DE GRUYTER J. Numer. Math. 2023; 31(2):105–123

Haoran Liu, Michael Neilan*, and M. Baris Otus

A divergence-free finite element method for the Stokes problem with boundary correction

https://doi.org/10.1515/jnma-2021-0125
Received November 03, 2021; revised March 17, 2022; accepted March 18, 2022

Abstract: This paper constructs and analyzes a boundary correction finite element method for the Stokes problem based on the Scott–Vogelius pair on Clough–Tocher splits. The velocity space consists of continuous piecewise polynomials of degree k, and the pressure space consists of piecewise polynomials of degree $(k-1)$ without continuity constraints. A Lagrange multiplier space that consists of continuous piecewise polynomials with respect to the boundary partition is introduced to enforce boundary conditions and to mitigate the lack of pressure-robustness. We prove several inf-sup conditions, leading to the well-posedness of the method. In addition, we show that the method converges with optimal order and the velocity approximation is divergence-free.

Keywords: finite elements, Stokes problem, boundary correction, divergence-free

Classification: 65N30, 65N12, 76M10

1 Introduction

Boundary correction methods are a broad class of unfitted finite element methods, i.e., methods in which the computational mesh does not conform to the physical domain Ω. In contrast to, e.g., isoparametric methods, in which a domain is approximated via curved elements, boundary correction methods generally solve a PDE in a polytopal interior domain and transfer boundary conditions in a way such that the scheme still maintains optimal order convergence. This polytopal approximation is, in general, not an $O(h^2)$ approximation to the physical domain and in particular, the polytope's vertices are not necessarily on the boundary of Ω. This approach can be advantageous for, e.g., dynamic problems with moving boundaries, as remeshing is not needed at each time step. Another feature of boundary correction methods, in contrast to other unfitted schemes, is the absence of 'cut elements' which may require special quadrature formula and algebraic stabilization. Boundary correction methods were first introduced and analyzed nearly 50 years ago [7] for the Poisson problem, and the technique has been improved and refined recently resulting in practical and robust implementations [2–4, 13, 28, 32] (see also [8, 19] for variants).

In this article, we construct a boundary correction finite element method for the Stokes problem based on the Scott–Vogelius pair on Clough–Tocher (or Alfeld) splits. The velocity approximation is sought in the space of continuous piecewise polynomials of degree k ($k \geqslant 2$) whereas the pressure space is approximated by piecewise polynomials of degree $(k-1)$ without continuity constraints. From their definitions, we see that the divergence operator maps the velocity space into the pressure space, and therefore, the scheme yields divergence-free velocity approximations. As far as we are aware this is the first H^1-conforming divergence-free finite element method for incompressible flow on unfitted meshes.

The construction and analysis of divergence-free methods is an active area of research, and many schemes have been proposed [1, 16, 17, 20, 24, 38]. These schemes have several inherent advantages, e.g., exact conservation laws for any mesh size and long-time stability [5, 12]. Another potential feature of these schemes, in the case Ω is a polytope, is pressure-robustness; similar to the continuous setting, modifying the source term in the Stokes problem by a gradient field only affects the pressure approximation. This invariance leads to a decoupling in the velocity error, with abstract estimates independent of the viscosity. Thus, divergence-free schemes may

***Corresponding author: Michael Neilan,** Department of Mathematics, University of Pittsburgh, Pittsburgh, PA 15260, USA.
Email: neilan@pitt.edu
Haoran Liu, M. Baris Otus, Department of Mathematics, University of Pittsburgh, Pittsburgh, PA 15260, USA.

be advantageous for high Reynold number flows and/or flows with large pressure gradients [27, 36, 37]. Most of these divergence-free methods are applied to polytopal domains; notable exemptions are the isoparametric method in [29], DG-type methods (e.g., [6, 25]), and isogeometric discretizations (e.g., [10, 14]).

Let us describe the scheme in more detail and briefly summarize the context of our results. The method starts with a background mesh enveloping the domain Ω, and the computational mesh simply consists of those elements fully contained in $\overline{\Omega}$. The method is based on a standard Nitsche-based formulation, where the Dirichlet boundary conditions are enforced via penalization. As the computational domain does not conform to Ω, boundary conditions are corrected via simple applications of Taylor's theorem to reduce the inconsistency of the scheme.

The procedure described so far is relatively standard for the Poisson problem (cf. [2–4, 7, 28]), but leads to some pressing issues for the Stokes equations. First, because the computational domain explicitly depends on the mesh parameter h, inf-sup stability of the Stokes pair is not immediately obvious. As explained in [18], the standard proof of inf-sup stability in the continuous setting (which is needed for the discrete result) is based on a decomposition of the computational domain into a finite number of strictly star shaped domains; the number of star shaped domains is generally unbounded as $h \to 0$. This problem can be circumvented with pressure-stabilization [3, 28], but at the price of additional consistency errors and poor conservation properties. We address this stability issue by carefully designing the computational mesh such that it inherits a macro element structure and applying the framework developed in [18] for Stokes pairs on unfitted domains. Doing so, we show that the resulting pair is uniformly stable on the unfitted domain with respect to the discretization parameter. As far as we are aware, this is the first uniform inf-sup stability result of a divergence-free Stokes pair on unfitted meshes.

The second difficulty of a boundary correction method for the Stokes problem is its lack of pressure-robustness, i.e., the discrete velocity approximation of these schemes depends on the irrotational part of the source function. This feature leads to a pressure-dependent velocity error estimate that scales with the inverse of the viscosity. The lack of pressure robustness is not due to the boundary correction per se, but rather due to the weak enforcement of boundary conditions via penalization. In particular, a divergence-free method for the Stokes problem with weak enforcement of the boundary conditions is *not* pressure robust. This stems from the simple fact that divergence-free functions with non-zero normal boundary conditions are not L^2-orthogonal to gradients. We mitigate the lack of pressure robustness in the scheme by introducing an additional Lagrange multiplier that enforces the boundary conditions of the normal component of the velocity. The Lagrange multiplier space consists of continuous piecewise polynomials of degree of k with respect to the boundary partition, and the Lagrange multiplier is an approximation to the pressure (modulo an additive constant) restricted to the computational boundary. The Lagrange multiplier ameliorates the lack of pressure robustness of the method and leads to a weakly coupled velocity error estimate; the velocity error's dependence on the viscosity is compensated by a higher-order power of the discretization parameter h.

The rest of the paper is organized as follows. In the next section, we state the Stokes problem, the computational mesh, and the boundary transfer operator. In Section 3, we state the finite element method and show that the scheme yields exactly divergence-free velocity approximations. Section 4 proves several inf-sup conditions and the well-posedness of the method. In Section 5, we prove optimal order convergence provided the exact solution is sufficiently smooth. Finally, in Section 6 we perform some numerical experiments which verify the theoretical results, and give some concluding remarks in Section 7.

2 Preliminaries

For a two-dimensional bounded domain $\Omega \subset \mathbb{R}^2$, we consider the Stokes problem

$$-\nu\Delta \boldsymbol{u} + \nabla p = \boldsymbol{f} \qquad \text{in } \Omega \tag{2.1a}$$

$$\operatorname{div} \boldsymbol{u} = 0 \qquad \text{in } \Omega \tag{2.1b}$$

$$\boldsymbol{u} = \boldsymbol{g} \qquad \text{on } \partial\Omega \tag{2.1c}$$

where $\nu > 0$ is the viscosity, assumed to be constant. For simplicity in the presentation, and without loss of generality, we assume that $\boldsymbol{g} = 0$. The extension to non-homogeneous boundary conditions is relatively straightforward [21].

We assume the domain has smooth boundary $\partial\Omega$ with outward unit normal $\boldsymbol{n}$. We denote by φ the signed distance function of $\partial\Omega$ such that $\varphi(x) < 0$ for $x \in \Omega$ and $\varphi(x) \geqslant 0$ otherwise, so that $\boldsymbol{n} = \nabla\varphi/|\nabla\varphi|$ on $\partial\Omega$. For a positive number τ, denote by $\Gamma_\tau = \{x \in \mathbb{R}^2 : |\varphi(x)| \leqslant \tau|\}$ the tubular region around $\partial\Omega$. By [15, Lem. 14.16], there exists $\tau_0 > 0$ such the closest point projection $\boldsymbol{p} : \Gamma_{\tau_0} \to \partial\Omega$ is well defined and satisfies $\boldsymbol{p}(x) = x - \varphi(x)\boldsymbol{n}(\boldsymbol{p}(x))$ for all $x \in \Gamma_{\tau_0}$ (see [11]).

Let $S \subset \mathbb{R}^2$ be a polygon such that $\Omega \subset S$, and let $\mathcal{S}_h$ be a shape-regular simplicial triangulation of S. We define the computational mesh as

$$\mathcal{T}_h = \{T \in \mathcal{S}_h : \overline{T} \subset \overline{\Omega}\}$$

and set

$$\Omega_h = \operatorname{int}\Big(\bigcup_{T\in\mathcal{T}_h} \overline{T}\Big) \subset \Omega$$

to be the associated domain. We denote by $\mathcal{T}_h^{ct}$ the Clough–Tocher refinement of $\mathcal{T}_h$, obtained by connecting the vertices of each $T \in \mathcal{T}_h$ to its barycenter. The set of boundary edges of $\mathcal{T}_h$, which is also the set of boundary edges of $\mathcal{T}_h^{ct}$, is denoted by $\mathcal{E}_h^B$. With an abuse of notation, for a piecewise smooth function q (with respect to $\mathcal{E}_h^B$), we write

$$\int_{\partial\Omega_h} q \,\mathrm{d}s = \sum_{e\in\mathcal{E}_h^B} \int_e q \,\mathrm{d}s.$$

We use $\boldsymbol{n}_h$ to denote the outward unit normal with respect to the computational boundary $\partial\Omega_h$. For $K \in \mathcal{T}_h^{ct}$, we set $h_K = \operatorname{diam}(K)$ and $h = \max_{K\in\mathcal{T}_h^{ct}} h_K$. Likewise, for $e \in \mathcal{E}_h^B$, we set $h_e = \operatorname{diam}(e)$.

Remark 2.1. Denote by $\mathcal{S}_h^{ct}$ the Clough–Tocher refinement of the background mesh $\mathcal{S}_h$. We emphasize that $\mathcal{T}_h^{ct} \subset \mathcal{S}_h^{ct}$, however,

$$\mathcal{T}_h^{ct} \neq \{K \in \mathcal{S}_h^{ct} : \overline{K} \subset \overline{\Omega}\}.$$

In particular, $\mathcal{T}_h^{ct}$ inherits the macro-element structure needed to prove the stability of the Scott–Vogelius pair.

2.1 Boundary transfer operator

The main component of boundary correction methods is a well-defined mapping $M : \partial\Omega_h \to \partial\Omega$ that assigns each point on the computational boundary to physical one in order to ‘transfer’ the boundary information on $\partial\Omega$ to $\partial\Omega_h$. With such a mapping in hand, we can define the transfer direction as

$$\eth(x) = (M - I)x, \qquad x \in \partial\Omega_h$$

and transfer length

$$\delta(x) = |\eth(x)|. \tag{2.2}$$

Several choices of the mapping M and corresponding transfer directions have appeared in the literature. A common choice (and arguably the most natural) is to take M to be the closest point projection, i.e., $M = \boldsymbol{p}$. In this case, assuming Ω_h approximates Ω well enough, the distance vector $\eth$ defined above coincides (up to a multiplicative constant) with the outward unit normal vector $\boldsymbol{n}$ of the original boundary $\partial\Omega$. In particular, there holds $\eth(x) = -\varphi(x)\boldsymbol{n}(\boldsymbol{p}(x))$ and $\delta(x) = |\varphi(x)|$. Another common choice is to take the transfer direction to be parallel to the outward unit normal of the computational boundary, i.e., $\eth/\delta = \boldsymbol{n}_h$. In this case, we have $\delta(x) \geqslant |\varphi(x)|$ with possible large discrepancies between $\delta(x)$ and $|\varphi(x)|$, but it leads to a simpler implementation in the numerical method.

In the definition and analysis of the method below, we do not explicitly define the mapping M; rather, our main requirement for the mapping M is to satisfy the assumption (A) below. In particular, and similar to [2–4, 7, 11, 32], the stability and convergence analysis only assumes that the transfer distance $\delta(x)$ is sufficiently small relative to the mesh parameter h. In the numerical experiments provided in Section 6, we take M to be an approximation to the closest point projection.

Set $\boldsymbol{d} = \mathfrak{d}/\delta$, and for $x \in \partial\Omega_h$ define the boundary transfer operator

$$(S_h \boldsymbol{v})(x) = \sum_{j=0}^{k} \frac{1}{j!}(\delta(x))^j \frac{\partial^i \boldsymbol{v}}{\partial \boldsymbol{d}^j}(x).$$

Note that $(S_h \boldsymbol{v})(x)$ is the k^{th}-order Taylor expansion of the function $\boldsymbol{v}$.

Remark 2.2. Throughout this paper, the constants C and c (with or without subscripts) denote some positive constants that are independent of the mesh parameter h and the viscosity that may change values at each occurrence.

3 A divergence-free finite element method

For $D \subset \mathbb{R}^2$, denote by $\mathcal{P}_s(D)$ the space of polynomials of degree $\leqslant s$ with domain D. Analogous vector-valued spaces are denoted in boldface. For $k \geqslant 2$, we define the corresponding Scott–Vogelius finite element pair with respect to the Clough–Tocher triangulation $\mathcal{T}_h^{ct}$:

$$\boldsymbol{V}_h = \left\{ \boldsymbol{v} \in \boldsymbol{H}^1(\Omega_h) : \ \boldsymbol{v}|_K \in \boldsymbol{\mathcal{P}}_k(K) \ \ \forall K \in \mathcal{T}_h^{ct}, \ \int_{\partial\Omega_h} (\boldsymbol{v} \cdot \boldsymbol{n}_h) \, \mathrm{d}s = 0 \right\}$$
$$Q_h = \left\{ q \in L^2(\Omega_h) : \ q|_K \in \mathcal{P}_{k-1}(K) \ \ \forall K \in \mathcal{T}_h^{ct} \right\}$$

and the analogous spaces with boundary conditions

$$\mathring{\boldsymbol{V}}_h = \boldsymbol{V}_h \cap \boldsymbol{H}_0^1(\Omega_h), \qquad \mathring{Q}_h = Q_h \cap L_0^2(\Omega_h).$$

We further introduce a Lagrange multiplier space

$$X_h = \left\{ \mu \in C(\partial\Omega_h) : \ \mu|_e \in \mathcal{P}_k(e) \, \forall e \in \mathcal{E}_h^B \right\}$$

and its variant,

$$\mathring{X}_h = \left\{ \mu \in X_h : \ \int_{\partial\Omega_h} \mu \, \mathrm{d}s = 0 \right\}.$$

We define the bilinear form

$$a_h(\boldsymbol{u}, \boldsymbol{v}) = \nu \bigg(\int_{\Omega_h} \nabla \boldsymbol{u} : \nabla \boldsymbol{v} \, \mathrm{d}x - \int_{\partial\Omega_h} \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{n}_h} \cdot \boldsymbol{v} \, \mathrm{d}s + \int_{\partial\Omega_h} \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{n}_h} \cdot (S_h \boldsymbol{u}) \, \mathrm{d}s + \sum_{e \in \mathcal{E}_h^B} \int_e \frac{\sigma}{h_e} (S_h \boldsymbol{u}) \cdot (S_h \boldsymbol{v}) \, \mathrm{d}s \bigg) \tag{3.1}$$

where $\sigma > 0$ is a penalty parameter.

Remark 3.1. The bilinear form $a_h(\cdot, \cdot)$ is based on a standard 'Nitsche bilinear form' associated with the Laplace operator, but with boundary correction to improve the consistency of the scheme (cf. Lemma 5.1). In particular, if the boundary correction operator S_h in (3.1) is replaced with the identity operator, then we recover the Nitsche bilinear form [30, 35]. Note that the bilinear form is based on a non-symmetric version of Nitsche's method due to the positive sign in front of the third term in the bilinear form $a_h(\cdot, \cdot)$. However, boundary correction methods based on the symmetric version of Nitsche's method still yield a non-symmetric bilinear form [7, 28]. The non-symmetric version allows less restrictions on the penalty parameter σ to ensure stability. In particular, Lemma 4.3 below shows the bilinear form $a_h(\cdot, \cdot)$ is coercive on $\boldsymbol{V}_h$ for any $\sigma > 0$.

We define two bilinear forms associated with the continuity equations, one without and one with boundary correction:

$$b_h(\boldsymbol{v},(q,\mu)) = -\int_{\Omega_h} (\operatorname{div}\boldsymbol{v})q\,\mathrm{d}x + \int_{\partial\Omega_h} (\boldsymbol{v}\cdot\boldsymbol{n}_h)\mu\,\mathrm{d}s$$

$$b_h^e(\boldsymbol{v},(q,\mu)) = -\int_{\Omega_h} (\operatorname{div}\boldsymbol{v})q\,\mathrm{d}x + \int_{\partial\Omega_h} ((S_h\boldsymbol{v})\cdot\boldsymbol{n}_h)\mu\,\mathrm{d}s.$$

We consider the method of finding $(\boldsymbol{u}_h, p_h, \lambda_h) \in \boldsymbol{V}_h \times \mathring{Q}_h \times \mathring{X}_h$ such that

$$a_h(\boldsymbol{u}_h,\boldsymbol{v}) + b_h(\boldsymbol{v},(p_h,\lambda_h)) = \int_{\Omega_h} \boldsymbol{f}\cdot\boldsymbol{v}\,\mathrm{d}x \quad \forall\,\boldsymbol{v}\in\boldsymbol{V}_h \tag{3.2a}$$

$$b_h^e(\boldsymbol{u}_h,(q,\mu)) = 0 \qquad \forall\,(q,\mu)\in\mathring{Q}_h\times\mathring{X}_h. \tag{3.2b}$$

Remark 3.2. The zero mean-value constant defined in the Lagrange multiplier space $\mathring{X}_h$ mods out constants, and is due to the condition $\int_{\partial\Omega_h}(\boldsymbol{v}\cdot\boldsymbol{n}_h)\,\mathrm{d}s = 0$ in the definition of the discrete velocity space $\boldsymbol{V}_h$. If this constraint is not imposed in the Lagrange multiplier space, then in general (3.2) is ill-posed since

$$b_h(\boldsymbol{v},(0,1)) = 0 \quad \forall\boldsymbol{v}\in\boldsymbol{V}_h.$$

On the other hand, the constraint $\int_{\partial\Omega_h}(\boldsymbol{v}\cdot\boldsymbol{n}_h)\,\mathrm{d}s = 0$ is needed to ensure that method (3.2) yields a divergence-free solution, as the next lemma shows.

Lemma 3.1 (divergence-free property). *If* $(\boldsymbol{u}_h, p_h, \lambda_h) \in \boldsymbol{V}_h \times \mathring{Q}_h \times \mathring{X}_h$ *satisfies* (3.2)*, then* $\operatorname{div}\boldsymbol{u}_h \equiv 0$ *in* Ω_h.

Proof. The definition of the Stokes pair $\boldsymbol{V}_h \times \mathring{Q}_h$ shows $\operatorname{div}\boldsymbol{u}_h \in \mathring{Q}_h$. Then, letting $q = \operatorname{div}\boldsymbol{u}_h$ and $\mu = 0$ in (3.2b) yields

$$0 = b_h^e(\boldsymbol{u}_h,(\operatorname{div}\boldsymbol{u}_h,0)) = -\|\operatorname{div}\boldsymbol{u}_h\|^2_{L^2(\Omega_h)}.$$

Thus, $\operatorname{div}\boldsymbol{u}_h \equiv 0$. □

4 Stability and continuity estimates

In our stability and convergence analysis, we make an assumption regarding the distance between the PDE domain Ω and the computational domain Ω_h. To state this assumption, we define for a boundary edge $e \in \mathcal{E}_h^B$,

$$\delta_e := \max_{x\in\bar{e}} \delta(x).$$

We make the assumption

$$\max_{e\in\mathcal{E}_h^B} h_e^{-1}\delta_e \leqslant c_\delta < 1 \quad \text{ for } c_\delta \text{ sufficiently small.} \tag{A}$$

Remark 4.1. Assumption (A) essentially states that the distance between $\partial\Omega$ and $\partial\Omega_h$ is of order h, i.e., $\delta = O(h)$ with (hidden) constant sufficiently small. Similar assumptions, in the context of boundary correction methods, are made in, e.g., [2, 3, 7, 28, 31]. As explained in [2, Rem. 3], the condition can be satisfied in practice by shifting the location of the nodes on the computational boundary along the direction $\boldsymbol{n}$. On the other hand, while assumption (A) is crucially used in the stability analysis (cf. Lemma 4.3 and Theorem 4.2), the numerical experiments presented in Section 6 indicate that the smallness assumption of c_δ can be relaxed, and a shifting of nodes on the computational boundary is not needed to ensure stability.

We define three H^1-type norms on $\boldsymbol{V}_h + \boldsymbol{H}^{k+1}(\Omega)$:

$$\|\boldsymbol{v}\|_h^2 = \|\nabla \boldsymbol{v}\|_{L^2(\Omega_h)}^2 + \sum_{e\in\mathcal{E}_h^B} h_e^{-1}\|S_h\boldsymbol{v}\|_{L^2(e)}^2$$
$$\|\boldsymbol{v}\|_{1,h}^2 = \|\nabla \boldsymbol{v}\|_{L^2(\Omega_h)}^2 + \sum_{e\in\mathcal{E}_h^B} h_e^{-1}\|\boldsymbol{v}\|_{L^2(e)}^2$$
$$|||\boldsymbol{v}|||_h^2 = \|\boldsymbol{v}\|_h^2 + \sum_{e\in\mathcal{E}_h^B} h_e\|\nabla\boldsymbol{v}\|_{L^2(e)}^2.$$

In addition, we define a $H^{-1/2}$-type norm on the Lagrange multiplier space $\mathring{X}_h$:

$$\|\mu\|_{-1/2,h}^2 = \sum_{e\in\mathcal{E}_h^B} h_e\|\mu\|_{L^2(e)}^2.$$

Finally, we define the norm on $\mathring{Q}_h \times \mathring{X}_h$ as

$$\|(q,\mu)\| := \|q\|_{L^2(\Omega_h)} + \|\mu\|_{-1/2,h}.$$

Lemma 4.1. *Assuming* (A), *there holds for all* $\boldsymbol{v}\in\boldsymbol{V}_h$,

$$\begin{aligned}\sum_{e\in\mathcal{E}_h^B} h_e^{-1}\|S_h\boldsymbol{v}-\boldsymbol{v}\|_{L^2(e)}^2 &\leqslant Cc_\delta^2\|\nabla\boldsymbol{v}\|_{L^2(\Omega_h)}^2\\ \sum_{e\in\mathcal{E}_h^B} h_e^{-1}\|S_h\boldsymbol{v}\|_{L^2(e)}^2 &\leqslant C\|\boldsymbol{v}\|_{1,h}^2.\end{aligned} \tag{4.1}$$

In particular, $\|\cdot\|_h$, $\|\cdot\|_{1,h}$, *and* $|||\cdot|||_h$ *are equivalent on* $\boldsymbol{V}_h$.

Proof. By trace and inverse inequalities, the shape-regularity of $\mathcal{T}_h$ and (A), there holds for $e\in\mathcal{E}_h^B$,

$$h_e^{-1}\int_e |\delta|^{2j}\Big|\frac{\partial^j\boldsymbol{v}}{\partial\boldsymbol{d}^j}\Big|^2\,\mathrm{d}s \leqslant C\delta_e^{2j}h_e^{-2j}\|\nabla\boldsymbol{v}\|_{L^2(T_e)}^2 \leqslant Cc_\delta^{2j}\|\nabla\boldsymbol{v}\|_{L^2(T_e)}^2, \qquad j=1,2,\ldots,k \tag{4.2}$$

where $T_e\in\mathcal{T}_h$ satisfies $e\subset\partial T$. The estimate (4.2) implies the first inequality in (4.1). The estimate (4.2) also implies the second inequality in (4.1) since

$$\sum_{e\in\mathcal{E}_h^B} h_e^{-1}\|S_h\boldsymbol{v}\|_{L^2(e)}^2 \leqslant C\sum_{e\in\mathcal{E}_h^B}\sum_{j=0}^{k} h_e^{-1}\int_e|\delta|^{2j}\Big|\frac{\partial^j\boldsymbol{v}}{\partial\boldsymbol{d}^j}\Big|^2\,\mathrm{d}s \leqslant C\|\boldsymbol{v}\|_{1,h}^2.$$

The second inequality in (4.1) immediately yields $\|\boldsymbol{v}\|_h\leqslant C\|\boldsymbol{v}\|_{1,h}$. Moreover, standard arguments involving the trace and inverse inequalities show $\|\boldsymbol{v}\|_h\leqslant|||\boldsymbol{v}|||_h\leqslant C\|\boldsymbol{v}\|_h$ on $\boldsymbol{V}_h$. Thus, to complete the proof, it suffices to show $\|\boldsymbol{v}\|_{1,h}\leqslant C\|\boldsymbol{v}\|_h$.

To this end, we once again use (4.2) to obtain

$$\begin{aligned}\sum_{e\in\mathcal{E}_h^B} h_e^{-1}\|\boldsymbol{v}\|_{L^2(e)}^2 &\leqslant 2\sum_{e\in\mathcal{E}_h^B} h_e^{-1}\|S_h\boldsymbol{v}\|_{L^2(e)}^2 + 2\sum_{e\in\mathcal{E}_h^B} h_e^{-1}\|S_h\boldsymbol{v}-\boldsymbol{v}\|_{L^2(e)}^2\\ &\leqslant 2\sum_{e\in\mathcal{E}_h^B} h_e^{-1}\|S_h\boldsymbol{v}\|_{L^2(e)}^2 + C\sum_{e\in\mathcal{E}_h^B} h_e^{-1}\sum_{j=1}^{k}\int_e|\delta|^{2j}\Big|\frac{\partial^j\boldsymbol{v}}{\partial\boldsymbol{d}^j}\Big|^2\,\mathrm{d}s\\ &\leqslant 2\sum_{e\in\mathcal{E}_h^B} h_e^{-1}\|S_h\boldsymbol{v}\|_{L^2(e)}^2 + C\|\nabla\boldsymbol{v}\|_{L^2(\Omega_h)}^2.\end{aligned}$$

This inequality implies $\|\boldsymbol{v}\|_{1,h}\leqslant C\|\boldsymbol{v}\|_h$. □

4.1 Continuity and coercivity estimates of bilinear forms

Lemma 4.2. *Assuming* (A), *there holds*

$$|a_h(\boldsymbol{v},\boldsymbol{w})| \leqslant c_2(1+\sigma)\nu|||\boldsymbol{v}|||_h|||\boldsymbol{w}|||_h \qquad \forall \boldsymbol{v},\boldsymbol{w}\in \boldsymbol{V}_h+\boldsymbol{H}^{k+1}(\Omega_h) \tag{4.3}$$

$$|b_h(\boldsymbol{v},(q,\mu))| \leqslant C\|\boldsymbol{v}\|_{1,h}\|(q,\mu)\| \qquad \forall (q,\mu)\in \mathring{Q}_h\times\mathring{X}_h \tag{4.4}$$

$$|b_h(\boldsymbol{v},(q,\mu))-b_h^e(\boldsymbol{v},(q,\mu))| \leqslant Cc_\delta\|\boldsymbol{v}\|_{1,h}\|(q,\mu)\| \qquad \forall \boldsymbol{v}\in\boldsymbol{V}_h,\quad \forall (q,\mu)\in \mathring{Q}_h\times\mathring{X}_h \tag{4.5}$$

$$|b_h^e(\boldsymbol{v},(q,\mu))| \leqslant C(1+c_\delta)\|\boldsymbol{v}\|_{1,h}\|(q,\mu)\| \qquad \forall (q,\mu)\in \mathring{Q}_h\times\mathring{X}_h. \tag{4.6}$$

Proof. The proof of the continuity estimate of (4.3) is given in [3, Prop. 2] (with superficial modifications). The continuity estimate of $b_h(\cdot,\cdot)$ (4.4) follows directly from the Cauchy–Schwarz inequality.

This third estimate (4.5) follows from the definition of the forms, the Cauchy–Schwarz inequality, and (4.1):

$$\begin{aligned}|b_h(\boldsymbol{v},(q,\mu))-b_h^e(\boldsymbol{v},(q,\mu))| &= \Big|\sum_{e\in\mathcal{E}_h^B}\int_e \big((\boldsymbol{v}-S_h\boldsymbol{v})\cdot\boldsymbol{n}_h\big)\mu\,\mathrm{d}s\Big|\\ &\leqslant \Big(\sum_{e\in\mathcal{E}_h^B} h_e^{-1}\|\boldsymbol{v}-S_h\boldsymbol{v}\|^2_{L^2(e)}\Big)^{1/2}\|\mu\|_{-1/2,h}\\ &\leqslant Cc_\delta\|\boldsymbol{v}\|_{1,h}\|\mu\|_{-1/2,h}.\end{aligned}$$

The estimate (4.6) follows from the estimates (4.4) and (4.5) using the triangle inequality. □

Lemma 4.3. *Suppose that assumption* (A) *is satisfied for c_δ sufficiently small. Then there holds,*

$$c_1\nu\|\boldsymbol{v}\|^2_{1,h}\leqslant a_h(\boldsymbol{v},\boldsymbol{v}) \quad \forall \boldsymbol{v}\in\boldsymbol{V}_h$$

for $c_1>0$ independent of h and ν, and for any positive penalty parameter $\sigma>0$.

Proof. By definition of the bilinear form $a_h(\cdot,\cdot)$,

$$a_h(\boldsymbol{v},\boldsymbol{v}) = \nu\left(\|\nabla\boldsymbol{v}\|^2_{L^2(\Omega_h)}+\sum\nolimits_{e\in\mathcal{E}_h^B}\left(\int_e\frac{\partial\boldsymbol{v}}{\partial\boldsymbol{n}_h}\cdot(S_h\boldsymbol{v}-\boldsymbol{v})\,\mathrm{d}s+\frac{\sigma}{h_e}\|S_h\boldsymbol{v}\|^2_{L^2(e)}\right)\right).$$

A discrete trace inequality with (4.1) yields

$$\left|\sum\nolimits_{e\in\mathcal{E}_h^B}\int_e\frac{\partial\boldsymbol{v}}{\partial\boldsymbol{n}_h}\cdot(S_h\boldsymbol{v}-\boldsymbol{v})\,\mathrm{d}s\right|\leqslant Cc_\delta\|\nabla\boldsymbol{v}\|^2_{L^2(\Omega_h)}. \tag{4.7}$$

Thus, we find

$$a_h(\boldsymbol{v},\boldsymbol{v})\geqslant\nu\left((1-Cc_\delta)\|\nabla\boldsymbol{v}\|^2_{L^2(\Omega_h)}+\sum\nolimits_{e\in\mathcal{E}_h^B}\frac{\sigma}{h_e}\|S_h\boldsymbol{v}\|^2_{L^2(e)}\right)\geqslant C\nu\|\boldsymbol{v}\|^2_h\geqslant C\nu\|\boldsymbol{v}\|^2_{1,h}$$

for c_δ sufficiently small and for $\sigma>0$. □

Remark 4.2. The smallness of c_δ in Lemma 4.3 depends on the constant C in estimate (4.7), which itself depends on estimate (4.1) and discrete trace inequalities. The discrete trace and inverse estimates found in [22, 33], yield the explicit estimate

$$c_\delta<C_\dagger^{-1},\qquad C_\dagger:=\sqrt{\left(\sum\nolimits_{j=1}^{k}\prod\nolimits_{\ell=k-j+1}^{k-1}C_\ell\right)\frac{(k+1)(k+2)}{2}} \tag{4.8}$$

to ensure the coercivity of a_h. Here, $C_\ell>0$ is the maximum eigenvalue of a matrix defined in [33, Sect. 3], which numerically scales as $O(\ell^4)$. Details of the estimate (4.8) are found in Appendix A.

4.2 Inf-sup stability, I

In this section we prove the discrete inf-sup (LBB) condition for the Stokes pair $\mathring{\boldsymbol{V}}_h\times\mathring{Q}_h$ with stability constants independent of h. In the case of a fixed polygonal domain, the LBB stability for this pair is well-known (cf. [1, 16, 34]); however, the extension of these results to the unfitted domain Ω_h is not immediate. In particular, the proofs in [1, 16, 34] (directly or indirectly) rely on the Nečas inequality:

$$\mathfrak{c}_h \|q\|_{L^2(\Omega_h)} \leqslant \sup_{\boldsymbol{v}\in \boldsymbol{H}_0^1(\Omega_h)\setminus\{0\}} \frac{\int_{\Omega_h} (\operatorname{div} \boldsymbol{v}) q \,\mathrm{d}x}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega_h)}} \quad \forall q \in L_0^2(\Omega_h)$$

for some $\mathfrak{c}_h > 0$ depending on the domain Ω_h. As explained in [18], it is unclear if the constant $\mathfrak{c}_h$ in this inequality is independent of h.

Our approach is to combine the local stability of the Scott–Vogelius pair with the stability of the $\boldsymbol{\mathcal{P}}_k \times \mathcal{P}_{k-2}$ pair. For a (macro) element $T \in \mathcal{T}_h$, we define the local spaces with boundary conditions

$$\boldsymbol{V}_0(T) = \{\boldsymbol{v} \in \boldsymbol{H}_0^1(T) : \boldsymbol{v}|_K \in \boldsymbol{\mathcal{P}}_k(K) \ \ \forall K \subset T, \ \ K \in \mathcal{T}_h^{ct}\}$$
$$Q_0(T) = \{q \in L_0^2(T) : q|_K \in \mathcal{P}_{k-1}(K) \ \ \forall K \subset T, \ \ K \in \mathcal{T}_h^{ct}\}.$$

We state a local surjectivity of the divergence operator acting on these spaces. The proof is found, e.g., in [16].

Lemma 4.4. *For every $q \in Q_0(T)$, there exists $\boldsymbol{v} \in \boldsymbol{V}_0(T)$ such that* $\operatorname{div} \boldsymbol{v} = q$ *and* $\|\nabla \boldsymbol{v}\|_{L^2(T)} \leqslant \beta_T^{-1} \|q\|_{L^2(T)}$. *Here, the constant $\beta_T > 0$ depends only on the shape-regularity of T.*

Next, we state the recent stability result of the $\boldsymbol{\mathcal{P}}_k \times \mathcal{P}_{k-2}$ pair on unfitted domains (cf. [18, Thm. 1, Sect. 6.3, Rem. 1]).

Lemma 4.5. *Define the space of piecewise polynomials of degree $(k-2)$ with respect to the mesh $\mathcal{T}_h$:*

$$\mathring{Y}_h = \{q \in L_0^2(\Omega_h) : q|_T \in \mathcal{P}_{k-2}(T) \ \ \forall T \in \mathcal{T}_h\} \subset \mathring{Q}_h.$$

There exist $\beta_0 > 0$ and $h_0 > 0$ such that for $h \leqslant h_0$, there holds

$$\sup_{\boldsymbol{v}\in \mathring{\boldsymbol{V}}_h\setminus\{0\}} \frac{\int_{\Omega_h} (\operatorname{div} \boldsymbol{v}) q \,\mathrm{d}x}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega_h)}} \geqslant \beta_0 \|q\|_{L^2(\Omega_h)} \quad \forall q \in \mathring{Y}_h.$$

Combining Lemmas 4.4–4.5 yields the following stability result for the $\mathring{\boldsymbol{V}}_h \times \mathring{Q}_h$ Stokes pair.

Lemma 4.6. *There exists $\beta_1 > 0$ independent of h such that*

$$\sup_{\boldsymbol{v}\in \mathring{\boldsymbol{V}}_h\setminus\{0\}} \frac{\int_{\Omega_h} (\operatorname{div} \boldsymbol{v}) q \,\mathrm{d}x}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega_h)}} \geqslant \beta_1 \|q\|_{L^2(\Omega_h)} \quad \forall q \in \mathring{Q}_h$$

for $h \leqslant h_0$.

Proof. We combine Lemmas 4.4–4.5 with the arguments in [1, 16, 34].

Let $q \in \mathring{Q}_h$, and let $\overline{q} \in \mathring{Y}_h$ be its piecewise average, i.e., $\overline{q}|_T = |T|^{-1} \int_T q \,\mathrm{d}x$ for all $T \in \mathcal{T}_h$. We then have $(q-\overline{q})|_T \in Q_0(T)$ for all $T \in \mathcal{T}_h$, and therefore, by Lemma 4.4, there exists $\boldsymbol{v}_{1,T} \in \boldsymbol{V}_0(T)$ such that $\operatorname{div} \boldsymbol{v}_{1,T} = (q-\overline{q})|_T$ and $\|\nabla \boldsymbol{v}\|_{L^2(T)} \leqslant \beta_T^{-1} \|q\|_{L^2(T)}$. Defining $\boldsymbol{v}_1 \in \mathring{\boldsymbol{V}}_h$ by $\boldsymbol{v}_1|_T = \boldsymbol{v}_{1,T}$ for all $T \in \mathcal{T}_h$, we have $\operatorname{div} \boldsymbol{v}_1 = (q - \overline{q})$ in Ω_h and $\|\nabla \boldsymbol{v}_1\|_{L^2(\Omega_h)} \leqslant \beta_*^{-1} \|q - \overline{q}\|_{L^2(\Omega_h)}$, where $\beta_* = \min_{T\in\mathcal{T}_h} \beta_T$.

With this result, and by Lemma 4.5, we conclude

$$\begin{aligned}
\beta_0 \|\overline{q}\|_{L^2(\Omega_h)} &\leqslant \sup_{\boldsymbol{v}\in \mathring{\boldsymbol{V}}_h\setminus\{0\}} \frac{\int_{\Omega_h} (\operatorname{div} \boldsymbol{v}) \overline{q} \,\mathrm{d}x}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega_h)}} \\
&\leqslant \sup_{\boldsymbol{v}\in \mathring{\boldsymbol{V}}_h\setminus\{0\}} \frac{\int_{\Omega_h} (\operatorname{div} \boldsymbol{v}) q \,\mathrm{d}x}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega_h)}} + \|q - \overline{q}\|_{L^2(\Omega_h)} \\
&\leqslant (1 + \beta_*^{-1}) \sup_{\boldsymbol{v}\in \mathring{\boldsymbol{V}}_h\setminus\{0\}} \frac{\int_{\Omega_h} (\operatorname{div} \boldsymbol{v}) q \,\mathrm{d}x}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega_h)}}.
\end{aligned}$$

Thus,

$$\|q\|_{L^2(\Omega_h)} \leqslant \|q - \overline{q}\|_{L^2(\Omega_h)} + \|\overline{q}\|_{L^2(\Omega_h)} \leqslant \big(\beta_*^{-1} + \beta_0^{-1}(1 + \beta_*^{-1})\big) \sup_{\boldsymbol{v}\in \mathring{\boldsymbol{V}}_h\setminus\{0\}} \frac{\int_{\Omega_h} (\operatorname{div} \boldsymbol{v}) q \,\mathrm{d}x}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega_h)}}.$$

Setting $\beta_1 = \big(\beta_*^{-1} + \beta_0^{-1}(1 + \beta_*^{-1})\big)^{-1}$ completes the proof. □

4.3 Inf-sup stability, II

The following lemma proves inf-sup stability for the Lagrange multiplier part of the bilinear form $b_h(\cdot,\cdot)$.

Lemma 4.7. *Assume the triangulation $\mathcal{T}_h$ is quasi-uniform. Then there holds*

$$\sup_{\boldsymbol{v}\in\boldsymbol{V}_h\setminus\{0\}} \frac{\int_{\partial\Omega_h}(\boldsymbol{v}\cdot\boldsymbol{n}_h)\mu\,\mathrm{d}s}{\|\boldsymbol{v}\|_{1,h}} \geqslant \beta_2\|\mu\|_{-1/2,h} \quad \forall\mu\in\mathring{X}_h \tag{4.9}$$

for some $\beta_2>0$ independent of h.

Proof. We label the boundary edges as $\{e_j\}_{j=1}^N = \mathcal{E}_h^B$, and denote the boundary vertices by $\{a_j\}_{j=1}^N = \mathcal{V}_h^B$, labeled such that e_j has vertices a_j and a_{j+1}, with the convention that $a_{N+1} = a_1$. For a boundary edge $e\in\mathcal{E}_h^B$, let $\mathcal{M}_h^e = \{m_j\}_{j=1}^{k-1}$ denote the canonical interior degrees of freedom on the edge e, and set $\mathcal{M}_h^B = \bigcup_{e\in\mathcal{E}_h^B}\mathcal{M}_h^e$. Let $\boldsymbol{n}_j$ be the normal vector of $\partial\Omega_h$ restricted to the edge e_j, and let $\boldsymbol{t}_j$ be the tangent vector obtained by rotating $\boldsymbol{n}_j$ 90 degrees clockwise. Without loss of generality, we assume that $\boldsymbol{t}_j$ is parallel to $a_{j+1}-a_j$. We further denote by $\mathcal{V}_h^C$ the set of boundary corner vertices, i.e., if $a_j\in\mathcal{V}_h^C$, then the outward unit normals $\boldsymbol{n}_j,\boldsymbol{n}_{j-1}$ of the edges touching a_j are linearly independent. The set of flat boundary vertices are defined as $\mathcal{V}_h^F = \mathcal{V}_h^B\setminus\mathcal{V}_h^C$. Note that $\boldsymbol{n}_j=\boldsymbol{n}_{j-1}$ and $\boldsymbol{t}_j=\boldsymbol{t}_{j-1}$ for $a_j\in\mathcal{V}_h^F$.

We let $h_I\in X_h$ denote the continuous, piecewise linear polynomial with respect to the partition $\mathcal{E}_h^B$ satisfying $h_I(a_j) = (h_{e_{j-1}}+h_{e_j})/2$. Given $\mu\in\mathring{X}_h$, we let $P_h(h_I\mu)\in\mathring{X}_h$ be the L^2-projection of $h_I\mu$, i.e.,

$$\int_{\partial\Omega_h} P_h(h_I\mu)\varkappa\,\mathrm{d}s = \int_{\partial\Omega_h} h_I\mu\varkappa\,\mathrm{d}s \qquad \forall\varkappa\in\mathring{X}_h.$$

We then define $\boldsymbol{v}\in\boldsymbol{V}_h$ by the conditions

$$\begin{aligned}
(\boldsymbol{v}\cdot\boldsymbol{n}_j)(a_j) &= P_h(h_I\mu)(a_j), & (\boldsymbol{v}\cdot\boldsymbol{n}_{j-1})(a_j) &= P_h(h_I\mu)(a_j) & &\forall a_j\in\mathcal{V}_h^C\\
(\boldsymbol{v}\cdot\boldsymbol{n}_j)(a_j) &= P_h(h_I\mu)(a_j), & (\boldsymbol{v}\cdot\boldsymbol{t}_j)(a_j) &= 0 & &\forall a_j\in\mathcal{V}_h^F\\
(\boldsymbol{v}\cdot\boldsymbol{n}_j)(m_j) &= P_h(h_I\mu)(m_j), & (\boldsymbol{v}\cdot\boldsymbol{t}_j)(m_j) &= 0 & &\forall m_j\in\mathcal{M}_h^e,\quad \forall e\in\mathcal{E}_h^B.
\end{aligned} \tag{4.10}$$

All other (Lagrange) degrees of freedom of $\boldsymbol{v}$ are set to zero.

Since $(\boldsymbol{v}\cdot\boldsymbol{n}_j - P_h(h_I\mu))|_{e_j}$ is a polynomial of degree k on each $e_j\in\mathcal{E}_h^B$, and $\boldsymbol{v}\cdot\boldsymbol{n}_j = P_h(h_I\mu)$ at $(k+1)$ distinct points on e_j, we have $\boldsymbol{v}\cdot\boldsymbol{n}_j - P_h(h_I\mu)|_{e_j}=0$. Thus by shape regularity,

$$\int_{\partial\Omega_h}(\boldsymbol{v}\cdot\boldsymbol{n}_h)\mu\,\mathrm{d}s = \int_{\partial\Omega_h}P_h(h_I\mu)\mu\,\mathrm{d}s = \int_{\partial\Omega_h}h_I\mu^2\,\mathrm{d}s \geqslant C\|\mu\|_{-1/2,h}^2. \tag{4.11}$$

It remains to show that $\|\boldsymbol{v}\|_{1,h}\leqslant C\|\mu\|_{-1/2,h}$ to complete the proof.

For $K\in\mathcal{T}_h^{ct}$, let $\mathcal{V}_K^B,\mathcal{V}_K^C,\mathcal{V}_K^F,\mathcal{M}_K^B$ be the sets of elements in $\mathcal{V}_h^B,\mathcal{V}_h^C,\mathcal{V}_h^F,\mathcal{M}_h^B$ contained in $\overline{K}$, respectively. By a standard scaling argument and (4.10), we get for $m=0,1$:

$$\begin{aligned}
\|\boldsymbol{v}\|_{H^m(K)}^2 &\leqslant C\sum_{c_j\in\mathcal{V}_K^B\cup\mathcal{M}_K^B} h_{e_j}^{2-2m}|\boldsymbol{v}(c_j)|^2\\
&= C\Big(\sum\nolimits_{a_j\in\mathcal{V}_K^C} h_{e_j}^{2-2m}|\boldsymbol{v}(a_j)|^2 + \sum\nolimits_{c_j\in\mathcal{V}_K^F\cup\mathcal{M}_K^B} h_{e_j}^{2-2m}|P_h(h_I\mu)(c_j)|^2\Big).
\end{aligned} \tag{4.12}$$

Claim: $|\boldsymbol{v}(a_j)|\leqslant C|P_h(h_I\mu)(a_j)|$ for all $a_j\in\mathcal{V}_K^C$, where $C>0$ is uniformly bounded and independent of h, $\boldsymbol{n}_j$, and $\boldsymbol{n}_{j-1}$.

Proof of the claim. Assume that $\mathcal{V}_K^C$ is non-empty for otherwise the proof is trivial. For $a_j\in\mathcal{V}_K^C$, we write $\boldsymbol{v}(a_j)$ in terms of the basis $\{\boldsymbol{t}_j,\boldsymbol{t}_{j-1}\}$, use (4.10), and apply some elementary vector identities:

$$\begin{aligned}
\boldsymbol{v}(a_j) &= \frac{1}{\boldsymbol{t}_{j-1}\cdot\boldsymbol{n}_j}(\boldsymbol{v}\cdot\boldsymbol{n}_j)(a_j)\boldsymbol{t}_{j-1} + \frac{1}{\boldsymbol{t}_j\cdot\boldsymbol{n}_{j-1}}(\boldsymbol{v}\cdot\boldsymbol{n}_{j-1})(a_j)\boldsymbol{t}_j\\
&= P_h(h_I\mu)(a_j)\Big(\frac{1}{\boldsymbol{t}_{j-1}\cdot\boldsymbol{n}_j}\boldsymbol{t}_{j-1} + \frac{1}{\boldsymbol{t}_j\cdot\boldsymbol{n}_{j-1}}\boldsymbol{t}_j\Big)\\
&= P_h(h_I\mu)(a_j)\Big(\frac{\boldsymbol{t}_j-\boldsymbol{t}_{j-1}}{\boldsymbol{t}_j\cdot\boldsymbol{n}_{j-1}}\Big).
\end{aligned} \tag{4.13}$$

We now show that $|(\boldsymbol{t}_j - \boldsymbol{t}_{j-1})/(\boldsymbol{t}_j \cdot \boldsymbol{n}_{j-1})|$ is bounded. Write $\boldsymbol{t}_j = (\cos(\vartheta_j), \sin(\vartheta_j))^T$ with $\vartheta_{j-1}, \vartheta_j \in [-\pi, \pi]$, so that

$$\frac{\boldsymbol{t}_j - \boldsymbol{t}_{j-1}}{\boldsymbol{t}_j \cdot \boldsymbol{n}_{j-1}} = \frac{(\cos(\vartheta_j) - \cos(\vartheta_{j-1}), \sin(\vartheta_j) - \sin(\vartheta_{j-1}))^T}{\sin(\vartheta_j - \vartheta_{j-1})}.$$

Since

$$\lim_{\vartheta_j \to \vartheta_{j-1}} \frac{(\cos\vartheta_j - \cos\vartheta_{j-1}, \sin\vartheta_j - \sin\vartheta_{j-1})^T}{\sin(\vartheta_j - \vartheta_{j-1})} = \lim_{\vartheta_j \to \vartheta_{j-1}} \frac{(-\sin\vartheta_j, \cos\vartheta_j)^T}{\cos(\vartheta_j - \vartheta_{j-1})} = (-\sin\vartheta_{j-1}, \cos\vartheta_{j-1})^T$$

and due to the shape regularity of the mesh, we conclude $|(\boldsymbol{t}_j - \boldsymbol{t}_{j-1})/(\boldsymbol{t}_j \cdot \boldsymbol{n}_{j-1})|$ is bounded in the case $|\boldsymbol{t}_j \cdot \boldsymbol{n}_{j-1}| \ll 1$, in particular, for 'nearly flat boundary vertices'. Therefore, $|(\boldsymbol{t}_j - \boldsymbol{t}_{j-1})/(\boldsymbol{t}_j \cdot \boldsymbol{n}_{j-1})| \leqslant C$ on shape-regular triangulations for some $C > 0$ independent of h and $\{\boldsymbol{n}_{j-1}, \boldsymbol{n}_j\}$. With (4.13), this yields $|\boldsymbol{v}(a_j)| \leqslant C|P_h(h_I\mu)(a_j)|$ for all $a_j \in \mathcal{V}_K^C$, which concludes the proof of the claim.

Applying the claim to (4.12) and a scaling argument yields

$$\|\boldsymbol{v}\|_{H^m(K)}^2 \leqslant C \sum_{c_j \in \mathcal{V}_K^B \cup \mathcal{M}_K^B} h_{e_j}^{2-2m} |P_h(h_I\mu)(c_j)|^2 \leqslant C \sum_{\substack{e \in \mathcal{E}_h^B \\ a_j \in \bar{e}:\ a_j \in \mathcal{V}_K^B}} h_e^{1-2m} \|P_h(h_I\mu)\|_{L^2(e)}^2.$$

Therefore, by an inverse inequality and shape-regularity of $\mathcal{T}_h^{ct}$,

$$\begin{aligned}
\|\boldsymbol{v}\|_{1,h}^2 &= \|\nabla \boldsymbol{v}\|_{L^2(\Omega_h)}^2 + \sum_{e \in \mathcal{E}_h^B} \frac{1}{h_e} \|\boldsymbol{v}\|_{L^2(e)}^2 \\
&\leqslant \|\nabla \boldsymbol{v}\|_{L^2(\Omega_h)}^2 + C \sum_{K \in \mathcal{T}_h^{ct}} h_K^{-2} \|\boldsymbol{v}\|_{L^2(K)}^2 \leqslant C \sum_{e \in \mathcal{E}_h^B} h_e^{-1} \|P_h(h_I\mu)\|_{L^2(e)}^2.
\end{aligned}$$

Finally, using the L^2-stability of $P_h(h_I\mu)$ and the quasi-uniform assumption, we have

$$\begin{aligned}
\|\boldsymbol{v}\|_{1,h}^2 &\leqslant C \sum_{e \in \mathcal{E}_h^B} h_e^{-1} \|P_h(h_I\mu)\|_{L^2(e)}^2 \\
&\leqslant Ch^{-1} \|P_h(h_I\mu)\|_{L^2(\partial\Omega_h)}^2 \leqslant Ch^{-1} \|h_I\mu\|_{L^2(\partial\Omega_h)}^2 \leqslant C\|\mu\|_{-1/2,h}^2.
\end{aligned} \tag{4.14}$$

Combining this estimate with (4.11) yields the desired inf-sup condition (4.9). □

Remark 4.3. The proof of Lemma 4.7, and in particular the proof of the claim, relies on the continuity properties of the Lagrange multiplier space at nearly flat corner vertices.

4.4 Main stability estimates

Combining Lemmas 4.6 and 4.7 yields inf-sup stability for the bilinear form $b_h(\cdot,\cdot)$. We also show that this result implies inf-sup stability for the bilinear form with boundary correction $b_h^e(\cdot,\cdot)$.

Theorem 4.1. *Assume $\mathcal{T}_h$ is quasi-uniform. Then there exists $\beta > 0$ depending only on β_1 and β_2 such that*

$$\beta\|(q,\mu)\| \leqslant \sup_{\boldsymbol{v} \in \boldsymbol{V}_h \setminus \{0\}} \frac{b_h(\boldsymbol{v},(q,\mu))}{\|\boldsymbol{v}\|_{1,h}} \quad \forall (q,\mu) \in \mathring{Q}_h \times \mathring{X}_h. \tag{4.15}$$

Proof. We use Lemmas 4.6 and 4.7 and follow the arguments in [23, Thm. 3.1].

Fix $(q,\mu) \in \mathring{Q}_h \times \mathring{X}_h$. The statement (4.9) implies the existence of $\boldsymbol{v}_2 \in \boldsymbol{V}_h$ such that $\|\boldsymbol{v}_2\|_{1,h} \leqslant 1$ and

$$\int_{\partial\Omega_h} (\boldsymbol{v}_2 \cdot \boldsymbol{n}_h)\mu \, ds \geqslant \beta_2 \|\mu\|_{-1/2,h}.$$

By Lemma 4.6, there exists $\boldsymbol{v}_1 \in \mathring{\boldsymbol{V}}_h$ satisfying $\|\nabla \boldsymbol{v}_1\|_{L^2(\Omega_h)} = \|\boldsymbol{v}_1\|_{1,h} \leqslant 1$ and

$$-\int_{\Omega_h} (\operatorname{div} \boldsymbol{v}_1) q \geqslant \beta_1 \|q\|_{L^2(\Omega_h)}.$$

Set $\boldsymbol{v} = c\boldsymbol{v}_1 + \boldsymbol{v}_2$ for some $c > 0$, so that $\|\boldsymbol{v}\|_{1,h} \leqslant (1 + c)$, and

$$\begin{aligned}-\int_{\Omega_h} (\operatorname{div} \boldsymbol{v})q \,\mathrm{d}x &\geqslant c\beta_1\|q\|_{L^2(\Omega_h)} - \|\operatorname{div} \boldsymbol{v}_2\|_{L^2(\Omega_h)}\|q\|_{L^2(\Omega_h)}\\ &\geqslant c\beta_1\|q\|_{L^2(\Omega_h)} - \sqrt{2}\|\nabla \boldsymbol{v}_2\|_{L^2(\Omega_h)}\|q\|_{L^2(\Omega_h)}\\ &\geqslant c\beta_1\|q\|_{L^2(\Omega_h)} - \sqrt{2}\|\boldsymbol{v}_2\|_{1,h}\|q\|_{L^2(\Omega_h)}\\ &= \big(c\beta_1 - \sqrt{2}\big)\|q\|_{L^2(\Omega_h)}.\end{aligned}$$

Due to $\boldsymbol{v}_1|_{\partial\Omega_h} = 0$, we have

$$\int_{\partial\Omega_h} (\boldsymbol{v} \cdot \boldsymbol{n}_h)\mu \,\mathrm{d}s = \int_{\partial\Omega_h} (\boldsymbol{v}_2 \cdot \boldsymbol{n}_h)\mu \,\mathrm{d}s \geqslant \beta_2\|\mu\|_{-1/2,h}.$$

Therefore,

$$\begin{aligned}b_h(\boldsymbol{v}, (q, \mu)) &\geqslant \big(c\beta_1 - \sqrt{2}\big)\|q\|_{L^2(\Omega_h)} + \beta_2\|\mu\|_{-1/2,h}\\ &\geqslant (1 + c)^{-1}\Big(\big(c\beta_1 - \sqrt{2}\big)\|q\|_{L^2(\Omega_h)} + \beta_2\|\mu\|_{-1/2,h}\Big)\|\boldsymbol{v}\|_{1,h}.\end{aligned}$$

We now choose $c > 0$ sufficiently large to obtain the desired result. □

Corollary 4.1. Provided assumption (A) is satisfied and the mesh $\mathcal{T}_h$ is quasi-uniform, there exists $\beta_e > 0$ independent of h such that there holds

$$\beta_e\|(q, \mu)\| \leqslant \sup_{\boldsymbol{v}\in \boldsymbol{V}_h\setminus\{0\}} \frac{b_h^e(\boldsymbol{v}, (q, \mu))}{\|\boldsymbol{v}\|_{1,h}} \qquad \forall (q, \mu) \in \mathring{Q}_h \times \mathring{X}_h. \tag{4.16}$$

Proof. Combining Theorem 4.1 and Lemma 4.2, we have

$$\beta\|(q, \mu)\| \leqslant \sup_{\boldsymbol{v}\in \boldsymbol{V}_h\setminus\{0\}} \frac{b_h^e(\boldsymbol{v}, (q, \mu))}{\|\boldsymbol{v}\|_{1,h}} + Cc_\delta\|(q, \mu)\| \qquad \forall (q, \mu) \in \mathring{Q}_h \times \mathring{X}_h.$$

This result implies (4.16) for c_δ sufficiently small with $\beta_e = \beta - Cc_\delta$. □

Theorem 4.2. *Let $(\boldsymbol{u}_h, p_h, \lambda_h) \in \boldsymbol{V}_h \times \mathring{Q}_h \times \mathring{X}_h$ satisfy* (3.2)*. Then, provided c_δ in assumption* (A) *is sufficiently small and the mesh $\mathcal{T}_h$ is quasi-uniform, there holds*

$$\nu\|\boldsymbol{u}_h\|_{1,h} + \|(p_h, \lambda_h)\| \leqslant C\|\boldsymbol{f}\|_{-1,h} \tag{4.17}$$

where $\|\boldsymbol{f}\|_{-1,h} = \sup_{\boldsymbol{v}\in \boldsymbol{V}_h\setminus\{0\}} \int_{\Omega_h} \boldsymbol{f} \cdot \boldsymbol{v} \,\mathrm{d}x/\|\boldsymbol{v}\|_{1,h}$. Consequently, there exists a unique solution to (3.2)*.*

Proof. Setting $\boldsymbol{v} = \boldsymbol{u}_h$ in (3.2a), $(q, \mu) = (p_h, \lambda_h)$ in (3.2b), and subtracting the resulting expressions yields

$$a_h(\boldsymbol{u}_h, \boldsymbol{u}_h) = \int_{\Omega_h} \boldsymbol{f} \cdot \boldsymbol{u}_h \,\mathrm{d}x + \int_{\partial\Omega_h} \big((S_h\boldsymbol{u}_h - \boldsymbol{u}_h) \cdot \boldsymbol{n}_h\big)\lambda_h \,\mathrm{d}s.$$

We apply the coercivity result in Lemma 4.3, the Cauchy–Schwarz inequality, and (4.1) to get

$$\nu c_1\|\boldsymbol{u}_h\|_{1,h}^2 \leqslant \|\boldsymbol{f}\|_{-1,h}\|\boldsymbol{u}_h\|_{1,h} + Cc_\delta\|\boldsymbol{u}_h\|_{1,h}\|\lambda_h\|_{-1/2,h}. \tag{4.18}$$

On the other hand, we use inf-sup stability (4.15) to conclude

$$\begin{aligned}\beta\|(p_h, \lambda_h)\|_{-1/2,h} &\leqslant \sup_{\boldsymbol{v}\in \boldsymbol{V}_h\setminus\{0\}} \frac{b_h(\boldsymbol{v}, (p_h, \lambda_h))}{\|\boldsymbol{v}\|_{1,h}}\\ &\leqslant \sup_{\boldsymbol{v}\in \boldsymbol{V}_h\setminus\{0\}} \frac{\int_{\Omega_h} \boldsymbol{f} \cdot \boldsymbol{v} \,\mathrm{d}x - a_h(\boldsymbol{u}_h, \boldsymbol{v})}{\|\boldsymbol{v}\|_{1,h}}.\end{aligned}$$

Using the continuity estimate (4.3) gets

$$\beta\|\lambda_h\|_{-1/2,h} \leqslant \beta\|(p_h, \lambda_h)\| \leqslant \|\boldsymbol{f}\|_{-1,h} + C(1 + \sigma)\nu\|\boldsymbol{u}_h\|_{1,h}. \tag{4.19}$$

Inserting this estimate into (4.18), we obtain

$$\nu\big(c_1 - Cc_\delta\beta^{-1}(1 + \sigma)\big)\|\boldsymbol{u}_h\|_{1,h} \leqslant (1 + Cc_\delta\beta^{-1})\|\boldsymbol{f}\|_{-1,h}.$$

Thus, $\|\boldsymbol{u}_h\|_{1,h} \leqslant C\nu^{-1}\|\boldsymbol{f}\|_{-1,h}$ for c_δ sufficiently small. This, combined with (4.19), yields the desired stability result (4.17). □

5 Convergence analysis

In this section, we show that the solution to the finite element method (3.2) converges with optimal order provided the exact solution is sufficiently smooth. Throughout this section, we assume that the hypotheses of Theorem 4.2 are satisfied, i.e., assumption A is satisfied and the mesh $\mathcal{T}_h$ is quasi-uniform.

5.1 Consistency estimates

Notice that since we assume homogeneous Dirichlet boundary condition on $\partial\Omega$, there holds $S_h\boldsymbol{u} + R_h\boldsymbol{u} = 0$, where $R_h\boldsymbol{u}$ is the Taylor remainder. The following lemma bounds the boundary correction operator acting on the exact velocity function. The result is essentially an estimate on $R_h\boldsymbol{u}$ and follows from similar arguments in [4, Prop. 3] (see also [7]). For this reason, we just give a sketch of the proof in Appendix B.

Lemma 5.1. *For any $\boldsymbol{u} \in \boldsymbol{H}^{k+1}(\Omega) \cap \boldsymbol{H}_0^1(\Omega)$, there holds*

$$\sum_{e\in\mathcal{E}_h^B} h_e^{-1} \int_e |S_h\boldsymbol{u}|^2\,\mathrm{d}s \leqslant Ch^{2k}\|\boldsymbol{u}\|^2_{H^{k+1}(\Omega)}.$$

Lemma 5.2. *There holds for all $\boldsymbol{u} \in \boldsymbol{H}^{k+1}(\Omega) \cap \boldsymbol{H}_0^1(\Omega)$,*

$$\Big| -\nu \int_{\Omega_h} \Delta\boldsymbol{u}\cdot\boldsymbol{v}\,\mathrm{d}x - a_h(\boldsymbol{u},\boldsymbol{v})\Big| \leqslant C\nu h^k\|\boldsymbol{u}\|_{H^{k+1}(\Omega)}\|\boldsymbol{v}\|_{1,h} \qquad \forall \boldsymbol{v}\in\boldsymbol{V}_h. \tag{5.1}$$

If $\operatorname{div}\boldsymbol{u} = 0$ *in Ω, then*

$$|b_h^e(\boldsymbol{u},(q,\mu))| \leqslant Ch^k\|\boldsymbol{u}\|_{H^{k+1}(\Omega)}\|(q,\mu)\| \qquad \forall (q,\mu)\in\mathring{Q}_h\times\mathring{X}_h.$$

Proof. We integrate-by-parts to write

$$\Big| -\nu \int_{\Omega_h} \Delta\boldsymbol{u}\cdot\boldsymbol{v}\,\mathrm{d}x - a_h(\boldsymbol{u},\boldsymbol{v})\Big| = \nu\Big|\sum_{e\in\mathcal{E}_h^B}\int_e \frac{\partial\boldsymbol{v}}{\partial\boldsymbol{n}_h}\cdot(S_h\boldsymbol{u})\,\mathrm{d}s + \sum_{e\in\mathcal{E}_h^B}\frac{\sigma}{h_e}\int_e (S_h\boldsymbol{u})\cdot(S_h\boldsymbol{v})\,\mathrm{d}s\Big|.$$

Next, we estimate the two terms on the right hand side of the above equality by using the Cauchy–Schwarz inequality, trace and inverse inequalities, along with Lemmas 4.1 and 5.1 as follows:

$$\begin{aligned}\Big|\sum_{e\in\mathcal{E}_h^B}\int_e \frac{\partial\boldsymbol{v}}{\partial\boldsymbol{n}_h}\cdot(S_h\boldsymbol{u})\,\mathrm{d}s\Big| &\leqslant \Big(\sum_{e\in\mathcal{E}_h^B} h_e\int_e\Big|\frac{\partial\boldsymbol{v}}{\partial\boldsymbol{n}_h}\Big|^2\,\mathrm{d}s\Big)^{1/2}\Big(\sum_{e\in\mathcal{E}_h^B} h_e^{-1}\int_e|S_h\boldsymbol{u}|^2\,\mathrm{d}s\Big)^{1/2}\\ &\leqslant Ch^k\|\boldsymbol{u}\|_{H^{k+1}(\Omega)}\|\boldsymbol{v}\|_{1,h}\end{aligned}$$

and

$$\begin{aligned}\Big|\sum_{e\in\mathcal{E}_h^B}\frac{\sigma}{h_e}\int_e(S_h\boldsymbol{u})\cdot(S_h\boldsymbol{v})\,\mathrm{d}s\Big| &\leqslant \sigma\Big(\sum_{e\in\mathcal{E}_h^B} h_e^{-1}\int_e|S_h\boldsymbol{u}|^2\,\mathrm{d}s\Big)^{1/2}\Big(\sum_{e\in\mathcal{E}_h^B} h_e^{-1}\int_e|S_h\boldsymbol{v}|^2\,\mathrm{d}s\Big)^{1/2}\\ &\leqslant Ch^k\|\boldsymbol{u}\|_{H^{k+1}(\Omega)}\|\boldsymbol{v}\|_{1,h}.\end{aligned}$$

Thus, the first estimate (5.1) holds.

Similarly, another use of the Cauchy–Schwarz inequality with Lemma 5.1 yields

$$|b_h^e(\boldsymbol{u},(q,\mu))| = \Big|\sum_{e\in\mathcal{E}_h^B}\int_e(S_h\boldsymbol{u}\cdot\boldsymbol{n}_h)\mu\,\mathrm{d}s\Big| \leqslant Ch^k\|\boldsymbol{u}\|_{H^{k+1}(\Omega)}\|\mu\|_{-1/2,h}$$

and this completes the proof. □

5.2 Approximation properties of the kernel

We define the discrete kernel as

$$\boldsymbol{Z}_h = \{\boldsymbol{v} \in \boldsymbol{V}_h : \ b_h^e(\boldsymbol{v}, (q, \mu)) = 0 \quad \forall (q, \mu) \in \mathring{Q}_h \times \mathring{X}_h\}.$$

Note that if $\boldsymbol{v} \in \boldsymbol{Z}_h$, then $\operatorname{div} \boldsymbol{v} = 0$ in Ω_h (cf. Lemma 3.1), and

$$\int_{\partial\Omega_h} ((S_h \boldsymbol{v}) \cdot \boldsymbol{n}_h)\mu \, ds = 0 \qquad \forall \mu \in \mathring{X}_h. \tag{5.2}$$

In this section, we show that the kernel $\boldsymbol{Z}_h$ has optimal order approximation properties with respect to divergence-free smooth functions. To this end, we define the orthogonal complement of $\boldsymbol{Z}_h$ as

$$\boldsymbol{Z}_h^\perp := \{\boldsymbol{v} \in \boldsymbol{V}_h : \ (\boldsymbol{v}, \boldsymbol{w})_{1,h} = 0 \quad \forall \boldsymbol{w} \in \boldsymbol{Z}_h\}$$

where $(\cdot, \cdot)_{1,h}$ is the inner product on $\boldsymbol{V}_h$ that induces the norm $\|\cdot\|_{1,h}$.

Lemma 5.3. *There holds*

$$\beta_e \|\boldsymbol{w}\|_{1,h} \leqslant \sup_{(q,\mu) \in \mathring{Q}_h \times \mathring{X}_h \setminus \{0\}} \frac{b_h^e(\boldsymbol{w}, (q, \mu))}{\|(q, \mu)\|} \quad \forall \boldsymbol{w} \in \boldsymbol{Z}_h^\perp.$$

Proof. The result immediately follows from Corollary 4.1 and standard results in mixed finite element theory (cf. [9, Lem. 12.5.10]). □

The following theorem states the approximation properties of the discrete kernel.

Theorem 5.1. *For any $\boldsymbol{u} \in \boldsymbol{H}^{k+1}(\Omega) \cap \boldsymbol{H}_0^1(\Omega)$ with* $\operatorname{div} \boldsymbol{u} = 0$*, there holds*

$$\inf_{\boldsymbol{w} \in \boldsymbol{Z}_h} |||\boldsymbol{u} - \boldsymbol{w}|||_h \leqslant Ch^k \|\boldsymbol{u}\|_{H^{k+1}(\Omega)}. \tag{5.3}$$

Proof. Let $\boldsymbol{v} \in \boldsymbol{V}_h$ be arbitrary. By Lemma 5.3, there exists $\boldsymbol{y} \in \boldsymbol{Z}_h^\perp$ such that

$$b_h^e(\boldsymbol{y}, (q, \mu)) = b_h^e(\boldsymbol{u} - \boldsymbol{v}, (q, \mu)) \qquad \forall (q, \mu) \in \mathring{Q}_h \times \mathring{X}_h$$

and $\|\boldsymbol{y}\|_{1,h} \leqslant C\beta_e^{-1} \|\boldsymbol{u} - \boldsymbol{v}\|_{1,h}$, where $C > 0$ is the continuity constant of the bilinear form b_h^e (cf. (4.6)). We then let $\boldsymbol{z} \in \boldsymbol{Z}_h^\perp$ satisfy

$$b_h^e(\boldsymbol{z}, (q, \mu)) = -b_h^e(\boldsymbol{u}, (q, \mu)) \qquad \forall (q, \mu) \in \mathring{Q}_h \times \mathring{X}_h.$$

Then $\boldsymbol{w} := \boldsymbol{v} + \boldsymbol{y} + \boldsymbol{z} \in \boldsymbol{Z}_h$, and

$$\begin{aligned} \|\boldsymbol{u} - \boldsymbol{w}\|_{1,h} &\leqslant \|\boldsymbol{u} - \boldsymbol{v}\|_{1,h} + \|\boldsymbol{y}\|_{1,h} + \|\boldsymbol{z}\|_{1,h} \\ &\leqslant (1 + C\beta_e^{-1}) \|\boldsymbol{u} - \boldsymbol{v}\|_{1,h} + \|\boldsymbol{z}\|_{1,h}. \end{aligned}$$

By Lemmas 5.3 and 5.2,

$$\beta_e \|\boldsymbol{z}\|_{1,h} \leqslant \sup_{(q,\mu) \in \mathring{Q}_h \times \mathring{X}_h \setminus \{0\}} \frac{b_h^e(\boldsymbol{u}, (q, \mu))}{\|(q, \mu)\|} \leqslant Ch^k \|\boldsymbol{u}\|_{H^{k+1}(\Omega)}$$

and so, by Lemma 4.1,

$$\begin{aligned} |||\boldsymbol{u} - \boldsymbol{w}|||_h &\leqslant |||\boldsymbol{u} - \boldsymbol{v}|||_h + C\|\boldsymbol{v} - \boldsymbol{w}\|_{1,h} \leqslant C\big(|||\boldsymbol{u} - \boldsymbol{v}|||_h + \|\boldsymbol{u} - \boldsymbol{w}\|_{1,h}\big) \\ &\leqslant C(1 + \beta_e^{-1})\big(|||\boldsymbol{u} - \boldsymbol{v}|||_h + \|\boldsymbol{u} - \boldsymbol{v}\|_{1,h} + h^k \|\boldsymbol{u}\|_{H^{k+1}(\Omega)}\big) \quad \forall \boldsymbol{v} \in \boldsymbol{V}_h. \end{aligned}$$

Taking $\boldsymbol{v}$ to be the nodal interpolant of $\boldsymbol{u}$, we obtain the desired result. □

Theorem 5.2. *Suppose that the solution to* (2.1) *has regularity* $(\boldsymbol{u}, p) \in \boldsymbol{H}^{k+1}(\Omega) \cap \boldsymbol{H}_0^1(\Omega) \times H^1(\Omega)$. *Furthermore, without loss of generality, assume that* $p|_{\Omega_h} \in L_0^2(\Omega_h)$. *Then,*

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{1,h} \leqslant C\big(h^k \|\boldsymbol{u}\|_{H^{k+1}(\Omega)} + \nu^{-1} \inf_{\mu \in X_h} \|p - \mu\|_{-1/2,h}\big) \tag{5.4a}$$

$$\|p - p_h\|_{L^2(\Omega_h)} \leqslant C(\nu h^k \|\boldsymbol{u}\|_{H^{k+1}(\Omega)} + \inf_{\mu \in X_h} \|p - \mu\|_{-1/2,h} + \inf_{q_h \in \mathring{Q}_h} \|p - q_h\|_{L^2(\Omega)}) \tag{5.4b}$$

$$\|\mathring{p} - \lambda_h\|_{-1/2,h} \leqslant C\Big(\nu h^k \|\boldsymbol{u}\|_{H^{k+1}(\Omega)} + \inf_{\mu \in X_h} \|p - \mu\|_{-1/2,h}\Big) \tag{5.4c}$$

where $\mathring{p} := p - \frac{1}{|\partial\Omega_h|}\int_{\partial\Omega_h} p \,\mathrm{d}s$. *In particular, if* $p \in H^{k+1}(\Omega)$ *there holds*

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{1,h} \leqslant C\big(h^k \|\boldsymbol{u}\|_{H^{k+1}(\Omega)} + \nu^{-1} h^{k+1} \|p\|_{H^{k+1}(\Omega)}\big) \tag{5.5a}$$

$$\|p - p_h\|_{L^2(\Omega_h)} \leqslant C\big(\nu h^k \|\boldsymbol{u}\|_{H^{k+1}(\Omega)} + h^k \|p\|_{H^k(\Omega)}\big) \tag{5.5b}$$

$$\|\mathring{p} - \lambda_h\|_{-1/2,h} \leqslant C\big(\nu h^k \|\boldsymbol{u}\|_{H^{k+1}(\Omega)} + h^{k+1} \|p\|_{H^{k+1}(\Omega)}\big). \tag{5.5c}$$

Remark 5.1. We again emphasis that the inclusion of the Lagrange multiplier in the method yields an additional power of h in the velocity error, which compensates its dependence on the inverse of the viscosity.

Proof. Let $\boldsymbol{w} \in \boldsymbol{Z}_h$ be arbitrary. We then have, for all $\boldsymbol{v} \in \boldsymbol{Z}_h$ and $\mu \in X_h$,

$$\begin{aligned} a_h(\boldsymbol{u}_h - \boldsymbol{w}, \boldsymbol{v}) &= \int_{\Omega_h} \boldsymbol{f} \cdot \boldsymbol{v} - a_h(\boldsymbol{w}, \boldsymbol{v}) - b_h(\boldsymbol{v}, (p_h, \lambda_h)) \\ &= -\nu \int_{\Omega_h} \Delta \boldsymbol{u} \cdot \boldsymbol{v} \,\mathrm{d}x - a_h(\boldsymbol{w}, \boldsymbol{v}) - \int_{\partial\Omega_h} (\boldsymbol{v} \cdot \boldsymbol{n}_h)(\lambda_h - p) \,\mathrm{d}s \\ &= -\nu \int_{\Omega_h} \Delta \boldsymbol{u} \cdot \boldsymbol{v} \,\mathrm{d}x - a_h(\boldsymbol{w}, \boldsymbol{v}) - \int_{\partial\Omega_h} (\boldsymbol{v} \cdot \boldsymbol{n}_h)(\mu - p) \,\mathrm{d}s - \int_{\partial\Omega_h} (\boldsymbol{v} \cdot \boldsymbol{n}_h)(\lambda_h - \mathring{\mu}) \,\mathrm{d}s \end{aligned}$$

where $\mathring{\mu} = \mu - \frac{1}{|\partial\Omega_h|}\int_{\partial\Omega_h} \mu \,\mathrm{d}s \in \mathring{X}_h$.

Therefore by Lemma 5.2, the continuity of $a_h(\cdot,\cdot)$ (cf. (4.3)), and the Cauchy–Schwarz inequality,

$$\begin{aligned} a_h(\boldsymbol{u}_h - \boldsymbol{w}, \boldsymbol{v}) &\leqslant C\big(\nu h^k \|\boldsymbol{u}\|_{H^{k+1}(\Omega)} + \|p - \mu\|_{-1/2,h}\big)\|\boldsymbol{v}\|_{1,h} + a_h(\boldsymbol{u} - \boldsymbol{w}, \boldsymbol{v}) - \int_{\partial\Omega_h} (\boldsymbol{v} \cdot \boldsymbol{n}_h)(\lambda_h - \mathring{\mu}) \,\mathrm{d}s \\ &\leqslant C\big(\nu h^k \|\boldsymbol{u}\|_{H^{k+1}(\Omega)} + \nu(1+\sigma)|||\boldsymbol{u} - \boldsymbol{w}|||_h + \|p - \mu\|_{-1/2,h}\big)\|\boldsymbol{v}\|_{1,h} - \int_{\partial\Omega_h} (\boldsymbol{v} \cdot \boldsymbol{n}_h)(\lambda_h - \mathring{\mu}) \,\mathrm{d}s. \end{aligned}$$

We then use (5.2) and (4.1) to obtain

$$\int_{\partial\Omega_h} (\boldsymbol{v} \cdot \boldsymbol{n}_h)(\lambda_h - \mathring{\mu}) \,\mathrm{d}s = \int_{\partial\Omega_h} \big((\boldsymbol{v} - S_h \boldsymbol{v}) \cdot \boldsymbol{n}_h\big)(\lambda_h - \mathring{\mu}) \,\mathrm{d}s \leqslant C c_\delta \|\boldsymbol{v}\|_{1,h} \|\lambda_h - \mathring{\mu}\|_{-1/2,h}.$$

Setting $\boldsymbol{v} = \boldsymbol{u}_h - \boldsymbol{w}$, applying the coercivity of $a_h(\cdot,\cdot)$ and Theorem 5.1, we obtain

$$c_1 \nu \|\boldsymbol{u}_h - \boldsymbol{w}\|_{1,h} \leqslant C\big(\nu(1+\sigma) h^k \|\boldsymbol{u}\|_{H^{k+1}(\Omega)} + \|p - \mu\|_{-1/2,h} + c_\delta \|\lambda_h - \mathring{\mu}\|_{-1/2,h}\big) \tag{5.6}$$

for $\boldsymbol{w} \in \boldsymbol{Z}_h$ satisfying (5.3).

Next, let $P_h \in \mathring{Q}_h$ be the L^2-projection of p and note that, due to the definitions of the finite element spaces, $\int_{\Omega_h} (\operatorname{div} \boldsymbol{v})(p - P_h) \,\mathrm{d}x = 0$ for all $\boldsymbol{v} \in \boldsymbol{V}_h$. This identity, along with the inf-sup stability estimate given in Theorem 4.1 yields

$$\beta \|(p_h - P_h, \lambda_h - \mathring{\mu})\| \leqslant \sup_{\boldsymbol{v} \in \boldsymbol{V}_h \setminus \{0\}} \frac{b_h(\boldsymbol{v}, (p_h - P_h, \lambda_h - \mathring{\mu}))}{\|\boldsymbol{v}\|_{1,h}} = \sup_{\boldsymbol{v} \in \boldsymbol{V}_h \setminus \{0\}} \frac{b_h(\boldsymbol{v}, (p_h - p, \lambda_h - \mathring{\mu}))}{\|\boldsymbol{v}\|_{1,h}}.$$

Using Lemma 5.2, we write the numerator as

$$\begin{aligned} b_h(\boldsymbol{v}, (p_h - p, \lambda_h - \mathring{\mu})) &= b_h(\boldsymbol{v}, (p_h, \lambda_h)) - b_h(\boldsymbol{v}, (p, \mathring{\mu})) \\ &= \int_{\Omega_h} \boldsymbol{f} \cdot \boldsymbol{v} \,\mathrm{d}x - a_h(\boldsymbol{u}_h, \boldsymbol{v}) + \int_{\Omega_h} (\operatorname{div} \boldsymbol{v}) p \,\mathrm{d}x - \int_{\partial\Omega_h} (\boldsymbol{v} \cdot \boldsymbol{n}_h) \mu \,\mathrm{d}s \\ &\leqslant C\nu h^k \|\boldsymbol{u}\|_{H^{k+1}(\Omega)} \|\boldsymbol{v}\|_{1,h} + a_h(\boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{v}) - \int_{\partial\Omega_h} (\boldsymbol{v} \cdot \boldsymbol{n}_h)(\mu - p) \,\mathrm{d}s. \end{aligned}$$

By continuity and the Cauchy–Schwarz inequality,

$$\begin{aligned}\beta\|(p_h - P_h, \lambda_h - \mathring{\mu})\| &\leqslant C\big(\nu h^k \|\boldsymbol{u}\|_{H^{k+1}(\Omega)} + c_2\nu(1+\sigma)\|\!|\boldsymbol{u} - \boldsymbol{u}_h|\!\|_h + \|p - \mu\|_{-1/2,h}\big)\\ &\leqslant C\big(\nu h^k \|\boldsymbol{u}\|_{H^{k+1}(\Omega)} + c_2\nu(1+\sigma)\big(\|\!|\boldsymbol{u} - \boldsymbol{w}|\!\|_h + \|\boldsymbol{u}_h - \boldsymbol{w}\|_{1,h}\big) + \|p - \mu\|_{-1/2,h}\big)\\ &\leqslant C\big(\nu(1+\sigma)h^k \|\boldsymbol{u}\|_{H^{k+1}(\Omega)} + c_2\nu(1+\sigma)\|\boldsymbol{u}_h - \boldsymbol{w}\|_{1,h} + \|p - \mu\|_{-1/2,h}\big).\end{aligned} \tag{5.7}$$

Inserting this estimate into (5.6), we get

$$\nu\big(c_1 - C\beta^{-1}c_2(1+\sigma)c_\delta\big)\|\boldsymbol{u}_h - \boldsymbol{w}\|_{1,h} \leqslant C\nu(1+\sigma)h^k\|\boldsymbol{u}\|_{H^{k+1}(\Omega)} + C\|p - \mu\|_{-1/2,h}. \tag{5.8}$$

Using the approximation properties of the discrete kernel once again (cf. Theorem 5.1), and for c_δ sufficiently small,

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{1,h} \leqslant C\big(h^k\|\boldsymbol{u}\|_{H^{k+1}(\Omega)} + \nu^{-1}\inf_{\mu\in X_h}\|p - \mu\|_{-1/2,h}\big).$$

This establishes the velocity estimate (5.4a).

To obtain the estimate for the pressure approximation (5.4b), we use the triangle inequality and the approximation properties of the L^2-projection:

$$\|p - p_h\|_{L^2(\Omega_h)} \leqslant \|p_h - P_h\|_{L^2(\Omega_h)} + \inf_{q_h\in\mathring{Q}_h}\|p - q_h\|_{L^2(\Omega_h)}.$$

Inserting (5.7) and (5.8) into the right-hand side yields the desired bound for the pressure. Likewise, combining (5.7) and (5.8) yields

$$\|\mathring{p} - \lambda_h\|_{-1/2,h} \leqslant C\Big(\nu h^k\|\boldsymbol{u}\|_{H^{k+1}(\Omega)} + \inf_{\mu\in X_h}\big(\|p - \mu\|_{-1/2,h} + \|\mathring{p} - \mathring{\mu}\|_{-1/2,h}\big)\Big).$$

Applications of the Cauchy–Schwarz inequality show $\|\mathring{p} - \mathring{\mu}\|_{-1/2,h} \leqslant C\|p - \mu\|_{-1/2,h}$ on quasi-uniform meshes, and therefore (5.4c) holds.

Next, we estimate the term $\inf_{\mu\in X_h}\|p - \mu\|_{-1/2,h}$ for $p \in H^{k+1}(\Omega)$. With an abuse of notation, let μ_I denote the kth degree nodal Lagrange interpolant of p on Ω_h with respect to $\mathcal{T}_h^{ct}$. Notice that $\mu_I|_{\partial\Omega_h} \in X_h$. Applying a trace inequality, followed by standard interpolation estimates and shape regularity of $\mathcal{T}_h^{ct}$, we obtain for each $e \in \mathcal{E}_h^B$,

$$\|p - \mu_I\|_{L^2(e)}^2 \leqslant C\big(h_e^{-1}\|p - \mu_I\|_{L^2(T_e)}^2 + h_e\|\nabla(p - \mu_I)\|_{L^2(T_e)}^2\big) \leqslant Ch_e^{2k+1}\|p\|_{H^{k+1}(T_e)}^2$$

where $T_e \in \mathcal{T}_h^{ct}$ satisfies $e \subset \partial T_e$. We thus conclude from the definition of $\|\cdot\|_{-1/2,h}$ that

$$\inf_{\mu\in X_h}\|p - \mu\|_{-1/2,h} \leqslant Ch^{k+1}\|p\|_{H^{k+1}(\Omega)}. \tag{5.9}$$

Finally, the estimates (5.5a)–(5.5c) follow from (5.4a)–(5.4c), interpolation estimates, and (5.9). □

6 Numerical experiments

In this section we perform simple numerical experiments of the finite element method (3.2) which verify the theoretical rates of convergence established in the previous sections.

In the series of tests, the domain is defined via a level set function [26]:

$$\Omega = \{x \in \mathbb{R}^2 : \varphi(x) < 0\}, \qquad \varphi = r - 0.3723423423343 - 0.1\sin(6\vartheta) \tag{6.1}$$

with $r = \sqrt{(x_1 - 0.5)^2 + (x_2 - 0.5)^2}$, and $\vartheta = \tan^{-1}((x_2 - 0.5)/(x_1 - 0.5))$. We take $k = 2$, $S = (0,1)^2$, and the background mesh $\mathcal{S}_h$ to be a sequence of type I triangulations of S, i.e., a mesh obtained by drawing diagonals of a Cartesian mesh (cf. Fig. 1). For all tests, the Nitsche penalty parameter in the bilinear form $a_h(\cdot,\cdot)$ takes the value $\sigma = 40$.

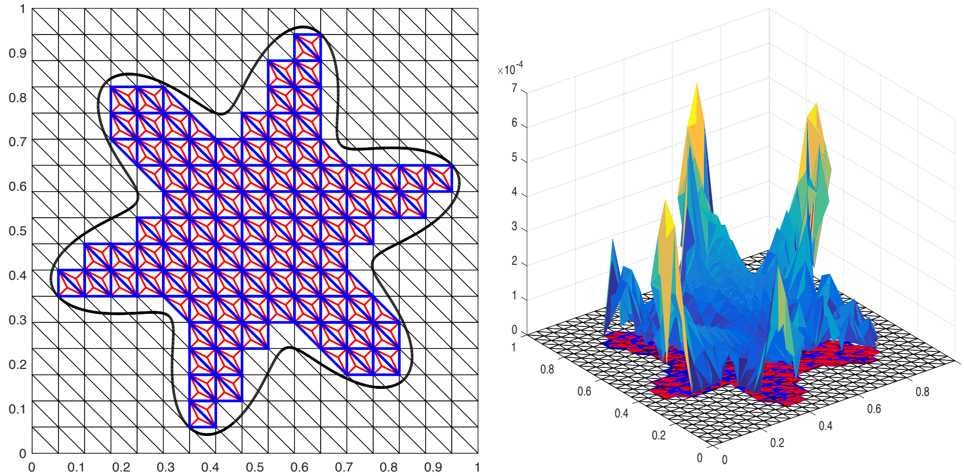

Fig. 1: Left: The domain and mesh with $h = 1/24$. Right: The graph of the error $|\boldsymbol{u} - \boldsymbol{u}_h|$ with exact solution (6.2).

The extension direction $\boldsymbol{d}$ is obtained by solving an auxiliary 2×2 nonlinear system at each quadrature point of each boundary edge of $\mathcal{T}_h^{ct}$. In particular, for each quadrature point $x \in \partial\Omega_h$, we find $x_* \in \partial\Omega$ such that

$$\varphi(x_*) = 0, \qquad (\nabla\varphi(x_*))^{\perp} \cdot (x - x_*) = 0$$

and set $\boldsymbol{d} = (x - x_*)/|x - x_*|$ and $\delta(x) = |x - x_*|$. The first equation ensures that x_* is on the boundary $\partial\Omega$, whereas the second equation states that $\boldsymbol{d}$ is parallel to the outward unit normal of $\partial\Omega$ at x_*.

We choose the data such that the exact solution to the Stokes problem is given by

$$\boldsymbol{u} = \begin{pmatrix} 2(x_1^2 - x_1 + \frac{1}{4} + x_2^2 - x_2)(2x_2 - 1) \\ -2(x_1^2 - x_1 + \frac{1}{4} + x_2^2 - x_2)(2x_1 - 1) \end{pmatrix}, \qquad p = 10(x_1^2 - x_2^2)^2. \tag{6.2}$$

Because the exact solution is smooth, Theorem 5.2 predicts the convergence rates

$$\|\nabla(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2(\Omega_h)} = \mathcal{O}(h^2 + \nu^{-1}h^3), \qquad \|p - p_h\|_{L^2(\Omega_h)} = \mathcal{O}(h^2). \tag{6.3}$$

The velocity and pressure errors are plotted in Fig. 2 for mesh parameters $h = 2^{-j}$ $(j = 3, 4, 5, 6, 7)$ and viscosities $\nu = 10^{-k}$ $(k = 1, 3, 5)$. The results show that, for the moderately sized viscosities $\nu = 10^{-1}$ and $\nu = 10^{-3}$, the L^2 and H^1 velocities converge with the optimal order three and two, respectively. We also observe larger velocity errors for viscosity value $\nu = 10^{-5}$, although, rates of convergence are higher; Figure 2 shows fourth and third order convergence in the L^2 and H^1 norms. This behavior is consistent with the theoretical estimate (6.3). Finally, the numerical experiments show second order convergence for the pressure approximation (with only marginal differences for different viscosity values) and divergence errors comparable to machine epsilon.

7 Concluding remarks

This paper constructed a uniformly stable and divergence-free method for the Stokes problem on unfitted meshes using a boundary correction approach. While the method is not pressure-robust, a Lagrange multiplier enforcing the normal boundary conditions is included to mitigate the affect of the pressure contribution in the velocity error. Theoretical results and numerical experiments show that the method converges with optimal order.

The presentation is confined to the two dimensional setting, however many of the results extend to 3D as well. For example, the proof of inf-sup stability given in Lemma 4.6 applies mutatis mutandis to the three-dimensional Scott–Vogelius pair. On the other hand, inf-sup stability of the velocity-Lagrange multiplier pairing (cf. Lemma 4.7), and its dependence on the geometry of the computational mesh is less obvious. We plan to address this issue in the near future.

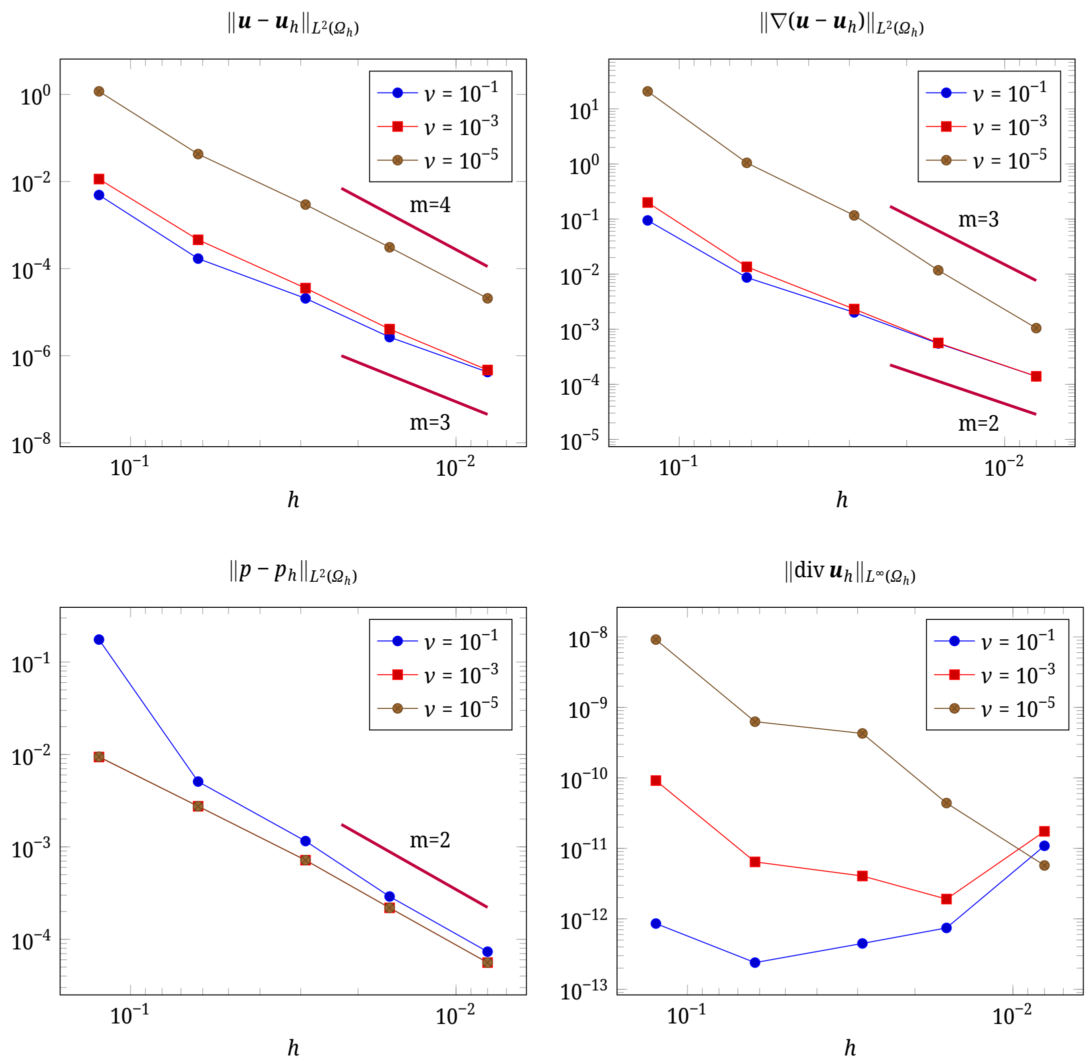

Fig. 2: Errors for the velocity and pressure for a sequence of meshes on domain (6.1) and exact solution (6.2).

Funding: This work was supported in part by the National Science Foundation through grant number DMS-2011733.

References

[1] D. N. Arnold and J. Qin, Quadratic velocity/linear pressure Stokes elements. In: *Advances in Computer Methods for Partial Differential Equations, VII*, (Eds. R. Vichnevetsky, D. Knight, and G. Richter), IMACS, New Brunswick, NJ, 1992, pp. 28–34.

[2] N. M. Atallah, C. Canuto, and G. Scovazzi, The second-generation shifted boundary method and its numerical analysis, *Comput. Methods Appl. Mech. Engrg.* **372** (2020) 113–3471.

[3] N. M. Atallah, C. Canuto, and G. Scovazzi, Analysis of the shifted boundary method for the Stokes problem, *Comput. Methods Appl. Mech. Engrg.* **358** (2020), 112609.

[4] N. M. Atallah, C. Canuto, and G. Scovazzi, Analysis of the shifted boundary method for the Poisson problem in domains with corners, *Math. Comp.* **90** (2021), No. 331, 2041–2069.

[5] M. A. Belenli, L. G. Rebholz, and F. Tone, A note on the importance of mass conservation in long-time stability of Navier–Stokes simulations using finite elements, *Appl. Math. Lett.* **45** (2015), 98–102.

[6] A. Bonito, A. Demlow, and M. Licht, A divergence-conforming finite element method for the surface Stokes equation, *SIAM J. Numer. Anal.* **58** (2020) No. 5, 2764–2798.

[7] J. H. Bramble, T. Dupont, and V. Thomée, Projection methods for Dirichlet's problem in approximating polygonal domains with boundary-value corrections, *Math. Comp.* **26** (1972) 869–879.

[8] J. H. Bramble and T. J. King, A robust finite element method for nonhomogeneous Dirichlet problems in domains with curved boundaries, *Math. Comp.* **63** (1994) No. 207, 1–17.

[9] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, Texts in Applied Mathematics, Vol. 15, Springer, New York, 2008.

[10] A. Buffa, C. de Falco, and G. Sangalli, IsoGeometric analysis: stable elements for the 2D Stokes equation, *Int. J. Numer. Methods Fluids* **65** (2011), No. 11-12, 1407–1422.

[11] E. Burman, P. Hansbo, M. G. Larson, Dirichlet boundary value correction using Lagrange multipliers, *BIT* **60** (2020), No. 1, 235–260.

[12] S. Charnyi, T. Heister, M. A. Olshanskii, and L. G. Rebholz, On conservation laws of Navier–Stokes Galerkin discretizations, *J. Comput. Phys.* **337** (2017), 289–308.

[13] B. Cockburn and M. Solano, Solving Dirichlet boundary-value problems on curved domains by extensions from subdomains, *SIAM J. Sci. Comput.* **34** (2012), No. 1, A497–A519.

[14] J. A. Evans and T. J. R. Hughes, Isogeometric divergence-conforming B-splines for the steady Navier–Stokes equations, *Math. Models Methods Appl. Sci.* **23** (2013), No. 8, 1421–1478.

[15] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*, Reprint of the 1998 edition, Classics in Mathematics, Springer-Verlag, Berlin, 2001.

[16] J. Guzmán and M. Neilan, Inf-sup stable finite elements on barycentric refinements producing divergence-free approximations in arbitrary dimensions, *SIAM J. Numer. Anal.* **56** (2018), No. 5, 2826–2844.

[17] J. Guzmán and M. Neilan, Conforming and divergence-free Stokes elements on general triangular meshes, *Math. Comp.* **83** (2014), No. 285, 15–36.

[18] J. Guzmán and M. Olshanskii, Inf-sup stability of geometrically unfitted Stokes finite elements, *Math. Comp.* **87** (2018), No. 313, 2091–2112.

[19] T. Dupont, J. Guzmán, and R. Scott, Obtaining higher-order Galerkin accuracy when the boundary is polygonally approximated, *arXiv:2001.03082*, 2020.

[20] R. S. Falk and M. Neilan, Stokes complexes and the construction of stable finite elements with pointwise mass conservation, *SIAM J. Numer. Anal.* **51** (2013), No. 2, 1308–1326.

[21] T. Heister, L. G. Rebholz, and M. Xiao, Flux-preserving enforcement of inhomogeneous Dirichlet boundary conditions for strongly divergence-free mixed finite element methods for flow problems, *J. Math. Anal. Appl.* **438** (2016), No. 1, 507–513.

[22] J. S. Hesthaven, T. Warburton On the constants in hp-finite element trace inverse inequalities, *Comput. Methods Appl. Mech. Engrg.* **192** (2003), No. 25, 2765–2773.

[23] J. S. Howell and N. J. Walkington, Inf-sup conditions for twofold saddle point problems, *Numer. Math.* **118** (2011), No. 4, 663–693.

[24] 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 Rev.* **59** (2017), No. 3, 492–544.

[25] P. L. Lederer, C. Lehrenfeld, and J. Schöberl, Divergence-free tangential finite element methods for incompressible flows on surfaces, *Int. J. Numer. Methods Engrg.* **121** (2020), No. 11, 2503–2533.

[26] C. Lehrenfeld, High order unfitted finite element methods on level set domains using isoparametric mappings, *Comput. Methods Appl. Mech. Engrg.* **300** (2016), 716–733.

[27] A. Linke, Collision in a cross-shaped domain – a steady 2D Navier–Stokes example demonstrating the importance of mass conservation in CFD, *Comput. Methods Appl. Mech. Engrg.* **198** (2009), No. 41-44, 3278–3286.

[28] A. Main and G. Scovazzi, The shifted boundary method for embedded domain computations. Part I: Poisson and Stokes problems, *J. Comput. Phys.* **372** (2018), 972–995.

[29] M. Neilan and B. Otus, Divergence-free Scott–Vogelius elements on curved domains, *SIAM J. Numer. Anal.* **59** (2021), No. 2, 1090–1116.

[30] J. Nitsche, Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind, *Abh. Math. Sem. Univ. Hamburg* **36** (1971), 9–15.

[31] R. Oyarzúa, M. Solano, and P. Zúñiga, A high order mixed-FEM for diffusion problems on curved domains, *J. Sci. Comput.* **79** (2019), No. 1, 49–78.

[32] R. Oyarzúa, M. Solano, and P. Zúñiga, A priori and a posteriori error analyses of a high order unfitted mixed-FEM for Stokes flow, *Comput. Methods Appl. Mech. Engrg.* **360** (2020), 112780.

[33] S. Ozisik, B. Riviere, and T. Warburton. On the constants in inverse inequalities in L_2, *Technical Report TR10-19*, Computational and Applied Mathematics Department, Rice University, 2010.

[34] J. Qin, On the convergence of some low order mixed finite elements for incompressible fluids, *Ph.D. Thesis*, The Pennsylvania State University, State College, PA, 1994.

[35] B. Rivière, *Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations*, Theory and Implementation. Frontiers in Applied Mathematics, Vol. 35, SIAM, Philadelphia, PA, 2008.

[36] P. W. Schroeder, C. Lehrenfeld, A. Linke, and G. Lube, Towards computable flows and robust estimates for inf-sup stable FEM applied to the time-dependent incompressible Navier–Stokes equations, *SeMA J.* **75** (2018), No. 4, 629–653.

[37] P. W. Schroeder and G. Lube, Divergence-free $H(div)$-FEM for time-dependent incompressible flows with applications to high Reynolds number vortex dynamics, *J. Sci. Comput.* **75** (2018), No. 2, 830–858.

[38] S. Zhang, A new family of stable mixed finite elements for the 3D Stokes equations, *Math. Comp.* **74** (2005), No. 250, 543–554.

A Details of estimate (4.8)

Here we provide the calculations in the estimate (4.8) that ensures the bilinear form a_h is coercive. As a first step, we provide an explicit estimate of the constant $C > 0$ in the first estimate of (4.1). To this end, we use the discrete trace inequality in [22, Thm. 3] to obtain

$$\begin{aligned} h_e^{-1}\int_e |\delta|^{2j}\left|\left(\frac{\partial^j \boldsymbol{v}}{\partial \boldsymbol{d}^j}\right)^2\right| \,\mathrm{d}s &\leqslant h_e^{-1}\delta_e^{2j}\|D^j\boldsymbol{v}\|_{L^2(e)}^2 \\ &\leqslant h_e^{-2}\delta_e^{2j}\frac{(k+1)(k+2)}{2}\|D^j\boldsymbol{v}\|_{L^2(T_e)}^2, \quad j=1,\dots,k. \end{aligned}$$

Combining this estimate with the inverse estimate in [33, Thm. 2] yields

$$\begin{aligned} h_e^{-1}\int_e |\delta|^{2j}\left|\left(\frac{\partial^j \boldsymbol{v}}{\partial \boldsymbol{d}^j}\right)^2\right| \,\mathrm{d}s &\leqslant \left(\prod\nolimits_{\ell=k-j+1}^{k-1} C_\ell\right)\frac{(k+1)(k+2)}{2}h_e^{-2j}\delta^{2j}\|\nabla\boldsymbol{v}\|_{L^2(T_e)}^2 \\ &\leqslant \left(\prod\nolimits_{\ell=k-j+1}^{k-1} C_\ell\right)\frac{(k+1)(k+2)}{2}c_\delta^{2j}\|\nabla\boldsymbol{v}\|_{L^2(T_e)}^2 \end{aligned}$$

where $C_\ell > 0$ is the maximum eigenvalue of of a matrix defined in [33, Sect. 3], which numerically scales as $O(\ell^4)$:

$$\sum_{e\in\mathcal{E}_h^B} h_e^{-1}\|S_h\boldsymbol{v}-\boldsymbol{v}\|_{L^2(e)}^2 \leqslant \left(\sum\nolimits_{j=1}^{k}\prod\nolimits_{\ell=k-j+1}^{k-1} C_\ell\right)\frac{(k+1)(k+2)}{2}c_\delta^2\|\nabla\boldsymbol{v}\|_{L^2(\Omega_h)}^2.$$

Combining this result with the same discrete trace inequality, we have

$$\left|\sum\nolimits_{e\in\mathcal{E}_h^B}\int_e \frac{\partial\boldsymbol{v}}{\partial\boldsymbol{n}_h}\cdot(S_h\boldsymbol{v}-\boldsymbol{v})\,\mathrm{d}s\right| C_\dagger^{-1}c_\delta\|\nabla\boldsymbol{v}\|_{L^2(\Omega_h)}^2$$

where $C_\dagger$ is given by (4.8). Applying this estimate in (4.7) of the proof of Lemma 4.3 shows that a_h is coercive provided $c_\delta < C_\dagger^{-1}$.

B Proof of Lemma 5.1

For a boundary edge $e \in \mathcal{E}_h^B$ with endpoints a_1, a_2, let $x(t) = a_1 + th_e^{-1}(a_2-a_1)$ $(0 \leqslant t \leqslant h_e)$ be its parameterization, and introduce the 2D parameterization $\varphi(t,s) = x(t) + s\boldsymbol{d}(x(t))$ for $0 \leqslant t \leqslant h_e$ and $0 \leqslant s \leqslant \delta(x(t))$. The Taylor remainder estimation with $S_h\boldsymbol{u} + R_h\boldsymbol{u} = 0$ yields

$$|S_h\boldsymbol{u}(x(t))| = |R_h\boldsymbol{u}(x(t))| = \frac{1}{k!}\left|\int_0^{\delta(x(t))}\frac{\partial^{k+1}\boldsymbol{u}}{\partial\boldsymbol{d}^{k+1}}(\varphi(t,s))(\delta(x(t))-s)^k\,\mathrm{d}s\right|.$$

Applying the Cauchy–Schwarz inequality, we obtain

$$|S_h\boldsymbol{u}(x(t))| \leqslant C\delta(x(t))^{k+1/2}\left(\int_0^{\delta(x(t))}\left|\frac{\partial^{k+1}\boldsymbol{u}}{\partial\boldsymbol{d}^{k+1}}(\varphi(t,s))\right|^2\mathrm{d}s\right)^{1/2}$$

and therefore

$$\begin{aligned} h_e^{-1}\|S_h\boldsymbol{u}\|_{L^2(e)}^2 &\leqslant Ch_e^{-1}\delta_e^{2k+1}\int_0^{h_e}\int_0^{\delta(x(t))}\left|\frac{\partial^{k+1}\boldsymbol{u}}{\partial\boldsymbol{d}^{k+1}}(\varphi(t,s))\right|^2\mathrm{d}s\,\mathrm{d}t \\ &\leqslant Ch_e^{2k}\int_0^{h_e}\int_0^{\delta(x(t))}\left|\frac{\partial^{k+1}\boldsymbol{u}}{\partial\boldsymbol{d}^{k+1}}(\varphi(t,s))\right|^2\mathrm{d}s\,\mathrm{d}t \end{aligned}$$

where we used assumption A in the last inequality. The estimate in Lemma 5.1 now follows from a change of variables (cf. [4, 31]) and summing over $e \in \mathcal{E}_h^B$.