ for all $T \in \mathcal{T}_h$.

2.2. **The mappings $F_{\tilde{T}}$ and F_T.** To define the finite element spaces, we must first construct mappings to and from the affine and curved triangulations, $\tilde{\mathcal{T}}_h$ and $\mathcal{T}_h$. To do so, we define $\hat{T}$ to be the reference triangle with vertices $(1,0)$, $(0,1)$, and $(0,0)$. Then for each $\tilde{T} \in \tilde{\mathcal{T}}_h$, we let $F_{\tilde{T}} : \hat{T} \to \tilde{T}$ be an affine bijection with $|F_{\tilde{T}}|_{W^{1,\infty}(\hat{T})} \le Ch_T$ and $|F_{\tilde{T}}^{-1}|_{W^{1,\infty}(\tilde{T})} \le Ch_{\tilde{T}}^{-1}$ for $h_{\tilde{T}} = \mathrm{diam}(\tilde{T})$. Subsequently, we may define $F_T : \hat{T} \to T$ by $F_T = G_h \circ F_{\tilde{T}}$. For each $T \in \mathcal{T}_h$, the polynomial diffeomorphism F_T and its inverse and assumed to satisfy the following estimates:

$$
\begin{aligned}
&|F_T|_{W^{m,\infty}(\hat{T})} \le Ch_T^m \quad (0 \le m \le k), \qquad |F_T^{-1}|_{W^{m,\infty}} \le Ch_T^{-m} \quad (0 \le m \le (k+1)),\\
&\qquad c_1 h_T^2 \le \det(DF_T) \le c_2 h_T^2.
\end{aligned} \tag{2.1}
$$

Here, we have $h_T = \mathrm{diam}(G_h^{-1}(T))$, and c_1, c_2 are generic constants independent of h_T. Furthermore, we note that, due to the assumptions on G, the mappings F_T and $F_{\tilde{T}}$ are oriented so that they match at the vertices of $\hat{T}$. Consequently, the mappings are the same on triangles with three interior edges, so that for all such triangles $T \in \mathcal{T}_h$ we have $T = G_h(\tilde{T}) = \tilde{T}$.

2.3. **The boundary regions of Ω and Ω_h.** With the isoparametric triangulation established, we define $\Omega\Delta\Omega_h = (\Omega \setminus \Omega_h) \cup (\Omega_h \setminus \Omega)$ and note it may be shown that (see e.g. [3, Equation 3.9] for proof)

$$
|\Omega\Delta\Omega_h| \le Ch^{k+1}. \tag{2.2}
$$

Next, by the construction of $\Omega\Delta\Omega_h$, we have a bound of the H^1 semi-norm in this boundary region.

Lemma 2.1. *Let $\boldsymbol{v} \in \boldsymbol{H}^2(\Omega)$ be extended into $\mathbb{R}^2$ in a way such that $\|\boldsymbol{v}\|_{H^2(\mathbb{R}^2)} \le C\|\boldsymbol{v}\|_{H^2(\Omega)}$. Then for h sufficiently small,*

$$
\|\nabla \boldsymbol{v}\|_{L^2(\Omega\triangle\Omega_h)} \le Ch^{\frac{k+1}{2}}\|\boldsymbol{v}\|_{H^2(\Omega)}.
$$

Proof. Let d be the signed distance function of Ω with the convention $d(x) < 0$ for $x \in \Omega$. For $\delta > 0$, define

$$
U_\delta := \{x \in \mathbb{R}^2 : \ d(x) < \delta\},
$$

and note that, because $\partial\Omega$ is C^2, there holds $\partial U_\delta \in C^2$ for $\delta > 0$ sufficiently small. We then set

$$
\mathcal{N}_\delta := \{x \in U : |d(x)| < \delta\}
$$

to be the tubular region around ∂U_δ. By [7, Lemma 4.10], there holds for all $w \in H^1(U_\delta)$:

$$
\|w\|_{L^2(\mathcal{N}_\delta)} \le C\delta^{1/2}\|w\|_{H^1(U_\delta)}.
$$

Now set $\delta_h = 2\mathrm{dist}\{\partial\Omega_h, \partial\Omega\} = \mathcal{O}(h^{k+1})$, so that $\Omega\triangle\Omega_h \subset \mathcal{N}_{\delta_h}$, and assume h is sufficiently small such that $\partial U_{\delta_h} \in C^2$. We then have

$$
\|\nabla \boldsymbol{v}\|_{L^2(\Omega\triangle\Omega_h)} \le C\|\nabla \boldsymbol{v}\|_{L^2(\mathcal{N}_{\delta_h})} \le C\delta_h^{1/2}\|\nabla \boldsymbol{v}\|_{H^1(U_{\delta_h})} \le Ch^{\frac{k+1}{2}}\|\boldsymbol{v}\|_{H^2(\Omega)}.
$$

□

2.4. **Clough-Tocher Split.** To guarantee inf-sup stability of the proposed divergence-free method, we introduce on each element a local triangulation given by the Clough-Tocher split. Let $\hat{T}^{ct} = \{\hat{K}_i\}_{i=1}^3$ be the Clough-Tocher triangulation of the reference triangle, obtained by connecting the vertices of $\hat{T}$ to its barycenter. We define analogous splits on our affine and curved triangulations via $F_{\tilde{T}}$ and F_T (cf. Figures 1–2):

$$
\tilde{T}^{ct} = \{F_{\tilde{T}}(\hat{K}) : \ \hat{K} \in \hat{T}^{ct}\}, \quad T^{ct} = \{F_T(\hat{K}) : \ \hat{K} \in \hat{T}^{ct}\}.
$$

From (2.1) and the shape-regularity of $\tilde{T}_h$, it follows that $|T| \le C|K|$ for all $K \in T^{ct}$.

Remark 2.2. We note that for the *macroelement* $T \in \mathcal{T}_h$, only edges containing both vertices on $\partial\Omega_h$ may be curved. However, it may be that interior edges of the local triangulations $K \in T^{ct}$ may indeed be curved as well.

2.5. **The Piola transform.** The final piece we need to construct the divergence-free method is the Piola transform. Given $T \in \mathcal{T}_h$, we define the matrix $A_T : \hat{T} \to \mathbb{R}^{2\times 2}$ to be the matrix arising in this transform

$$A_T(\hat{x}) := \frac{DF_T(\hat{x})}{\det(DF_T(\hat{x}))}. \tag{2.3}$$

In what follows, the local function spaces on each $T \in \mathcal{T}_h$ will be constructed through A_T. Specifically, given a function $\hat{\boldsymbol{v}} : \hat{T} \to \mathbb{R}^2$, its Piola transform yields the function $\boldsymbol{v} : T \to \mathbb{R}^2$ with $\boldsymbol{v} = (A_T \hat{\boldsymbol{v}}) \circ F_T^{-1}$ (cf. Section 3). It is well-known that this transform is divergence-preserving and normal-continuity preserving, and its use in the definition of the spaces given below allows us to maintain these properties of the Scott-Vogelius pair on curved triangulations. We emphasize that this transform is defined with respect to $T \in \mathcal{T}_h$, not with respect to the triangles in the Clough-Tocher split.

2.6. **Bounds and scaling results.** The following results give bounds on the matrix A_T and its inverse. We refer to the appendix of [15] for a proof for the case $k = 2$. The arguments given there generalize trivially for $k \geq 2$, and therefore the proof of the following lemma is omitted.

Lemma 2.3. *For each $T \in \mathcal{T}_h$, there holds*

$$|A_T|_{W^{m,\infty}(\hat{T})} \leq C h_T^{m-1} \ (m \geq 0), \qquad |A_T^{-1}|_{W^{m,\infty}(\hat{T})} \leq \begin{cases} C h_T^{1+m} & 0 \leq m \leq k-1, \\ 0 & k \leq m. \end{cases} \tag{2.4}$$

Additionally, we will make use of the following scaling results from [2].

Lemma 2.4. *Let $T \in \mathcal{T}_h$ and $\boldsymbol{w} \in \boldsymbol{W}^{m,p}(T)$ with $m \geq 0$ and $p \in [1, \infty]$. Let $\hat{\boldsymbol{w}} \in \boldsymbol{W}^{m,p}(\hat{T})$ be the image of $\boldsymbol{w}$ on $\hat{T}$ with $\hat{\boldsymbol{w}}(\hat{x}) = \boldsymbol{w}(x)$, $x = F_T(\hat{x})$ and set $\hat{K} = F_T^{-1}(K)$ for each $K \in T^{ct}$. Then for any $K \in T^{ct}$,*

$$\begin{aligned} |\boldsymbol{w}|_{W^{m,p}(K)} &\leq C h_T^{2/p-m} \sum_{r=0}^{m} h_T^{2(m-r)} |\hat{\boldsymbol{w}}|_{W^{r,p}(\hat{K})}, \\ |\hat{\boldsymbol{w}}|_{W^{m,p}(\hat{K})} &\leq C h_T^{m-2/p} \sum_{r=0}^{m} |\boldsymbol{w}|_{W^{r,p}(K)}. \end{aligned} \tag{2.5}$$

In the results that follow, we let $\boldsymbol{n}$ denote the outward unit normal of a given domain (understood from context), and set $\boldsymbol{t}$ to be the unit tangent vector obtained by rotating $\boldsymbol{n}$ 90 degrees counterclockwise.

3. Local spaces

The full derivation of the local spaces for the divergence-free isoparametric framework can be found in [15] for the case $k = 2$. Below, we have the analogous results for general k.

To begin, we define the local function space on the reference triangle $\hat{T} \subset \mathbb{R}^2$, with Clough-Tocher triangulation $\hat{T}^{ct} = \{\hat{K}_1, \hat{K}_2, \hat{K}_3\}$. Without including boundary conditions, the polynomial spaces on the reference triangle are

$$\begin{aligned} \hat{\boldsymbol{V}}_k =& \{\hat{\boldsymbol{v}} \in \boldsymbol{H}^1(\hat{T}) : \hat{\boldsymbol{v}}|_{\hat{K}} \in \boldsymbol{\mathcal{P}}_k(\hat{K}) \ \forall \hat{K} \in \hat{T}^{ct}\}, \\ \hat{Q}_{k-1} =& \{\hat{q} \in L^2(\hat{T}) : \hat{q}|_{\hat{K}} \in \mathcal{P}_{k-1}(\hat{K}) \ \forall \hat{K} \in \hat{T}^{ct}\}, \end{aligned}$$

where $\mathcal{P}_k(S)$ is the space of scalar polynomials of degree $\leq k$ on domain S, and $\boldsymbol{\mathcal{P}}_k(S) = [\mathcal{P}_k(S)]^2$.

With $\tilde{x} = F_{\tilde{T}}(\hat{x})$, we define the local spaces on the affine triangle $\tilde{T} \in \tilde{\mathcal{T}}_h$ via composition:

$$\begin{aligned} \tilde{\boldsymbol{V}}_k(\tilde{T}) =& \{\tilde{\boldsymbol{v}} \in \boldsymbol{H}^1(\tilde{T}) : \tilde{\boldsymbol{v}}(\tilde{x}) = \hat{\boldsymbol{v}}(\hat{x}), \ \exists \hat{\boldsymbol{v}} \in \hat{\boldsymbol{V}}_k\}, \\ \tilde{Q}_{k-1}(\tilde{T}) =& \{\tilde{q} \in L^2(\tilde{T}) : \tilde{q}(\tilde{x}) = \hat{q}(\hat{x}), \ \exists \hat{q} \in \hat{Q}_{k-1}\}. \end{aligned}$$

To incorporate boundary conditions, we further define

$$\begin{aligned}\hat{\boldsymbol{V}}_{k,0} &= \hat{\boldsymbol{V}}_k \cap \boldsymbol{H}_0^1(\hat{T}), & \hat{Q}_{k-1,0} &= \hat{Q} \cap L_0^2(\hat{T}),\\ \tilde{\boldsymbol{V}}_{k,0}(\tilde{T}) &= \tilde{\boldsymbol{V}}_k(\tilde{T}) \cap \boldsymbol{H}_0^1(\tilde{T}), & \tilde{Q}_{k-1,0}(\tilde{T}) &= \tilde{Q}(\tilde{T}) \cap L_0^2(\tilde{T}),\end{aligned}$$

where $L_0^2(\tilde{T})$ is the space of $L^2(\tilde{T})$-functions with vanishing mean.

We then define the function spaces on the triangles $T \in \mathcal{T}_h$ (which may have a curved edge), using the notation $x = F_T(\hat{x})$ and the Piola transform:

$$\begin{aligned}\boldsymbol{V}_k(T) =& \{\boldsymbol{v} \in \boldsymbol{H}^1(T) : \boldsymbol{v}(x) = A_T(\hat{x})\hat{\boldsymbol{v}}(\hat{x}),\ \exists \hat{\boldsymbol{v}} \in \hat{\boldsymbol{V}}_k\}, \quad \boldsymbol{V}_{k,0}(T) = \boldsymbol{V}_k(T) \cap \boldsymbol{H}_0^1(T),\\ Q_{k-1}(T) =& \{q \in L^2(T) : q(x) = \hat{q}(\hat{x}),\ \exists \hat{q} \in \hat{Q}_{k-1}\},\\ Q_{k-1,0}(T) =& \{q \in L^2(T) : q(x) = \hat{q}(\hat{x}),\ \exists \hat{q} \in \hat{Q}_{k-1,0}\},\end{aligned}$$

where the matrix A_T is given by (2.3). Note that if F_T is affine, then $\boldsymbol{V}(T) = \tilde{\boldsymbol{V}}(\tilde{T})$ and $Q(T) = \tilde{Q}(\tilde{T})$.

It is important to note that functions in $\boldsymbol{V}_{k-1}(T)$ and $Q_{k-1}(T)$ are not necessarily piecewise-polynomial spaces if T is not affine. In addition, on curved triangles the matrix A_T is not necessarily constant on straight edges, and therefore functions in $\boldsymbol{V}_{k-1}(T)$ are not necessarily polynomials on such edges. However, the following lemma shows that the the normal component of $\boldsymbol{v}$ will be a polynomial when restricted to a straight edge.

Lemma 3.1. *Let $\boldsymbol{v} \in \boldsymbol{V}_k(T)$, and suppose that e is a straight edge of ∂T with unit normal $\boldsymbol{n}$. Then $\boldsymbol{v} \cdot \boldsymbol{n}|_e \in \mathcal{P}_k(e)$.*

The proof of this result is found in [15, Lemma 3.1] for the case $k = 2$ and essentially uses the well-known normal-preserving properties of the Piola transform. However, the result extends trivially to $\hat{\boldsymbol{v}} \cdot \hat{\boldsymbol{n}}$ a polynomial of arbitrary degree k.

With the local spaces now defined, we may state the following lemma showing that functions in the finite element space $\boldsymbol{V}_k(T)$ enjoy inverse estimates similar to those for piecewise polynomials. In addition, similar to isoparametric (polynomial) elements defined via composition, high-order Sobolev norms of functions in $\boldsymbol{V}_k(T)$ are controlled by kth-order Sobolev norms. Its proof is based on scaling arguments and is found in Appendix A.

Lemma 3.2. *Let $p, q \in [1, \infty]$ and $0 \le m \le \ell$ be integers. Then for any $T \in \mathcal{T}_h$ and $\boldsymbol{v} \in \boldsymbol{V}_k(T)$,*

$$\|\boldsymbol{v}\|_{W^{\ell,p}(K)} \le C h_T^{m-\ell+2(\frac{1}{p}-\frac{1}{q})} \|\boldsymbol{v}\|_{W^{m,q}(K)} \qquad \forall K \in T^{ct}. \tag{3.1}$$

Moreover,

$$\|\boldsymbol{v}\|_{W^{\ell,p}(K)} \le C \|\boldsymbol{v}\|_{W^{k,p}(K)} \qquad \forall \ell \ge k,\ \forall K \in T^{ct}. \tag{3.2}$$

3.1. **Degrees of Freedom on $\boldsymbol{V}(T)$.** To describe the degrees of freedom of the local velocity space $\boldsymbol{V}(T)$, we first summarize the canonical degrees of freedom for the reference space $\hat{\boldsymbol{V}}_k$, i.e., the kth-degree Lagrange finite element space defined on the Clough-Tocher split. It is well known that a function in this space is uniquely determined by (i) its values at the four vertices in $\hat{T}^{ct}$ (4 nodes); (ii) its values at $(k-1)$ distinct points for each of the six (open) edges in $\hat{T}^{ct}$ ($6(k-1)$ nodes); and (iii) its values at $\frac{1}{2}(k-1)(k-2)$ distinct points for each of the three (open) subtriangles in $\hat{T}^{ct}$ ($\frac{3}{2}(k-1)(k-2)$ nodes). In (iii), the $\frac{1}{2}(k-1)(k-2)$ points for each subtriangle must be chosen such that they uniquely determine a polynomial of degree $(k-3)$. We see that the total number of nodes is $M_k := 4 + 6(k-1) + \frac{3}{2}(k-1)(k-2) = \frac{3}{2}k(k+1) + 1$, and therefore the dimension of $\hat{\boldsymbol{V}}_k$ is $3k(k+1)+2$. By setting $\hat{\mathcal{N}}_k = \{\hat{a}_i\}_{i=1}^{M_k}$ to be the set of these points, then a function $\hat{\boldsymbol{v}} \in \hat{\boldsymbol{V}}_k$ is uniquely determined by the values $\hat{\boldsymbol{v}}(\hat{a}_i)$ for all $a_i \in \hat{\mathcal{N}}_k$.

To ensure sufficient weak continuity properties of the global finite element spaces defined below, we specify that the location of the points on the three boundary edges of $\hat{T}$ correspond to the nodes of the Gauss-Lobatto quadrature scheme. In particular, for a boundary edge $\hat{e} \subset \partial\hat{T}$, the nodes on the closure of the edge, denoted by $\{\hat{m}_i\}_{i=1}^{k+1} \subset \bar{\hat{e}}$, satisfy

$$\sum_{i=1}^{k+1} \hat{\omega}_i \hat{q}(\hat{m}_i) = \int_{\hat{e}} \hat{q} \qquad \forall \hat{q} \in \mathcal{P}_{2k-1}(\hat{e}),$$

and two of the nodes in this set correspond to the vertices of $\hat{e}$. The other nodes (i.e., nodes on interior edges and the interior of subtriangles) can be chosen such that they satisfy the above properties to form a unisolvent set of degrees of freedom on $\hat{\boldsymbol{V}}_k$. However, to simplify the implementation of the resulting finite element spaces, we also take nodes on the interior edges to be the Gauss-Lobttto points, and let the nodes in the interior of subtriangles to be the canonical Lagrange nodes (cf. Figures 1 and 2).

We map these nodes to $\tilde{T} \in \tilde{\mathcal{T}}_h$ and $T \in \mathcal{T}_h$ via the mappings $F_{\tilde{T}}$ and F_T, respectively:

$$\mathcal{N}_k(\tilde{T}) = \{F_{\tilde{T}}(\hat{a}_i) : \ \hat{a}_i \in \hat{\mathcal{N}}_k\}, \quad \mathcal{N}_k(T) = \{F_T(\hat{a}_i) : \ \hat{a}_i \in \hat{\mathcal{N}}_k\}.$$

Due to the invariance of spaces of polynomials under affine transformations, we see that the nodes in $\mathcal{N}_k(\tilde{T})$ that lie on an edge $\tilde{e} \subset \bar{\tilde{T}}$ correspond to the Gauss-Lobatto quadrature rule on that edge. Likewise, if $e \subset \partial T$ is a straight edge of T, then the nodes in $\mathcal{N}_k(T)$ that lie on $\bar{e}$ are the nodes of the Gauss-Lobatto quadrature rule on e.

Finally, we note that, since functions in $\hat{\boldsymbol{V}}_k$ are uniquely determined by their values at the nodes $\hat{\mathcal{N}}_k$, and since the matrix A_T is invertible, it follows that any $\boldsymbol{v} \in \boldsymbol{V}_k(T)$ is uniquely determined by the values $\boldsymbol{v}(a_i)$, $a_i \in \mathcal{N}_k(T)$.

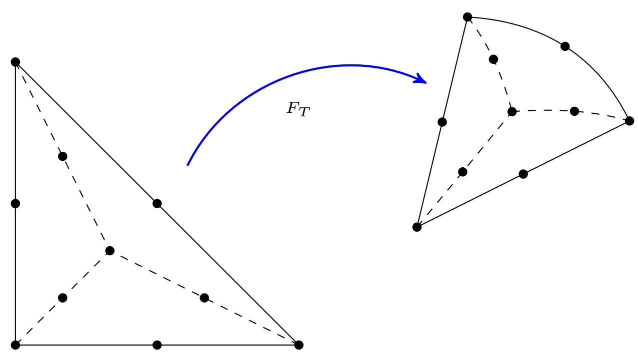

FIGURE 1. Clough-Tocher split and degrees of freedom for the quadratic Lagrange finite element space ($k = 2$). The local split is mapping to the curved element T via the polynomial diffeomorphism F_T.

4. Global Spaces

On the affine triangulation $\tilde{\mathcal{T}}_h$, we define the Scott-Vogelius pair

$$\begin{aligned}\tilde{\boldsymbol{V}}_k^h =& \{\tilde{\boldsymbol{v}} \in \boldsymbol{H}_0^1(\tilde{\Omega}_h) : \tilde{\boldsymbol{v}}|_{\tilde{T}} \in \tilde{\boldsymbol{V}}_k(\tilde{T}) \ \forall \tilde{T} \in \tilde{\mathcal{T}}_h\},\\ \tilde{Q}_{k-1}^h =& \{\tilde{q} \in L_0^2(\tilde{\Omega}_h) : \tilde{q}|_{\tilde{T}} \in \tilde{Q}_{k-1}(\tilde{T}) \ \forall \tilde{T} \in \tilde{\mathcal{T}}_h\}.\end{aligned}$$

We see that $\tilde{\boldsymbol{V}}_k^h$ is the kth degree Lagrange finite element space with respect to the Clough-Tocher refinement of $\tilde{\mathcal{T}}_h$, and $\tilde{Q}_{k-1}^h$ is the space of discontinuous polynomials of degree $(k-1)$, again with respect to the Clough-Tocher refinement. The finite elements $\tilde{\boldsymbol{V}}_h \times \tilde{Q}_h$ represents a stable and divergence–free Stokes pair [1], however its use formally leads to a suboptimal scheme on smooth domains due to geometric error.

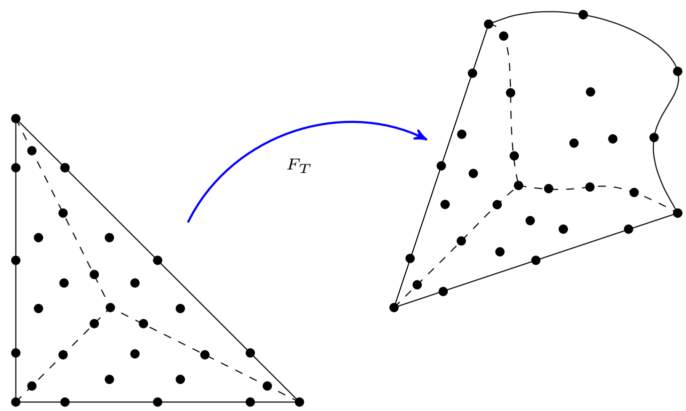

FIGURE 2. Clough-Tocher split and degrees of freedom for the quartic Lagrange finite element space ($k = 4$). Edge degrees of freedom on $\hat{T}$ are placed at Gauss-Lobatto points.

To define the isoparametric spaces, we define the operators $\boldsymbol{\Psi}_k$ and Υ_{k-1} such that $\boldsymbol{\Psi}_k|_T : \tilde{\boldsymbol{V}}(\tilde{T}) \to \boldsymbol{V}(T)$ and $\Upsilon_{k-1}|_T : \tilde{Q}(\tilde{T}) \to Q(T)$ and are uniquely determined on each $\tilde{T} \in \tilde{\mathcal{T}}_h$ and $T \in \mathcal{T}_h$ with $T = G_h(\tilde{T})$ as follows:

1. $(\boldsymbol{\Psi}_k|_T\tilde{\boldsymbol{v}})(a) = \tilde{\boldsymbol{v}}(\tilde{a})\ \forall \tilde{a} \in \mathcal{N}_k(\tilde{T})$, with $a = G_h(\tilde{a}) \in \mathcal{N}_k(T)$, and
2. $(\Upsilon_{k-1}|_T\tilde{q}) = \tilde{q} \circ G_h^{-1}$

Thus, $\boldsymbol{\Psi}_k$ maps functions in $\tilde{\boldsymbol{V}}_k^h$ to the isoparametric domain Ω_h via the Piola transform and interpolation, and Υ_{k-1} maps functions in $\tilde{Q}_{k-1}^h$ to Ω_h via composition.

In the following proposition, we state some properties of the mapping $\boldsymbol{\Psi}_k$ without proof, as the result is proven for $k = 2$ in [15, Theorem 3.7]. To extend the results to arbitrary degree k, one only needs to recognize that a k-th degree polynomial along an edge $e \subset \partial T$ is uniquely determined by its values at the $k + 1$ nodal points that lie on this edge.

Proposition 4.1. *The following properties are satisfied:*

1. *If F_T is affine, then $\boldsymbol{\Psi}_k|_T$ is the identity operator.*
2. *If $e \subset \partial T$ is a straight edge (so $e \subset \partial\tilde{T}$ with $T = G_h(\tilde{T})$), then*
$$(\boldsymbol{\Psi}_k|_T\tilde{\boldsymbol{v}}) \cdot \boldsymbol{n}|_e = \tilde{\boldsymbol{v}} \cdot \boldsymbol{n}|_e \qquad \forall \tilde{\boldsymbol{v}} \in \tilde{\boldsymbol{V}}_k^h.$$
3. *There holds $\|\boldsymbol{\Psi}_k|_T\tilde{\boldsymbol{v}}\|_{H^1(T)} \le C\|\tilde{\boldsymbol{v}}\|_{H^1(\tilde{T})}$ for all $\tilde{\boldsymbol{v}} \in \tilde{\boldsymbol{V}}_k^h$.*

Consequently, global function spaces defined on the isoparametric mesh $\mathcal{T}_h$ are given by
$$\boldsymbol{V}_k^h := \{\boldsymbol{v} : \boldsymbol{v} = \boldsymbol{\Psi}_k\tilde{\boldsymbol{v}},\ \exists \tilde{\boldsymbol{v}} \in \tilde{\boldsymbol{V}}_k^h\}, \quad Q_k^h := \{q : q = \Upsilon_{k-1}\tilde{q},\ \exists \tilde{q} \in \tilde{Q}_k^h\}.$$

Remark 4.2. From the boundary conditions applied to the space $\tilde{\boldsymbol{V}}_k^h$ and the definition of $\boldsymbol{\Psi}_k$, we see that functions in $\boldsymbol{V}_k^h$ are continuous at the degrees of freedom, and vanish on $\partial\Omega_h$.

With these spaces defined, we have the following results.

Lemma 4.3. *There holds $\boldsymbol{V}_k^h \subset \boldsymbol{H}_0(\mathrm{div};\Omega_h) = \{\boldsymbol{v} \in \boldsymbol{L}^2(\Omega_h) :\ \mathrm{div}\,\boldsymbol{v} \in L^2(\Omega_h),\ \boldsymbol{v} \cdot \boldsymbol{n}|_{\partial\Omega_h} = 0\}$.*

This result follows immediately from the construction of $\boldsymbol{V}_k^h$ and the continuity of the normal component on interior edges imposed by part 2 of Proposition 4.1. See [15, Theorem 4.2] for details.

Lemma 4.4. *There exists an operator $\boldsymbol{I}_k^h : \boldsymbol{H}^2(\Omega) \cap \boldsymbol{H}_0^1(\Omega_h) \to \boldsymbol{V}_k^h$ such that for $\boldsymbol{u} \in \boldsymbol{H}^s(\Omega_h) \cap \boldsymbol{H}_0^1(\Omega_h)$ $(s \ge 2)$ and for each $T \in \mathcal{T}_h$, there holds*
$$\|\boldsymbol{u} - \boldsymbol{I}_k^h\boldsymbol{u}\|_{H^m(T)} \le Ch_T^{\ell-m}\|\boldsymbol{u}\|_{H^\ell(T)} \quad 0 \le m \le \ell := \min\{k+1, s\}. \tag{4.1}$$

Proof. Recall $\mathcal{N}_k(T)$ and $\hat{\mathcal{N}}_k$ are the sets of nodes on T and $\hat{T}$, respectively. We uniquely define the operator $\boldsymbol{I}_k^h$ such that on each $T \in \mathcal{T}_h$,

$$(\boldsymbol{I}_k^h \boldsymbol{u})|_T(a) = \boldsymbol{u}(a) \quad \forall a \in \mathcal{N}_k(T).$$

Set $\boldsymbol{v} = \boldsymbol{I}_k^h \boldsymbol{u}|_T \in \boldsymbol{V}_k(T)$, for $T \in \mathcal{T}_h$. Then set $\hat{\boldsymbol{v}} \in \hat{\boldsymbol{V}}_k$ and $\hat{\boldsymbol{u}} \in \boldsymbol{H}^s(\hat{T})$ such that

$$\boldsymbol{v}(x) = (A_T \hat{\boldsymbol{v}})(\hat{x}), \quad \boldsymbol{u}(x) = (A_T \hat{\boldsymbol{u}})(\hat{x}).$$

Consequently,

$$(A_T \hat{\boldsymbol{v}})(\hat{a}) = (A_T \hat{\boldsymbol{u}})(\hat{a}) \qquad \Longrightarrow \qquad \hat{\boldsymbol{v}}(\hat{a}) = \hat{\boldsymbol{u}}(\hat{a}) \quad \forall \hat{a} \in \hat{\mathcal{N}}_k$$

because the matrix A_T is invertible. Thus, $\hat{\boldsymbol{v}}$ is the kth degree nodal interpolant of $\hat{\boldsymbol{u}}$ with respect to $\hat{T}^{ct}$, so, by standard interpolation theory, we have

$$\|\hat{\boldsymbol{u}} - \hat{\boldsymbol{v}}\|_{H^m(\hat{T})} \le C|\hat{\boldsymbol{u}}|_{H^\ell(\hat{T})} \quad 0 \le m \le \ell = \min\{k+1, s\}. \tag{4.2}$$

Thus it follows from (4.2), Lemmas 2.3 and 2.4, and an application of the product rule that

$$|\boldsymbol{u} - \boldsymbol{v}|_{H^m(T)} \le Ch_T^{1-m} \|A_T\|_{W^{m,\infty}(\hat{T})} \|\hat{\boldsymbol{u}} - \hat{\boldsymbol{v}}\|_{H^m(\hat{T})} \le Ch_T^{-m} |\hat{\boldsymbol{u}}|_{H^\ell(\hat{T})},$$

and so, by using Lemmas 2.3 and 2.4 once again,

$$\begin{aligned}
|\boldsymbol{u} - \boldsymbol{v}|_{H^m(T)} &\le Ch_T^{-m} |A_T^{-1} A_T \hat{\boldsymbol{u}}|_{H^\ell(\hat{T})} \\
&\le Ch_T^{-m} \sum_{j=0}^{\ell} |A_T^{-1}|_{W^{j,\infty}(\hat{T})} |A_T \hat{\boldsymbol{u}}|_{H^{\ell-j}(\hat{T})} \\
&\le Ch_T^{-m} \sum_{j=0}^{\ell} h_T^{1+j} |A_T \hat{\boldsymbol{u}}|_{H^{\ell-j}(\hat{T})} \\
&\le Ch_T^{\ell-m} \|\boldsymbol{u}\|_{H^\ell(T)}.
\end{aligned}$$

□

4.1. **Weak continuity properties.** The next result shows that, while functions in $\boldsymbol{V}_k^h$ are only $\boldsymbol{H}_0(\mathrm{div};\Omega_h)$-conforming, they do have weak continuity properties across interior edges of the mesh. In particular, they are "close" to an $\boldsymbol{H}_0^1(\Omega_h)$-conforming relative. The lemma is a generalization of [15, Lemma 4.5] to general polynomial degree and to higher-order Sobolev norms; its proof is given in Appendix B.

Lemma 4.5. *There exists an operator $\boldsymbol{E}_h : \boldsymbol{V}_k^h \to \boldsymbol{H}_0^1(\Omega_h)$ such that for all $\boldsymbol{v} \in \boldsymbol{V}_k^h$*

$$\|\boldsymbol{v} - \boldsymbol{E}_h \boldsymbol{v}\|_{L^2(T)} + h_T \|\nabla(\boldsymbol{v} - \boldsymbol{E}_h \boldsymbol{v})\|_{L^2(T)} \le Ch_T^{m+1} \|\boldsymbol{v}\|_{H^m(T)} \qquad \forall T \in \mathcal{T}_h, \tag{4.3}$$

for $m = 0, 1, \ldots, k$. Moreover, $\boldsymbol{E}_h \boldsymbol{v}|_T = \boldsymbol{v}$ if T is affine.

Corollary 4.6. *For $\boldsymbol{v} \in \boldsymbol{V}^h$, it holds*

$$\|\boldsymbol{v}\|_{L^2(\Omega_h)} \le C\|\nabla_h \boldsymbol{v}\|_{L^2(\Omega_h)},$$

where ∇_h denotes the piecewise gradient operator with respect to $\mathcal{T}_h$, and $C > 0$ is a constant depending only on the size of Ω_h and the shape regularity of $\mathcal{T}_h$.

Proof. Recall that $\boldsymbol{v} = \boldsymbol{E}_h \boldsymbol{v}$ on affine triangles, so $\|\boldsymbol{v} - \boldsymbol{E}_h\|_{L^2(T)}$ may only be nonzero on curved $T \in \mathcal{T}_h$, all of which will have at least two vertices on the boundary. We denote the set of triangles with two boundary vertices as $\mathcal{T}_h^\partial$ so that $\boldsymbol{v}|_T = \boldsymbol{E}_h \boldsymbol{v}|_T$ for $T \in \mathcal{T}_h \backslash \mathcal{T}_h^\partial$. Because $\boldsymbol{v}|_{\partial\Omega_h} = 0$, we have $\|\boldsymbol{v}\|_{L^2(T)} \le Ch_T \|\nabla \boldsymbol{v}\|_{L^2(T)}$ for $T \in \mathcal{T}_h^\partial$.

Thus, recalling that $\boldsymbol{E}_h\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega_h)$, we may apply the triangle inequality, Lemma 4.5 (twice with $m = 0$), and the Poincaré inequality (twice) to determine

$$\begin{aligned}
\|\boldsymbol{v}\|_{L^2(\Omega_h)}^2 \leq & 2\left(\|\boldsymbol{E}_h\boldsymbol{v}\|_{L^2(\Omega_h)}^2 + \sum_{T\in\mathcal{T}_h^\partial} \|\boldsymbol{v} - \boldsymbol{E}_h\boldsymbol{v}\|_{L^2(T)}^2\right)\\
\leq & C\left(\|\nabla\boldsymbol{E}_h\boldsymbol{v}\|_{L^2(\Omega_h)}^2 + \sum_{T\in\mathcal{T}_h^\partial} h_T^2\|\nabla\boldsymbol{v}\|_{L^2(T)}^2\right)\\
\leq & C\left(\|\nabla_h\boldsymbol{v}\|_{L^2(\Omega_h)}^2 + \sum_{T\in\mathcal{T}_h^\partial} \|\nabla(\boldsymbol{v} - \boldsymbol{E}_h\boldsymbol{v})\|_{L^2(T)}^2\right)\\
\leq & C\|\nabla_h\boldsymbol{v}\|_{L^2(\Omega_h)}^2.
\end{aligned}$$

□

Using the H^1-conforming relative in Lemma 4.5 and the fact that the Lagrange DOFs are Gauss-Lobatto nodes, we show that functions in $\boldsymbol{V}_k^h$ possess weak continuity properties across interior edges. To describe the result, we set $\mathcal{E}_h^I$ to denote the set of interior edges of $\mathcal{T}_h$, and define the jump of a vector-valued function across an edge $e = \partial T_+ \cap \partial T_- \in \mathcal{E}_h^I$ ($T_\pm \in \mathcal{T}_h$) as

$$[\boldsymbol{v}]|_e = \boldsymbol{v}_+ \otimes \boldsymbol{n}_+|_e + \boldsymbol{v}_- \otimes \boldsymbol{n}_-|_e,$$

where $\boldsymbol{v}_\pm = \boldsymbol{v}|_{T_\pm}$ and $\boldsymbol{n}_\pm$ is the outward unit normal of $\partial T_\pm$ restricted to e.

Lemma 4.7. *Let $\boldsymbol{w} \in \boldsymbol{H}^s(\Omega)$ with $s \geq 2$, and set $r = \min\{s-1, k-1\}$. We extend $\boldsymbol{w}$ to $\mathbb{R}^2$ such that $\|\boldsymbol{w}\|_{H^{r+1}(\mathbb{R}^2)} \leq C\|\boldsymbol{w}\|_{H^{r+1}(\Omega)}$. Then there holds for all $\boldsymbol{v} \in \boldsymbol{V}_k^h$, and $m = 0, 1, \ldots, k$,*

$$\left|\sum_{e\in\mathcal{E}_h^I}\int_e \nabla\boldsymbol{w} : [\boldsymbol{v}]\right| \leq Ch^{r+m}\|\boldsymbol{w}\|_{H^{r+1}(\Omega)}\|\boldsymbol{v}\|_{H_h^m(\Omega_h)}. \tag{4.4}$$

Proof. For $e \in \mathcal{E}_h^I$, let $T_+, T_- \in \mathcal{T}_h$ such that $e = \partial T_+ \cap \partial T_-$. We let $G_e \in [H^1(T_+ \cup T_-)]^{2\times 2}$ such that $G_e|_{T_\pm} \circ F_{T_\pm} \in [\mathcal{P}_{k-2}(\hat{T})]^{2\times 2}$ and

$$\int_{T_+\cup T_-} G_e : Q = \int_{T_+\cup T_-} \nabla\boldsymbol{w} : Q$$

for all $Q \in [H^1(T_+\cup T_-)]^{2\times 2}$ with $Q|_{T_\pm} \circ F_{T_\pm} \in [\mathcal{P}_{k-2}(\hat{T})]^{2\times 2}$. That is G_e is the $L^2(T_+\cup T_-)$ projection of $\nabla\boldsymbol{w}$ with respect to the local $(k-2)$-degree Lagrange (isoparametric) finite element space. Note that because $F_{T_\pm}$ is affine on the interior edge e, there holds $G_e|_e \in [\mathcal{P}_{k-2}(e)]^{2\times 2}$. We also have by standard approximation theory,

$$\|\nabla\boldsymbol{w} - G_e\|_{H^m(T_\pm)} \leq Ch_T^{r-m}\|\nabla\boldsymbol{w}\|_{H^r(T_+\cup T_-)} \leq Ch_T^{r-m}\|\boldsymbol{w}\|_{H^{r+1}(T_+\cup T_-)} \quad m = 0, 1, \ldots, r, \tag{4.5}$$

where $h_T = \max\{h_{T_+}, h_{T_-}\}$. Thus, by a trace inequality,

$$\|\nabla\boldsymbol{w} - G_e\|_{L^2(e)} \leq Ch_T^{r-1/2}\|\boldsymbol{w}\|_{H^{r+1}(T_+\cup T_-)}. \tag{4.6}$$

We then write

$$\begin{aligned}
\left|\sum_{e\in\mathcal{E}_h^I}\int_e \nabla\boldsymbol{w} : [\boldsymbol{v}]\right| \leq & \left|\sum_{e\in\mathcal{E}_h^I}\int_e (\nabla\boldsymbol{w} - G_e) : [\boldsymbol{v} - \boldsymbol{E}_h\boldsymbol{v}]\right| + \left|\sum_{e\in\mathcal{E}_h^I}\int_e G_e : [\boldsymbol{v}]\right|\\
=: & I_1 + I_2.
\end{aligned} \tag{4.7}$$

To estimate I_1, we use (4.6), Lemma 4.5, and a trace inequality:

$$
\begin{aligned}
I_1 &\le \left(\sum_{e\in\mathcal{E}_h^I} h_e\|\nabla \boldsymbol{w} - G_e\|_{L^2(e)}^2\right)^{1/2} \left(\sum_{e\in\mathcal{E}_h^I} h_e^{-1}\|\boldsymbol{v} - \boldsymbol{E}_h\boldsymbol{v}\|_{L^2(e)}^2\right)^{1/2} \\
&\le C\left(\sum_{T\in\mathcal{T}_h} h_T^{2r}\|\boldsymbol{w}\|_{H^{r+1}(T)}^2\right)^{1/2} \left(\sum_{T\in\mathcal{T}_h} h_T^{2m}\|\boldsymbol{v}\|_{H^m(T)}^2\right)^{1/2} \le Ch^{r+m}\|\boldsymbol{w}\|_{H^{r+1}(\Omega)}\|\boldsymbol{v}\|_{H_h^m(\Omega_h)}.
\end{aligned}
\tag{4.8}
$$

To estimate I_2, we first observe that, by construction, for $\boldsymbol{v}\in \boldsymbol{V}_k^h$ and $e\in\mathcal{E}_h^I$, we have $[\boldsymbol{v}]|_e(a) = 0$ for all $a\in\mathcal{N}_k(T)$ with $a\in\bar{e}$. Recalling that these edge degrees of freedom are placed at Gauss-Lobatto nodes, it follows from the error of the $(k+1)$-point Gauss-Lobatto rule and the fact that G_e is a polynomial of degree $(k-2)$ on e that

$$
\begin{aligned}
\left|\int_e [\boldsymbol{v}] : G_e\right| &\le C|e|^{2k+1}\,|[\boldsymbol{v}] : G_e|_{W^{2k,\infty}(e)} \\
&\le Ch_T^{2k+1}\Big(\|\boldsymbol{v}\|_{W^{2k,\infty}(K_+)} + \|\boldsymbol{v}\|_{W^{2k,\infty}(K_-)}\Big)\|G_e\|_{W^{k-2,\infty}(e)} \quad \forall e\in\mathcal{E}_h^I,
\end{aligned}
\tag{4.9}
$$

where $K_\pm \in T_\pm^{ct}$ share edge e.

Note that a standard inverse/trace estimate yields

$$
\|G_e\|_{W^{k-2,\infty}(e)} \le Ch_T^{-1}\|G_e\|_{H^{k-2}(T_\pm)}. \tag{4.10}
$$

From here, we consider two cases:

Case 1: $k-2\le r$. For this case, we recall that $r=\min\{s-1,k-1\}$, and so $r-k\le -1$. It therefore holds that we have $h_T^{-1}\le h_T^{r-k}$. With this, (4.5), and $(k-1)\le(r+1)$, we have

$$
\begin{aligned}
\|G_e\|_{W^{k-2,\infty}(e)} \le& Ch_T^{-1}\big(\|G_e - \nabla\boldsymbol{w}\|_{H^{k-2}(T_\pm)} + \|\boldsymbol{w}\|_{H^{k-1}(T_\pm)}\big) \\
\le& Ch_T^{r-k}\|\boldsymbol{w}\|_{H^{r+1}(T_+\cup T_-)}.
\end{aligned}
$$

Case 2: $r\le k-2$. For the second case, we may apply another inverse estimate to (4.10) before applying (4.5). This yields

$$
\begin{aligned}
\|G_e\|_{W^{k-2,\infty}(e)} \le& Ch_T^{-1}h_{T_\pm}^{r-(k-2)}\|G_e\|_{H^r(T_\pm)} \\
\le& Ch_T^{r-k+1}\big(\|G_e - \nabla\boldsymbol{w}\|_{H^r(T_\pm)} + \|\boldsymbol{w}\|_{H^{r+1}(T_\pm)}\big) \\
\le& Ch_T^{r-k+1}\|\boldsymbol{w}\|_{H^{r+1}(T_+\cup T_-)}.
\end{aligned}
$$

Consequently, we take the less sharp estimate in these cases in (4.9) and apply the inverse estimates (3.1)-(3.2) to $\|\boldsymbol{v}\|_{W^{2k,\infty}(K_\pm)}$ to obtain

$$
\begin{aligned}
\left|\int_e [\boldsymbol{v}] : G_e\right| &\le Ch_T^{k+r}\Big(\|\boldsymbol{v}\|_{H^{2k}(K_+)} + \|\boldsymbol{v}\|_{H^{2k}(K_-)}\Big)\|\boldsymbol{w}\|_{H^{r+1}(T_+\cup T_-)} \\
&\le Ch_T^{k+r}\Big(\|\boldsymbol{v}\|_{H^{k}(K_+)} + \|\boldsymbol{v}\|_{H^{k}(K_-)}\Big)\|\boldsymbol{w}\|_{H^{r+1}(T_+\cup T_-)} \\
&\le Ch_T^{r+m}\Big(\|\boldsymbol{v}\|_{H^{m}(K_+)} + \|\boldsymbol{v}\|_{H^{m}(K_-)}\Big)\|\boldsymbol{w}\|_{H^{r+1}(T_+\cup T_-)}.
\end{aligned}
$$

Summing this expression over $\mathcal{E}_h^I$ we obtain an upper bound for I_2:

$$
I_2 \le Ch^{r+m}\|\boldsymbol{w}\|_{H^{r+1}(\Omega)}\|\boldsymbol{v}\|_{H_h^m(\Omega_h)}. \tag{4.11}
$$

Applying the estimates (4.8) and (4.11) towards (4.7) yields the result. ☐

Remark 4.8. We note that the result above is not as sharp if Newton-Cotes (uniformly spaced) nodes are used instead of Gauss-Lobatto nodes. Indeed, Newton-Cotes integration on m points is exact on polynomials in $\mathcal{P}^m$, if m is odd, and $\mathcal{P}^{m-1}$ if m is even, so the bound on the right-hand side of (4.9) becomes

$$Ch_T^{k+2}\bigg(\|\boldsymbol{v}\|_{W^{k+1,\infty}(K_+)} + \|\boldsymbol{v}\|_{W^{k+1,\infty}(K_-)}\bigg)\|G_e\|_{W^{k-2,\infty}(e)} \quad \forall e \in \mathcal{E}_h^I,$$

if k is odd, and

$$Ch_T^{k+3}\bigg(\|\boldsymbol{v}\|_{W^{k+2,\infty}(K_+)} + \|\boldsymbol{v}\|_{W^{k+2,\infty}(K_-)}\bigg)\|G_e\|_{W^{k-2,\infty}(e)} \quad \forall e \in \mathcal{E}_h^I,$$

if k is even. Thus, if we use equidistant points, the bound loses $k-1$ powers of h if k is odd, and $k-2$ if k is even.

4.2. **Inf-sup stability.** An inf-sup stability result for the finite element pair $\boldsymbol{V}_k^h \times Q_k^h$ was proven in [15, Theorem 4.4] in the case $k = 2$. The argument given there are essentially valid for all $k \geq 2$. Consequently, we only provide a sketch of the proof in the general case.

Theorem 4.9. *There holds*

$$\sup_{\boldsymbol{v}\in\boldsymbol{V}_k^h\setminus\{0\}} \frac{\int_{\Omega_h}(\operatorname{div}\boldsymbol{v})q}{\|\nabla_h\boldsymbol{v}\|_{L^2(\Omega_h)}} \geq C\|q\|_{L^2(\Omega_h)} \quad \forall q \in Q_{k-1}^h. \tag{4.12}$$

Sketch of proof for Theorem 4.9. Fix $q \in Q_{k-1}^h$, and set $\bar{q} \in Q_{k-1}^h$ to be piecewise constant with respect to $\mathcal{T}_h$ satisfying $\int_T (q-\bar{q})/\det(DF_T \circ F_T^{-1}) = 0$ for all $T \in \mathcal{T}_h$. By a change of variables, we see that $(q-\bar{q}) \circ F_T \in \hat{Q}_{k-1,0}$.

Next, the results in, e.g., [9] show that $\widehat{\operatorname{div}} : \hat{\boldsymbol{V}}_{k,0} \to \hat{Q}_{k-1,0}$ is surjective with bounded right inverse. Consequently, for each $T \in \mathcal{T}_h$, there exists $\hat{\boldsymbol{v}}_{1,T} \in \hat{\boldsymbol{V}}_{k,0}$ such that $\widehat{\operatorname{div}}\,\hat{\boldsymbol{v}}_{1,T} = h_T^2(q-\bar{q})|_T \circ F_T$. and $\|\hat{\boldsymbol{v}}_{1,T}\|_{H^1(\hat{T})} \leq Ch_T^2\|(q-\bar{q})|_T \circ F_T\|_{L^2(\hat{T})} \leq Ch_T\|q-\bar{q}\|_{L^2(T)}$. Setting $\boldsymbol{v}_{1,T} = (A_T\hat{\boldsymbol{v}}) \circ F_T^{-1} \in \boldsymbol{V}_{k,0}$, we have $\operatorname{div}\boldsymbol{v}_{1,T} = h_T^2(q-\bar{q})/(\det(DF_T \circ F_T^{-1}))$ by the divergence-preserving properties of the Piola transform, and $\|\nabla\boldsymbol{v}_{1,T}\|_{L^2(T)} \leq C\|q-\bar{q}\|_{L^2(T)}$ by a scaling argument.

We then define $\boldsymbol{v}_1 \in \boldsymbol{V}_k^h$ such that $\boldsymbol{v}_1|_T = \boldsymbol{v}_{1,T}$ for all $T \in \mathcal{T}_h$. Thus $\|\boldsymbol{v}_1\|_{L^2(\Omega_h)} \leq C\|q-\bar{q}\|_{L^2(\Omega_h)}$, and

$$\begin{aligned}
\sup_{\boldsymbol{v}\in\boldsymbol{V}_k^h\setminus\{0\}} \frac{\int_{\Omega_h}(\operatorname{div}\boldsymbol{v})q}{\|\nabla\boldsymbol{v}\|_{L^2(\Omega_h)}} \geq \frac{\int_{\Omega_h}(\operatorname{div}\boldsymbol{v}_1)q}{\|\nabla\boldsymbol{v}_1\|_{L^2(\Omega_h)}} &= \frac{\int_{\Omega_h}(\operatorname{div}\boldsymbol{v}_1)(q-\bar{q})}{\|\nabla\boldsymbol{v}_1\|_{L^2(\Omega_h)}} \\
&= \frac{\sum_{T\in\mathcal{T}_h} h_T^2 \int_T |q-\bar{q}|^2/(\det(DF_T \circ F_T^{-1}))}{\|\nabla\boldsymbol{v}_1\|_{L^2(\Omega_h)}} \\
&\geq \gamma_0\|q-\bar{q}\|_{L^2(\Omega_h)}.
\end{aligned}$$

Next, Theorem 4.4 in [15] shows that

$$\sup_{\boldsymbol{v}\in\boldsymbol{V}_k^h\setminus\{0\}} \frac{\int_{\Omega_h}(\operatorname{div}\boldsymbol{v})\bar{q}}{\|\nabla_h\boldsymbol{v}\|_{L^2(\Omega_h)}} \geq \gamma_1\|\bar{q}\|_{L^2(\Omega_h)}.$$

Consequently, it follows that

$$\begin{aligned}
\|q\|_{L^2(\Omega_h)} \leq& \|q-\bar{q}\|_{L^2(\Omega_h)} + \|\bar{q}\|_{L^2(\Omega_h)} \\
\leq& (\gamma_0^{-1} + \gamma_1^{-1}(1+\gamma_0^{-1})) \sup_{\boldsymbol{v}\in\boldsymbol{V}_h^k\setminus\{0\}} \frac{\int_{\Omega_h}(\operatorname{div}\boldsymbol{v})q}{\|\nabla\boldsymbol{v}\|_{L^2(\Omega_h)}}.
\end{aligned}$$

5. The Stokes System and Finite Element Method

We let $(\boldsymbol{u}, p) \in \boldsymbol{H}_0^1(\Omega) \times L_0^2(\Omega)$ be the solution to the Stokes problem

$$\begin{cases} -\nu \Delta \boldsymbol{u} + \nabla p = \boldsymbol{f} & \text{in } \Omega, \\ \operatorname{div} \boldsymbol{u} = 0 & \text{in } \Omega, \\ \boldsymbol{u} = 0 & \text{on } \partial\Omega, \end{cases} \tag{5.1}$$

where $\nu > 0$ is the viscosity. We assume the domain Ω and source function $\boldsymbol{f}$ are sufficiently smooth such that $(\boldsymbol{u}, p) \in \boldsymbol{H}^s(\Omega) \times H^{s-1}(\Omega)$ with $s \geq 2$. We then extend the velocity solution to $\mathbb{R}^2$ such that the extension (still denoted by $\boldsymbol{u}$) is divergence-free and satisfies [11]

$$\|\boldsymbol{u}\|_{H^s(\mathbb{R}^2)} \leq C \|\boldsymbol{u}\|_{H^s(\Omega)}. \tag{5.2}$$

Likewise, we extend the pressure solution p to $\mathbb{R}^2$ with $\|p\|_{H^{s-1}(\mathbb{R}^2)} \leq C\|p\|_{H^{s-1}(\Omega)}$ and extend the source function by setting $\boldsymbol{f} = -\nu\Delta\boldsymbol{u} + \nabla p$ in $\mathbb{R}^2$.

We define the continuous bilinear forms

$$a(\boldsymbol{u}, \boldsymbol{v}) := \int_\Omega \nu \nabla \boldsymbol{u} : \nabla \boldsymbol{v}, \qquad b(\boldsymbol{v}, p) := -\int_\Omega (\operatorname{div} \boldsymbol{v}) p,$$

and the discrete bilinear forms

$$a_h(\boldsymbol{u}_h, \boldsymbol{v}) := \int_{\Omega_h} \nu \nabla_h \boldsymbol{u}_h : \nabla_h \boldsymbol{v}, \qquad b_h(\boldsymbol{v}, p_h) := -\int_{\Omega_h} (\operatorname{div} \boldsymbol{v}) p_h.$$

Clearly, the solution to (5.1) solves the variational problem

$$a(\boldsymbol{u}, \boldsymbol{v}) = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v} \quad \forall \boldsymbol{v} \in \boldsymbol{X} := \{\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega) : \ \operatorname{div} \boldsymbol{v} = 0\}. \tag{5.3}$$

We define the finite element method as finding $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_k^h \times Q_{k-1}^h$ such that

$$a_h(\boldsymbol{u}_h, \boldsymbol{v}) + b_h(\boldsymbol{v}, p_h) = \int_{\Omega_h} \boldsymbol{f}_h \cdot \boldsymbol{v} \quad \forall \boldsymbol{v} \in \boldsymbol{V}_k^h, \tag{5.4a}$$

$$-b_h(\boldsymbol{u}_h, q) = 0 \qquad \forall q \in Q_k^h, \tag{5.4b}$$

where $\boldsymbol{f}_h \in \boldsymbol{L}^2(\Omega_h)$ is a suitable (and computable) approximation to $\boldsymbol{f}|_\Omega$. It follows from the inf-sup condition in Theorem 4.9 and the Poincare inequality in Corollary 4.6 that there exists a unique solution to (5.4). Moreover, by a simple generalization of [15, Lemma 5.2], the method (5.4) yields divergence-free velocity approximations.

Lemma 5.1. *Let $\boldsymbol{u}_h \in \boldsymbol{V}_k^h$ satisfy* (5.4b). *Then $\operatorname{div} \boldsymbol{u}_h = 0$ in Ω_h.*

5.1. Energy estimates.

In this section, we derive error estimates for the approximation velocity and pressure solutions in the H^1 and L^2 norms, respectively. To this end, we define the discrete space of divergence-free functions

$$\boldsymbol{X}_k^h := \{\boldsymbol{v} \in \boldsymbol{V}_k^h : \operatorname{div} \boldsymbol{v} = 0\} \not\subseteq \boldsymbol{X} := \{\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega) : \operatorname{div} \boldsymbol{v} = 0\},$$

and note that functions in this space are not necessarily in $\boldsymbol{H}_0^1$. Lemma 5.1 shows that $\boldsymbol{u}_h \in \boldsymbol{X}_h^k$, and thus the velocity solution $\boldsymbol{u}_h$ is uniquely characterized as the solution of the Poisson-type problem

$$a_h(\boldsymbol{u}_h, \boldsymbol{v}) = \int_{\Omega_h} \boldsymbol{f}_h \cdot \boldsymbol{v} \quad \forall \boldsymbol{v} \in \boldsymbol{V}_k^h. \tag{5.5}$$

Theorem 5.2. *Let $(\boldsymbol{u}, p) \in \boldsymbol{H}^s(\Omega) \times H^{s-1}(\Omega)$ satisfy* (5.1) *with $s \geq 2$. Then there holds*

$$\|\nabla_h(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2(\Omega_h)} \leq C \left(h^{\ell-1} \|\boldsymbol{u}\|_{H^\ell(\Omega)} + \nu^{-1} |\boldsymbol{f} - \boldsymbol{f}_h|_{\boldsymbol{X}_k^*} \right), \tag{5.6}$$

where $\ell = \min\{k+1, s\}$, and

$$|\boldsymbol{f} - \boldsymbol{f}_h|_{\boldsymbol{X}_k^*} = \sup_{\boldsymbol{v} \in \boldsymbol{X}_k^h \setminus \{0\}} \frac{\int_{\Omega_h} (\boldsymbol{f} - \boldsymbol{f}_h) \cdot \boldsymbol{v}}{\|\nabla_h \boldsymbol{v}\|_{L^2(\Omega_h)}}.$$

The pressure approximation, p_h, satisfies

$$\|p - p_h\|_{L^2(\Omega_h)} \le C\left(h^{\ell-1}(\nu\|\boldsymbol{u}\|_{H^\ell(\Omega)} + \|p\|_{H^{\ell-1}(\Omega)}) + \|\boldsymbol{f} - \boldsymbol{f}_h\|_{L^2(\Omega_h)}\right). \tag{5.7}$$

Proof. From standard theory of non-conforming finite elements (see, for example, [4]) and the inf-sup condition (4.12),

$$\begin{aligned}
\nu\|\nabla_h(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2(\Omega_h)} \le& \inf_{\boldsymbol{w}\in \boldsymbol{X}_k^h} \nu\|\nabla_h(\boldsymbol{u} - \boldsymbol{w})\|_{L^2(\Omega_h)} + \sup_{\boldsymbol{v}\in \boldsymbol{X}_k^h\setminus\{0\}} \frac{a_h(\boldsymbol{u}_h - \boldsymbol{u}, \boldsymbol{v})}{\|\nabla_h \boldsymbol{v}\|_{L^2(\Omega_h)}} \\
\le& C \inf_{\boldsymbol{w}\in \boldsymbol{V}_k^h} \nu\|\nabla_h(\boldsymbol{u} - \boldsymbol{w})\|_{L^2(\Omega_h)} + \sup_{\boldsymbol{v}\in \boldsymbol{X}_k^h\setminus\{0\}} \frac{a_h(\boldsymbol{u}_h - \boldsymbol{u}, \boldsymbol{v})}{\|\nabla_h \boldsymbol{v}\|_{L^2(\Omega_h)}} \\
\le& C h^{\ell-1}\nu\|\boldsymbol{u}\|_{H^\ell(\Omega)} + \sup_{\boldsymbol{v}\in \boldsymbol{X}_k^h\setminus\{0\}} \frac{a_h(\boldsymbol{u}_h - \boldsymbol{u}, \boldsymbol{v})}{\|\nabla_h \boldsymbol{v}\|_{L^2(\Omega_h)}},
\end{aligned} \tag{5.8}$$

where the final step follows from Lemma 4.4.

To address the consistency term, we note that we have $\forall \boldsymbol{v} \in \boldsymbol{X}_k^h$

$$\begin{aligned}
a_h(\boldsymbol{u}_h - \boldsymbol{u}, \boldsymbol{v}) =& \int_{\Omega_h} \boldsymbol{f}\cdot \boldsymbol{v} - a_h(\boldsymbol{u}, \boldsymbol{v}) + \int_{\Omega_h} (\boldsymbol{f}_h - \boldsymbol{f})\cdot \boldsymbol{v} \\
=& -\nu \int_{\Omega_h} \Delta \boldsymbol{u}\cdot \boldsymbol{v} - a_h(\boldsymbol{u}, \boldsymbol{v}) + \int_{\Omega_h} (\boldsymbol{f}_h - \boldsymbol{f})\cdot \boldsymbol{v}.
\end{aligned} \tag{5.9}$$

Note that the last step uses the fact that $\boldsymbol{v} \in \boldsymbol{X}_k^h$, therefore $\operatorname{div} \boldsymbol{v} = 0$ and $\boldsymbol{v} = 0$ on $\partial\Omega_h$.

We then apply a standard integration-by-parts formula in (5.9), Lemma 4.7 (with $m = 1$, and noting $r = \min\{s-1, k-1\} \le \ell - 1$) and Corollary 4.6 to obtain

$$\begin{aligned}
a_h(\boldsymbol{u}_h - \boldsymbol{u}, \boldsymbol{v}) =& -\nu \sum_{e\in\mathcal{E}_h^I} \int_e \nabla \boldsymbol{u} : [\boldsymbol{v}] + \int_{\Omega_h} (\boldsymbol{f}_h - \boldsymbol{f})\cdot \boldsymbol{v} \\
\le& C\nu h^{\ell-1}\|\boldsymbol{u}\|_{H^\ell(\Omega)}\|\nabla_h \boldsymbol{v}\|_{L^2(\Omega_h)} + \|\boldsymbol{f}_h - \boldsymbol{f}\|_{X_k^*}\|\nabla_h \boldsymbol{v}\|_{L^2(\Omega_h)}.
\end{aligned} \tag{5.10}$$

Finally, to complete the velocity bound (5.6), we apply this estimate to (5.8).

To prove the pressure bound (5.7), we fix $q \in Q_{k-1}^h$. For any $\boldsymbol{v} \in \boldsymbol{V}_k^h$, we then have the following identity, using integration by parts and (5.1):

$$\begin{aligned}
\int_{\Omega_h} (\operatorname{div} \boldsymbol{v})(p_h - q) =& a_h(\boldsymbol{u}_h, \boldsymbol{v}) - \int_{\Omega_h} (\operatorname{div} \boldsymbol{v})q - \int_{\Omega_h} (\boldsymbol{f}_h - \boldsymbol{f})\cdot \boldsymbol{v} - \int_{\Omega_h} \boldsymbol{f}\cdot \boldsymbol{v} \\
=& a_h(\boldsymbol{u}_h, \boldsymbol{v}) + \nu \int_{\Omega_h} \Delta \boldsymbol{u}\cdot \boldsymbol{v} - \int_{\Omega_h} \nabla p\cdot \boldsymbol{v} - \int_{\Omega_h} (\operatorname{div} \boldsymbol{v})q - \int_{\Omega_h} (\boldsymbol{f}_h - \boldsymbol{f})\cdot \boldsymbol{v} \\
=& a_h(\boldsymbol{u}_h - \boldsymbol{u}, \boldsymbol{v}) - \int_{\Omega_h} (\operatorname{div} \boldsymbol{v})(q - p) - \int_{\Omega_h} (\boldsymbol{f}_h - \boldsymbol{f})\cdot \boldsymbol{v} + \nu \sum_{e\in\mathcal{E}_h^I} \int_e \nabla \boldsymbol{u} : [\boldsymbol{v}].
\end{aligned} \tag{5.11}$$

Then, applying (4.4) to (5.11) and Corollary 4.6, we have

$$\begin{aligned}
\int_{\Omega_h} (\operatorname{div} \boldsymbol{v})(p_h - q) \le& C\Big(\nu\|\nabla_h(\boldsymbol{u}_h - \boldsymbol{u})\|_{L^2(\Omega_h)} + \|q - p\|_{L^2(\Omega_h)} \\
&+ \nu h^{\ell-1}\|\boldsymbol{u}\|_{H^\ell(\Omega_h)} + \|\boldsymbol{f} - \boldsymbol{f}_h\|_{L^2(\Omega_h)}\Big)\|\nabla_h \boldsymbol{v}\|_{L^2(\Omega_h)}.
\end{aligned} \tag{5.12}$$

Finally, by triangle inequality and Theorem 4.9, we have

$$\begin{aligned}
\|p - p_h\|_{L^2(\Omega_h)} \le& \|p - q\|_{L^2(\Omega_h)} + \|p_h - q\|_{L^2(\Omega_h)} \\
\le& \|p - q\|_{L^2(\Omega_h)} + \sup_{\boldsymbol{v}\in \boldsymbol{V}^h\setminus\{0\}} \frac{\int_{\Omega_h} (\operatorname{div} \boldsymbol{v})(p_h - q)}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega_h)}}.
\end{aligned}$$

Applying (5.12) to this result, taking the infimum over $q \in Q_{k-1}^h$, and using (5.6) completes the proof. $\square$

Corollary 5.3. *Assume the conditions in Theorem 5.2 are satisfied, and in addition, assume the mesh $\mathcal{T}_h$ is quasi-uniform. Then the solution $\boldsymbol{u}_h \in \boldsymbol{V}_k^h$ to* (5.4) *satisfies*

$$\|\boldsymbol{u}_h\|_{H_h^\ell(\Omega_h)} \le C\bigg(\|\boldsymbol{u}\|_{H^\ell(\Omega)} + h^{1-\ell}\nu^{-1}|\boldsymbol{f} - \boldsymbol{f}_h|_{\boldsymbol{X}_k^*}\bigg).$$

Proof. Define $\boldsymbol{I}_k^h\boldsymbol{u} \in \boldsymbol{V}_k^h$ to be the approximation to $\boldsymbol{u}$ given in Lemma 4.4. Then, applying the inverse inequality (3.1), Lemma 4.4, and Theorem 5.2, we have

$$\begin{aligned}\|\boldsymbol{u}_h\|_{H_h^\ell(\Omega_h)} \le& C\big(\|\boldsymbol{u} - \boldsymbol{I}_k^h\boldsymbol{u}\|_{H_h^\ell(\Omega_h)} + h^{1-\ell}\|\boldsymbol{I}_k^h\boldsymbol{u} - \boldsymbol{u}_h\|_{H_h^1(\Omega_h)} + \|\boldsymbol{u}\|_{H^\ell(\Omega)}\big)\\ \le& C\big(\|\boldsymbol{u} - \boldsymbol{I}_k^h\boldsymbol{u}\|_{H_h^\ell(\Omega_h)} + h^{1-\ell}\|\boldsymbol{I}_k^h\boldsymbol{u} - \boldsymbol{u}\|_{H_h^1(\Omega_h)} + h^{1-\ell}\|\boldsymbol{u} - \boldsymbol{u}_h\|_{H_h^1(\Omega_h)} + \|\boldsymbol{u}\|_{H^\ell(\Omega)}\big)\\ \le& C\big(\|\boldsymbol{u}\|_{H^\ell(\Omega)} + h^{1-\ell}\nu^{-1}|\boldsymbol{f} - \boldsymbol{f}_h|_{\boldsymbol{X}_k^*}\big).\end{aligned}$$

$\square$

Remark 5.4. If $\boldsymbol{f}$ is sufficiently smooth, and $\boldsymbol{f}_h$ is, for example, the kth degree nodal (isoparametric) interpolant, then $|\boldsymbol{f} - \boldsymbol{f}_h|_{\boldsymbol{X}_k^*} \le \|\boldsymbol{f} - \boldsymbol{f}_h\|_{L^2(\Omega_h)} \le Ch^{k+1}\|\boldsymbol{f}\|_{H^{k+1}(\Omega_h)}$. Thus, Corollary 5.3 yields

$$\|\boldsymbol{u}_h\|_{H_h^\ell(\Omega_h)} \le C\left(\|\boldsymbol{u}\|_{H^\ell(\Omega)} + h^{k-\ell+2}\nu^{-1}\|\boldsymbol{f}\|_{H^{k+1}(\Omega_h)}\right). \tag{5.13}$$

6. Convergence analysis in L^2

In this section, we prove the following optimal-order L^2 error estimate.

Theorem 6.1. *Assume the conditions in Theorem 5.2 are satisfied, and in addition, assume the mesh $\mathcal{T}_h$ is quasi-uniform. We have*

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega_h)} \le C\bigg(h^\ell\|\boldsymbol{u}\|_{H^\ell(\Omega)} + (\nu^{-1}h + 1)|\boldsymbol{f} - \boldsymbol{f}_h|_{\boldsymbol{X}_h^*}\bigg), \tag{6.1}$$

where C is a constant that does not depend on the mesh parameter h, and we recall $\ell = \min\{s, k+1\}$.

Proof. To derive (6.1), we first write

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega_h)} \le \|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega_h\setminus\Omega)} + \|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega_h\cap\Omega)} =: J_1 + J_2. \tag{6.2}$$

To bound J_1, we introduce $\boldsymbol{E}_h : \boldsymbol{V}_k^h \to \boldsymbol{H}_0^1(\Omega_h)$ as defined in Lemma 4.5. Consequently, we may write

$$J_1 \le \|\boldsymbol{u} - \boldsymbol{E}_h\boldsymbol{u}_h\|_{L^2(\Omega_h\setminus\Omega)} + \|\boldsymbol{E}_h\boldsymbol{u}_h - \boldsymbol{u}_h\|_{L^2(\Omega_h\setminus\Omega)}.$$

A bound of the second term in this sum follows from Lemma 4.5 and Corollary 5.3:

$$\begin{aligned}J_1 \le& \|\boldsymbol{u} - \boldsymbol{E}_h\boldsymbol{u}_h\|_{L^2(\Omega_h\setminus\Omega)} + \|\boldsymbol{E}_h\boldsymbol{u}_h - \boldsymbol{u}_h\|_{L^2(\Omega_h)}\\ \le& \|\boldsymbol{u} - \boldsymbol{E}_h\boldsymbol{u}_h\|_{L^2(\Omega_h\setminus\Omega)} + Ch^\ell\|\boldsymbol{u}_h\|_{H_h^{\ell-1}(\Omega_h)}\\ \le& \|\boldsymbol{u} - \boldsymbol{E}_h\boldsymbol{u}_h\|_{L^2(\Omega_h\setminus\Omega)} + C\bigg(h^\ell\|\boldsymbol{u}\|_{H^\ell(\Omega_h)} + h\nu^{-1}|\boldsymbol{f} - \boldsymbol{f}_h|_{\boldsymbol{X}_h^*}\bigg).\end{aligned}$$

To bound the remaining term, begin with Hölder's inequality and recall that H^1 embeds in L^6 and $k \ge 2$. Thus we have

$$\begin{aligned}\|\boldsymbol{u} - \boldsymbol{E}_h\boldsymbol{u}_h\|_{L^2(\Omega_h\setminus\Omega)} \le& |\Omega_h \setminus \Omega|^{1/3}\|\boldsymbol{u} - \boldsymbol{E}_h\boldsymbol{u}_h\|_{L^6(\Omega_h)}\\ \le& Ch^{(k+1)/3}\|\boldsymbol{u} - \boldsymbol{E}_h\boldsymbol{u}_h\|_{L^6(\Omega_h)}\\ \le& Ch\|\boldsymbol{u} - \boldsymbol{E}_h\boldsymbol{u}_h\|_{H^1(\Omega_h)}.\end{aligned}$$

It follows from Theorem 5.2, Lemma 4.5, and Corollary 5.3 that

$$\begin{aligned}\|\boldsymbol{u}-\boldsymbol{E}_h\boldsymbol{u}_h\|_{L^2(\Omega_h\setminus\Omega)} \le& Ch\Big(\|\boldsymbol{u}-\boldsymbol{u}_h\|_{H_h^1(\Omega_h)}+\|\boldsymbol{u}_h-\boldsymbol{E}_h\boldsymbol{u}_h\|_{H_h^1(\Omega_h)}\Big)\\ \le& C\Big(h^\ell\|\boldsymbol{u}\|_{H^\ell(\Omega)}+h\nu^{-1}|\boldsymbol{f}-\boldsymbol{f}_h|_{\boldsymbol{X}_k^*}\Big).\end{aligned}$$

Combining this with the result above, we have

$$J_1 \le C\Big(h^\ell\|\boldsymbol{u}\|_{H^\ell(\Omega)}+h\nu^{-1}|\boldsymbol{f}-\boldsymbol{f}_h|_{\boldsymbol{X}_k^*}\Big). \tag{6.3}$$

To bound J_2, we let $\boldsymbol{\phi}\in\boldsymbol{L}^2(\Omega\cup\Omega_h)$ such that $\boldsymbol{\phi}|_{\Omega\cap\Omega_h}=(\boldsymbol{u}-\boldsymbol{u}_h)|_{\Omega\cap\Omega_h}$ and $\boldsymbol{\phi}|_{\Omega\cup\Omega_h\setminus(\Omega\cap\Omega_h)}=0$. We then define $(\boldsymbol{\psi},r)\in\boldsymbol{H}_0^1(\Omega)\times L_0^2(\Omega)$ to be the solution to the auxiliary problem

$$\begin{cases}-\nu\Delta\boldsymbol{\psi}+\nabla r=\boldsymbol{\phi} & \text{in } \Omega,\\ \operatorname{div}\boldsymbol{\psi}=0 & \text{in } \Omega.\end{cases} \tag{6.4}$$

Because $\partial\Omega$ is smooth and $\boldsymbol{\phi}|_\Omega\in\boldsymbol{L}^2(\Omega)$, there holds $\boldsymbol{\psi}\in\boldsymbol{H}^2(\Omega)$ with $\|\boldsymbol{\psi}\|_{H^2(\Omega)}\le C\|\boldsymbol{\phi}\|_{L^2(\Omega)}=C\|\boldsymbol{\phi}\|_{L^2(\Omega\cap\Omega_h)}$ by elliptic regularity theory. Similar to the solution $\boldsymbol{u}$ in (5.2), we extend $\boldsymbol{\psi}$ to $\mathbb{R}^2$ in such a way that preserves the divergence–free condition (cf. [11]) and

$$\|\boldsymbol{\psi}\|_{H^2(\mathbb{R}^2)}\le C\|\boldsymbol{\psi}\|_{H^2(\Omega)}\le C\|\boldsymbol{\phi}\|_{L^2(\Omega\cap\Omega_h)}. \tag{6.5}$$

Finally, we define $\boldsymbol{\psi}_h\in\boldsymbol{X}_k^h$ to be the approximation of $\boldsymbol{\psi}$ on Ω_h satisfying

$$a_h(\boldsymbol{\psi}_h,\boldsymbol{v})=\int_{\Omega_h}\boldsymbol{\phi}\cdot\boldsymbol{v}\quad\forall\boldsymbol{v}\in\boldsymbol{V}_k^h. \tag{6.6}$$

We note that $\boldsymbol{\psi}$ and $\boldsymbol{\psi}_h$ are analogous to $\boldsymbol{u}$ and $\boldsymbol{u}_h$, respectively, in Theorem 5.2 when $s=2$ (so that $\ell=2$), with $\boldsymbol{\phi}$ replacing both $\boldsymbol{f}$ and $\boldsymbol{f}_h$. Therefore, the following estimate holds:

$$\|\boldsymbol{\psi}-\boldsymbol{\psi}_h\|_{H_h^1(\Omega_h)}\le Ch\|\boldsymbol{\psi}\|_{H^2(\Omega)}\le C\|\boldsymbol{\phi}\|_{L^2(\Omega\cap\Omega_h)}. \tag{6.7}$$

Additionally, applying Corollary 5.3 yields

$$\|\boldsymbol{\psi}_h\|_{H_h^2(\Omega_h)}\le C\|\boldsymbol{\psi}\|_{H^2(\Omega)}\le C\|\boldsymbol{\phi}\|_{L^2(\Omega\cap\Omega_h)}. \tag{6.8}$$

Next, we write

$$\begin{aligned}(J_2)^2=\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega\cap\Omega_h)}^2=&\int_{\Omega\cap\Omega_h}\boldsymbol{\phi}\cdot(\boldsymbol{u}-\boldsymbol{u}_h)\\ =&\int_\Omega\boldsymbol{\phi}\cdot\boldsymbol{u}-\int_{\Omega_h}\boldsymbol{\phi}\cdot\boldsymbol{u}_h\\ =&a(\boldsymbol{u},\boldsymbol{\psi})-a_h(\boldsymbol{u}_h,\boldsymbol{\psi}_h)\\ =&[a(\boldsymbol{u},\boldsymbol{\psi})-a_h(\boldsymbol{u},\boldsymbol{\psi})]+a_h(\boldsymbol{u},\boldsymbol{\psi}-\boldsymbol{\psi}_h)+a_h(\boldsymbol{u}-\boldsymbol{u}_h,\boldsymbol{\psi}_h).\end{aligned} \tag{6.9}$$

We now consider the terms of (6.9) separately.

Bound of $[a(\boldsymbol{u},\boldsymbol{\psi})-a_h(\boldsymbol{u},\boldsymbol{\psi})]$. To bound the first terms of (6.9), we begin with

$$\begin{aligned}|a(\boldsymbol{u},\boldsymbol{\psi})-a_h(\boldsymbol{u},\boldsymbol{\psi})|=&\nu\left|\int_\Omega\nabla\boldsymbol{u}:\nabla\boldsymbol{\psi}-\int_{\Omega_h}\nabla\boldsymbol{u}:\nabla\boldsymbol{\psi}\right|\\ =&\nu\left|\int_{\Omega\setminus\Omega_h}\nabla\boldsymbol{u}:\nabla\boldsymbol{\psi}-\int_{\Omega_h\setminus\Omega}\nabla\boldsymbol{u}:\nabla\boldsymbol{\psi}\right|\\ \le& C\nu\|\nabla\boldsymbol{u}\|_{L^2(\Omega\Delta\Omega_h)}\|\nabla\boldsymbol{\psi}\|_{L^2(\Omega\Delta\Omega_h)}.\end{aligned}$$

The result of Lemma 2.1 implies

$$\|\nabla\boldsymbol{u}\|_{L^2(\Omega\Delta\Omega_h)}\le Ch^{(k+1)/2}\|\boldsymbol{u}\|_{H^2(\Omega)},\quad \|\nabla\boldsymbol{\psi}\|_{L^2(\Omega\Delta\Omega)}\le Ch^{(k+1)/2}\|\boldsymbol{\psi}\|_{H^2(\Omega)},$$

from which we get

$$|a(\boldsymbol{u},\boldsymbol{\psi}) - a_h(\boldsymbol{u},\boldsymbol{\psi})| \le Ch^{k+1}\nu\|\boldsymbol{u}\|_{H^2(\Omega)}\|\boldsymbol{\psi}\|_{H^2(\Omega)}. \tag{6.10}$$

Bound of $a_h(\boldsymbol{u},\boldsymbol{\psi}-\boldsymbol{\psi}_h)+a_h(\boldsymbol{u}-\boldsymbol{u}_h,\boldsymbol{\psi}_h)$. It now remains to bound the last two terms in (6.9). To begin, we write

$$\begin{aligned} a_h(\boldsymbol{u},\boldsymbol{\psi}-\boldsymbol{\psi}_h)+a_h(\boldsymbol{u}-\boldsymbol{u}_h,\boldsymbol{\psi}_h) =& a_h(\boldsymbol{u}-\boldsymbol{u}_h,\boldsymbol{\psi}-\boldsymbol{\psi}_h)+a_h(\boldsymbol{u}_h,\boldsymbol{\psi}-\boldsymbol{\psi}_h)+a_h(\boldsymbol{u}-\boldsymbol{u}_h,\boldsymbol{\psi}_h)\\ \le& \nu\|\boldsymbol{u}-\boldsymbol{u}_h\|_{H^1(\Omega_h)}\|\boldsymbol{\psi}-\boldsymbol{\psi}_h\|_{H^1(\Omega_h)}+a_h(\boldsymbol{u}_h,\boldsymbol{\psi}-\boldsymbol{\psi}_h)+a_h(\boldsymbol{u}-\boldsymbol{u}_h,\boldsymbol{\psi}_h)\\ \le& C\left(\nu h^{\ell}\|\boldsymbol{u}\|_{H^{\ell}(\Omega)}+h|\boldsymbol{f}-\boldsymbol{f}_h|_{\boldsymbol{X}_h^*}\right)\|\boldsymbol{\phi}\|_{L^2(\Omega\cap\Omega_h)}\\ &+a_h(\boldsymbol{u}_h,\boldsymbol{\psi}-\boldsymbol{\psi}_h)+a_h(\boldsymbol{u}-\boldsymbol{u}_h,\boldsymbol{\psi}_h) \end{aligned} \tag{6.11}$$

by Theorem 5.2 and (6.7).

Recalling (5.10), we have by Lemma 4.7 (with $m=2$, and noting $\ell=\min\{s,k+1\}\le\min\{s-1,k-1\}+2=r+2$)

$$\begin{aligned} a_h(\boldsymbol{u}-\boldsymbol{u}_h,\boldsymbol{\psi}_h) =& -\nu\sum_{e\in\mathcal{E}_h^I}\int_e\nabla\boldsymbol{u}:[\boldsymbol{\psi}_h]+\int_{\Omega_h}(\boldsymbol{f}-\boldsymbol{f}_h)\cdot\boldsymbol{\psi}_h\\ \le& C\nu h^{\ell}\|\boldsymbol{u}\|_{H^{\ell}(\Omega)}\|\boldsymbol{\psi}_h\|_{H_h^2(\Omega_h)}+|\boldsymbol{f}-\boldsymbol{f}_h|_{X_h^*}\|\nabla\boldsymbol{\psi}_h\|_{L^2(\Omega_h)}\\ \le& C\left(\nu h^{\ell}\|\boldsymbol{u}\|_{H^{\ell}(\Omega)}+|\boldsymbol{f}-\boldsymbol{f}_h|_{X_h^*}\right)\|\boldsymbol{\phi}\|_{L^2(\Omega\cap\Omega_h)}. \end{aligned} \tag{6.12}$$

By an analogous argument, but with $s=2$ and $m=k$ in Lemma 4.7 (so that $r=1$) , we have

$$\begin{aligned} a_h(\boldsymbol{u}_h,\boldsymbol{\psi}-\boldsymbol{\psi}_h) =& -\nu\sum_{e\in\mathcal{E}_h^I}\int_e\nabla\boldsymbol{\psi}:[\boldsymbol{u}_h]\\ \le& C\nu h^{k+1}\|\boldsymbol{\psi}\|_{H^2(\Omega)}\|\boldsymbol{u}_h\|_{H_h^k(\Omega_h)}\\ \le& C\nu h^{\ell}\|\boldsymbol{\phi}\|_{L^2(\Omega\cap\Omega_h)}\|\boldsymbol{u}\|_{H^{\ell}(\Omega)}. \end{aligned} \tag{6.13}$$

Combining (6.11)–(6.13) yields

$$a_h(\boldsymbol{u},\boldsymbol{\psi}-\boldsymbol{\psi}_h)+a_h(\boldsymbol{u}-\boldsymbol{u}_h,\boldsymbol{\psi}_h)\le C\left(\nu h^{\ell}\|\boldsymbol{u}\|_{H^{\ell}(\Omega)}+|\boldsymbol{f}-\boldsymbol{f}_h|_{\boldsymbol{X}_h^*}\right)\|\boldsymbol{\phi}\|_{L^2(\Omega\cap\Omega_h)}, \tag{6.14}$$

and so applying this estimate and (6.10) to (6.9) (recalling that $\boldsymbol{\phi}=\boldsymbol{u}-\boldsymbol{u}_h$ on $\Omega\cap\Omega_h$), we obtain

$$J_2\le C\left(\nu h^{\ell}\|\boldsymbol{u}\|_{H^{\ell}(\Omega)}+|\boldsymbol{f}-\boldsymbol{f}_h|_{\boldsymbol{X}_h^*}\right). \tag{6.15}$$

Finally, applying (6.3) and (6.15) to (6.2) completes the proof. □

7. Numerical Experiments

We perform a series of numerical experiments to compare with the theoretical results presented in this paper. We focus on the $k=3$ case below. Numerical experiments for the $k=2$ case can be found in [15].

We define our domain to be the region bounded by the ellipse

$$\Omega=\{x\in\mathbb{R}^2:\ \frac{x_1^2}{2.25}+x_2^2<1\},$$

and assign data according to the exact solution

$$\boldsymbol{u}=\begin{pmatrix}1.5(\frac{x_1^2}{2.25}+x_2^2-1)(\frac{8x_1^2y}{2.25}+\frac{x_1^2}{2.25}+5x_2^2-1)\\ \frac{-4x_1}{1.5}(\frac{x_1^2}{2.25}+x_2^2-1)(\frac{3x_1^2}{2.25}+x_2^2+x_2-1)\end{pmatrix},\quad p=10(\frac{x_1^2}{2.25}+x_2^2-\frac{1}{2}). \tag{7.1}$$

We take $\boldsymbol{f}_h$ to be the cubic (nodal) Lagrange of $\boldsymbol{f}$ and set the viscosity to $\nu=1$ to compute the finite element method described in (5.4). We subsequently compute the errors for decreasing mesh parameter h.

7.1. **Isoparametric and affine comparison.** In Table 1, we compare the isoparametric approximation defined through the Piola transform described in this paper with the corresponding affine approximation. Both tests were run on $\boldsymbol{\mathcal{P}}^3 - \mathcal{P}^2$ Scott-Vogelius elements with all edge degrees of freedom placed at the Gauss-Lobatto points. For the isoparametric approximation, we observe the optimal convergence rates predicted by the theory:

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega_h)} = \mathcal{O}(h^4), \quad \|\nabla(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2(\Omega_h)} = \mathcal{O}(h^3)$$
$$\|p - p_h\|_{L^2(\Omega_h)} = \mathcal{O}(h^3).$$

For the affine approximation, we observe suboptimal convergence.

Isoparametric						
h	$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega_h)}$	rate	$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{H^1(\Omega_h)}$	rate	$\|p - p_h\|_{L^2(\Omega_h)}$	rate
0.654	$3.391 \cdot 10^{-1}$	–	3.363	–	7.071	–
0.318	$2.392 \cdot 10^{-2}$	3.672	$4.257 \cdot 10^{-1}$	2.862	$5.234 \cdot 10^{-1}$	3.604
0.158	$1.675 \cdot 10^{-3}$	3.791	$6.232 \cdot 10^{-2}$	2.740	$8.845 \cdot 10^{-2}$	2.537
0.079	$1.139 \cdot 10^{-4}$	3.866	$9.046 \cdot 10^{-3}$	2.776	$1.298 \cdot 10^{-2}$	2.761
0.039	$7.183 \cdot 10^{-6}$	3.985	$1.225 \cdot 10^{-3}$	2.882	$1.695 \cdot 10^{-3}$	2.935
Affine						
h	$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega_h)}$	rate	$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{H^1(\Omega_h)}$	rate	$\|p - p_h\|_{L^2(\Omega_h)}$	rate
0.654	$6.411 \cdot 10^{-1}$	–	3.705	–	8.289	–
0.318	$1.525 \cdot 10^{-1}$	1.989	1.248	1.507	2.288	1.782
0.158	$3.667 \cdot 10^{-2}$	2.032	$4.589 \cdot 10^{-1}$	1.427	$8.526 \cdot 10^{-1}$	1.408
0.079	$8.779 \cdot 10^{-3}$	2.056	$1.635 \cdot 10^{-1}$	1.484	$3.005 \cdot 10^{-1}$	1.500
0.039	$2.133 \cdot 10^{-3}$	2.040	$5.741 \cdot 10^{-2}$	1.509	$1.056 \cdot 10^{-1}$	1.508

TABLE 1. Errors and rates for the Isoparametric approximations with Gauss-Lobatto nodes compared to the affine approximation.

7.2. **Dependence on degrees of freedom.** In Remark 4.8, we note the error estimate may lose up to $k-1$ powers of h if equidistant nodes are used in places of Gauss-Lobatto nodes. To test this, we compute the errors for the isoparametric approximation with the standard, equidistant placement of degrees of freedom in order to test whether Gauss-Lobatto points are necessary or simply a tool for the analysis. We compare these results, shown in Table 2, with those in Table 1, and we see that the isoparametric approximation with equidistant points is indeed suboptimal.

Isoparametric with Equidistant Degrees of Freedom						
h	$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega_h)}$	rate	$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{H^1(\Omega_h)}$	rate	$\|p - p_h\|_{L^2(\Omega_h)}$	rate
0.654	$3.393 \cdot 10^{-1}$	–	3.298	–	6.906	–
0.318	$2.307 \cdot 10^{-2}$	3.723	$3.998 \cdot 10^{-1}$	2.922	$4.813 \cdot 10^{-1}$	3.689
0.158	$1.657 \cdot 10^{-3}$	3.755	$5.538 \cdot 10^{-2}$	2.819	$8.089 \cdot 10^{-2}$	2.543
0.079	$1.212 \cdot 10^{-4}$	3.762	$7.938 \cdot 10^{-2}$	2.794	$1.246 \cdot 10^{-2}$	2.691
0.039	$1.791 \cdot 10^{-5}$	2.756	$2.043 \cdot 10^{-3}$	1.957	$2.946 \cdot 10^{-3}$	2.079

TABLE 2. Errors and rates for the Isoparametric approximation with degrees of freedom placed at standard, equidistant points.

7.3. **Divergence errors and pressure robustness.** We also compare the maximum divergence values computed using isoparametric approximation presented in this paper with those computed with the standard isoparametric approach. The degrees of freedom for both approximations are taken at the Gauss-Lobatto points so that the only difference is the use of the Piola transform in the velocity space. As shown in Figure 3, the method described in this paper is divergence free, whereas the standard isoparametric method ($u_h^{standard}$) is not.

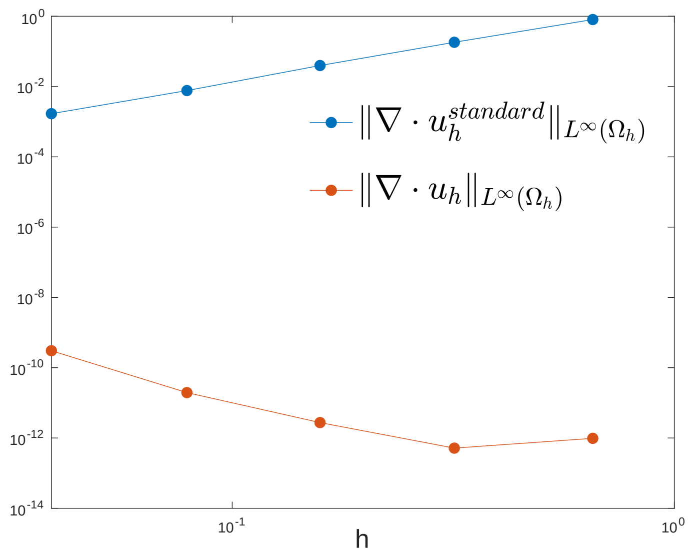

FIGURE 3. Divergence of the isoparametric method with Piola transform compared to the standard isoparametric method on Scott-Vogelius $\boldsymbol{\mathcal{P}}^3 - \mathcal{P}^2$ elements.

ν	$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega_h)}$	$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{H^1(\Omega_h)}$
10^{-7}	$1.139 \cdot 10^{-4}$	$9.046 \cdot 10^{-3}$
10^{-6}	$1.139 \cdot 10^{-4}$	$9.046 \cdot 10^{-3}$
10^{-3}	$1.139 \cdot 10^{-4}$	$9.046 \cdot 10^{-3}$
1	$1.139 \cdot 10^{-4}$	$9.046 \cdot 10^{-3}$

TABLE 3. Error tests for $h = 0.079$ with varying values of viscosity ν.

Finally, we check the behavior of the method for varying values of viscosity. We run the method on $\boldsymbol{\mathcal{P}}^3 - \mathcal{P}^2$ elements for data given by (7.1). In Table 3, we show the behavior of the error in the velocity as we vary viscosity ν. As we can see, the error remains nearly unchanged as we vary values of ν over several orders of magnitude, indicating that the scheme is pressure robust.

REFERENCES

[1] D. N. ARNOLD AND J. QIN, *Quadratic velocity/linear pressure stokes elements*, in Advances in Computer Methods for Partial Differential Equations-VII, R. Vichnevetsky, D. Knight, and G. Richter, eds., IMACS, 1992, pp. 28–34.
[2] C. BERNARDI, *Optimal finite-element interpolation on curved domains*, SIAM J. Numer. Anal., 26 (1989), pp. 1212–1240.
[3] S. C. BRENNER, M. NEILAN, AND L.-Y. SUNG, *Isoparametric C^0 interior penalty methods for plate bending problems on smooth domains*, Calcolo, 50 (2013), pp. 35–67.
[4] S. C. BRENNER AND L. R. SCOTT, *The mathematical theory of finite element methods*, vol. 3, Springer, 2008.

[5] P. G. Ciarlet and P.-A. Raviart, *Interpolation theory over curved elements, with applications to finite element methods*, Comput. Methods Appl. Mech. Engrg., 1 (1972), pp. 217–249.
[6] B. Cockburn, G. Kanschat, and D. Schötzau, *A note on discontinuous galerkin divergence-free solutions of the navier–stokes equations*, Journal of Scientific Computing, 31 (2007), pp. 61–73.
[7] C. M. Elliott and T. Ranner, *Finite element analysis for a coupled bulk-surface partial differential equation*, IMA J. Numer. Anal., 33 (2013), pp. 377–402.
[8] R. S. Falk and M. Neilan, *Stokes complexes and the construction of stable finite elements with pointwise mass conservation*, SIAM Journal on Numerical Analysis, 51 (2013), pp. 1308–1326.
[9] J. Guzmán and M. Neilan, *Inf-sup stable finite elements on barycentric refinements producing divergence–free approximations in arbitrary dimensions*, SIAM Journal on Numerical Analysis, 56 (2018), pp. 2826–2844.
[10] V. John, A. Linke, C. Merdon, M. Neilan, and L. G. Rebholz, *On the divergence constraint in mixed finite element methods for incompressible flows*, SIAM review, 59 (2017), pp. 492–544.
[11] T. Kato, M. Mitrea, G. Ponce, and M. Taylor, *Extension and representation of divergence-free vector fields on bounded domains*, Math. Res. Lett., 7 (2000), pp. 643–650.
[12] M. Lenoir, *Optimal isoparametric finite elements and error estimates for domains involving curved boundaries*, SIAM J. Numer. Anal., 23 (1986), pp. 562–580.
[13] H. Liu, M. Neilan, and M. Olshanskii, *A CutFEM divergence-free discretization for the Stokes problem*, ESAIM Math. Model. Numer. Anal., 57 (2023), pp. 143–165.
[14] H. Liu, M. Neilan, and M. B. Otus, *A divergence-free finite element method for the Stokes problem with boundary correction*, J. Numer. Math., 31 (2023), pp. 105–123.
[15] M. Neilan and B. Otus, *Divergence-free Scott-Vogelius elements on curved domains*, SIAM J. Numer. Anal., 59 (2021), pp. 1090–1116.
[16] L. R. Scott, *Finite element techniques for curved boundaries*, PhD thesis, Massachusetts Institute of Technology, 1973.
[17] L. R. Scott and M. Vogelius, *Norm estimates for a maximal right inverse of the divergence operator in spaces of piecewise polynomials*, ESAIM: Mathematical Modelling and Numerical Analysis, 19 (1985), pp. 111–143.

Appendix A. Proof of Lemma 3.2

Proof. Write $\boldsymbol{v}(x) = A_T\hat{\boldsymbol{v}}(\hat{x})$ for some $\hat{\boldsymbol{v}} \in \hat{\boldsymbol{V}}_k$. We then use Lemma 2.3, (2.5), and equivalence of norms to obtain

$$\begin{aligned}\|\boldsymbol{v}\|_{W^{\ell,p}(K)} &\le Ch_T^{2/p-\ell}\|A_T\hat{\boldsymbol{v}}\|_{W^{\ell,p}(\hat{K})}\\ &\le Ch_T^{2/p-\ell}\|A_T\|_{W^{j,\infty}(\hat{K})}\|\hat{\boldsymbol{v}}\|_{W^{\ell,p}(\hat{K})}\\ &\le Ch_T^{2/p-\ell-1}\|\hat{\boldsymbol{v}}\|_{W^{\ell,p}(\hat{K})} \le Ch_T^{2/p-\ell-1}\|\hat{\boldsymbol{v}}\|_{L^q(\hat{K})}.\end{aligned} \tag{A.1}$$

Likewise, we have

$$\|\hat{\boldsymbol{v}}\|_{L^q(\hat{K})} \le \|A_T^{-1}\|_{L^\infty(\hat{K})}\|A_T\hat{\boldsymbol{v}}\|_{L^q(\hat{K})} \le Ch_T^{1-2/q}\|\boldsymbol{v}\|_{L^q(K)}. \tag{A.2}$$

Combining (A.1)–(A.2) yields (3.1) for the case $m = 0$. The estimate (3.1) for general m then follows by standard arguments (cf. [4, Lemma 4.5.3]).

To prove (3.2), we first use (2.5):

$$|\boldsymbol{v}|_{W^{\ell,p}(K)} \le C\Big[\underbrace{h_T^{2/p+\ell}\sum_{r=0}^{k}h_T^{-2r}|A_T\hat{\boldsymbol{v}}|_{W^{r,p}(\hat{K})}}_{=:I} + \underbrace{h_T^{2/p+\ell}\sum_{r=k+1}^{\ell}h_T^{-2r}|A_T\hat{\boldsymbol{v}}|_{W^{r,p}(\hat{K})}}_{=:II}\Big].$$

To bound I, we use (2.5) once again to obtain

$$I \le h_T^{2/p+\ell}\sum_{r=0}^{k}h_T^{-2r}\cdot h_T^{r-2/p}\|\boldsymbol{v}\|_{W^{r,p}(K)} \le Ch_T^{\ell-k}\|\boldsymbol{v}\|_{W^{k,p}(K)}.$$

For II, we use the fact that $\hat{\boldsymbol{v}}$ is a polynomial of degree $\le k$ on K to obtain

$$\begin{aligned}
|A_T\hat{\boldsymbol{v}}|_{W^{r,p}(\hat{K})} &\le C\sum_{j=0}^{k} |A_T|_{W^{r-j,\infty}(\hat{K})}|\hat{\boldsymbol{v}}|_{W^{j,p}(\hat{K})}\\
&\le C\sum_{j=0}^{k} h_T^{r-j-1}|A_T^{-1}A_T\hat{\boldsymbol{v}}|_{W^{j,p}(\hat{K})}\\
&\le C\sum_{j=0}^{k}\sum_{i=0}^{j} h_T^{r-j-1}|A_T^{-1}|_{W^{j-i,\infty}(\hat{K})}|A_T\hat{\boldsymbol{v}}|_{W^{i,p}(\hat{K})}\\
&\le C\sum_{j=0}^{k}\sum_{i=0}^{j} h_T^{r-i}|A_T\hat{\boldsymbol{v}}|_{W^{i,p}(\hat{K})}\\
&\le C\sum_{j=0}^{k}\sum_{i=0}^{j} h_T^{r-i}\cdot h_T^{i-2/p}\|\boldsymbol{v}\|_{W^{i,p}(K)}\\
&\le Ch_T^{r-2/p}\|\boldsymbol{v}\|_{W^{k,p}(K)}.
\end{aligned}$$

Thus,

$$II \le Ch_T^{2/p+\ell}\sum_{r=k+1}^{\ell} h_T^{-r-2/p}\|\boldsymbol{v}\|_{W^{k,p}(K)} \le C\|\boldsymbol{v}\|_{W^{k,p}(K)}.$$

Combining the bounds for I and II completes the proof of (3.2). □

Appendix B. Proof of Lemma 4.5

Proof. Define $\boldsymbol{E}_h : \boldsymbol{V}^h \to \boldsymbol{H}_0^1(\Omega_h)$ such that, for $\boldsymbol{v} \in \boldsymbol{V}^h$,

$$\boldsymbol{E}_h\boldsymbol{v}|_T = (\tilde{\boldsymbol{v}} \circ F_{\tilde{T}} \circ F_T^{-1})|_T,$$

where $\tilde{\boldsymbol{v}}$ is the function in $\tilde{\boldsymbol{V}}$ uniquely defined by

$$\boldsymbol{v}|_T(a) = \tilde{\boldsymbol{v}}|_{\tilde{T}}(\tilde{a}) \quad \forall a \in \mathcal{N}_T, \quad \forall T \in \mathcal{T}_h,$$

where $T = G_h(\tilde{T})$. In other words, in a *standard* isoparametric, kth degree Lagrange finite element method, $\boldsymbol{E}_h\boldsymbol{v}$ would be the function on the isoparametric element associated with $\tilde{\boldsymbol{v}}$ on $\tilde{T}$. Thus, $\boldsymbol{E}_h\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega_h)$.

As shown in [15], $\tilde{\boldsymbol{v}} = \boldsymbol{E}_h\boldsymbol{v}$ on affine triangles, and we may conclude

$$\boldsymbol{E}_h\boldsymbol{v}|_T(a) = \boldsymbol{v}|_T(a) \quad \forall a \in \mathcal{N}_T, \quad \forall T \in \mathcal{T}_h.$$

Our goal is to estimate $\boldsymbol{v} - \boldsymbol{E}_h\boldsymbol{v}$, and our proof follows closely with the proof of Lemma 4.5 in [15]. However, here we provide a more general result.

As $\boldsymbol{v} = \boldsymbol{E}_h\boldsymbol{v}$ on affine triangles, we only consider $T \in \mathcal{T}_h$ with curved boundaries. Additionally, we know $\boldsymbol{v}|_{\partial T\cap\partial\Omega_h} = 0$. We may write $\boldsymbol{v}|_T(x) = A_T(\hat{x})\hat{\boldsymbol{v}}(\hat{x})$, for some $\hat{\boldsymbol{v}} \in \hat{\boldsymbol{V}}$, where $A_T = DF_T/\det(DF_T)$. Furthermore, there exists $\hat{\boldsymbol{w}} \in \hat{\boldsymbol{V}}$ such that $\hat{\boldsymbol{w}}(\hat{x}) = \boldsymbol{E}_h\boldsymbol{v}|_T(x)$. Consequently,

$$A_T(\hat{a})\hat{\boldsymbol{v}}(\hat{a}) = \hat{\boldsymbol{w}}(\hat{a}) \quad \forall \hat{a} \in \mathcal{N}_{\hat{T}},$$

so $\hat{\boldsymbol{w}}$ is the piecewise kth degree Lagrange interpolant of $A_T\hat{\boldsymbol{v}}$ on $\hat{T}^{CT}$.

By the Bramble-Hilbert lemma, we have

$$\|A_T\hat{\boldsymbol{v}} - \hat{\boldsymbol{w}}\|_{H^i(\hat{K})} \le C|A_T\hat{\boldsymbol{v}}|_{H^{k+1}(\hat{K})} \quad \forall \hat{K} \in \hat{T}^{CT}, \quad i = 0, 1, \ldots, k. \tag{B.1}$$

We may then bound the right-hand side using Lemma 2.3 and recognizing that $\hat{\boldsymbol{v}}$ is a polynomial of degree k. Thus we have

$$
\begin{aligned}
|A_T\hat{\boldsymbol{v}}|_{H^{k+1}(\hat{K})} \leq & C\sum_{j=0}^{k+1} |A_T|_{W^{k+1-j}(\hat{K})}|\hat{\boldsymbol{v}}|_{H^j(\hat{K})} = C\sum_{j=0}^{k} |A_T|_{W^{k+1-j}(\hat{K})}|\hat{\boldsymbol{v}}|_{H^j(\hat{K})} \\
\leq & C\sum_{j=0}^{k} h_T^{k-j}|\hat{\boldsymbol{v}}|_{H^j(\hat{K})}.
\end{aligned} \tag{B.2}
$$

Using Lemmas 2.3 and 2.4, we have

$$
\begin{aligned}
|\hat{\boldsymbol{v}}|_{H^j(\hat{K})} = |A_T^{-1}A_T\hat{\boldsymbol{v}}|_{H^j(\hat{K})} &\leq C\sum_{\ell=0}^{j} |A_T^{-1}|_{W^{j-\ell}(\hat{K})}|A_T\hat{\boldsymbol{v}}|_{H^\ell(\hat{K})} \\
&\leq C\sum_{\ell=0}^{j} h_T^{1+j-\ell}|A_T\hat{\boldsymbol{v}}|_{H^\ell(\hat{K})} \\
&\leq C\sum_{\ell=0}^{j} h_T^{1+j-\ell}h_T^{\ell-1}\|\boldsymbol{v}\|_{H^\ell(K)} \leq Ch_T^j\|\boldsymbol{v}\|_{H^j(K)}.
\end{aligned} \tag{B.3}
$$

Inserting this estimate into (B.2) yields

$$
|A_T\hat{\boldsymbol{v}}|_{H^{k+1}(\hat{K})} \leq Ch_T^k\|\boldsymbol{v}\|_{H^k(K)},
$$

and therefore by (B.1) and Lemma 2.4,

$$
|\boldsymbol{v} - \boldsymbol{E}_h\boldsymbol{v}|_{H^i(K)} \leq Ch_T^{1-i}\|A_T\hat{\boldsymbol{v}} - \hat{\boldsymbol{w}}\|_{H^i(\hat{K})} \leq Ch_T^{1-i}|A_T\hat{\boldsymbol{v}}|_{H^{k+1}(\hat{K})} \leq Ch_T^{k+1-i}\|\boldsymbol{v}\|_{H^k(K)}. \tag{B.4}
$$

An applicaiton of the inverse inequality (3.1) then yields the desired estimate (4.3). □

University of Pittsburgh, Department of Mathematics
Email address: RFD17@pitt.edu

University of Pittsburgh, Department of Mathematics
Email address: neilan@pitt.edu