

STAGGERED DG METHODS FOR THE PSEUDOSTRESS-VELOCITY FORMULATION OF THE STOKES EQUATIONS ON GENERAL MESHES*

DOHYUN KIM[†], LINA ZHAO[‡], AND EUN-JAE PARK[†]

Abstract. In this paper, we introduce staggered discontinuous Galerkin methods for the stationary Stokes flow on polygonal meshes. The proposed method is based on the pseudostress-velocity formulation. A Lagrange multiplier on dual edges is introduced to impose the continuity of the pseudostress, which reduces the size of the final system via hybridization and eases the construction of the finite element space for the approximation of the pseudostress. The resulting method is stable and optimally convergent even on distorted or concave polygonal meshes. In addition, hanging nodes can be automatically incorporated in the construction of the method, which favors adaptive mesh refinement. Two types of local postprocessing for the velocity field are proposed to obtain one order higher convergence. Numerical experiments are provided to validate the theoretical findings and demonstrate the performance of the proposed method.

Key words. staggered grid, discontinuous Galerkin method, pseudostress-velocity, Stokes equations, hybridization, general meshes

AMS subject classifications. 65N30, 65N50, 65N12

DOI. 10.1137/20M1322170

1. Introduction. Over the last few years, many advances have been made in the design of efficient numerical methods that can be flexibly applied to fairly general meshes with arbitrary polynomial orders. Such methods can handle meshes with arbitrary shape, and possibly include hanging nodes, which is well suited to adaptive mesh refinement. The use of general meshes enables us to deal with complex geometries, which show great potential in more practical problems. Among all the methods that have been successfully designed on general meshes, we mention in particular a family of conforming finite element methods based on generalized barycentric coordinates [34, 36], hybrid high order methods [23, 22], virtual element methods [4, 5], and weak Galerkin methods [33, 35].

Staggered discontinuous Galerkin (SDG) methods were originally proposed by Chung and Engquist [16] for the mixed formulation of the wave propagation on a triangular mesh. Since then, SDG methods have been successfully applied to various partial differential equations arising from practical applications [18, 15, 32, 31, 19]. The crux of SDG methods is to decompose the computational domain into the primal mesh and the dual mesh, then staggered continuity for all the variables involved is imposed. This staggered continuity makes SDG methods stable without the stabilization term in contrast to other DG methods. Recently, lowest order SDG methods have

*Submitted to the journal's Methods and Algorithms for Scientific Computing section February 27, 2020; accepted for publication (in revised form) May 19, 2020; published electronically August 31, 2020.

<https://doi.org/10.1137/20M1322170>

Funding: The work of the first author was supported by the Graduate School of YONSEI University Research Scholarship grants in 2019. The work of the third author was supported by the National Research Foundation of Korea (NRF) through grants by the Ministry of Science and ICT (NRF-2015R1A5A1009350 and NRF-2019R1A2C2090021).

[†]Department of Computational Science and Engineering, Yonsei University, Seoul 03722, Korea (kim92n@yonsei.ac.kr, ejpark@yonsei.ac.kr).

[‡]Department of Mathematics, The Chinese University of Hong Kong, Hong Kong Special Administrative Region (lzhao@math.cuhk.edu.hk).

been successfully designed on general meshes to solve the Poisson model problem and the Stokes equations [37, 40] and further studied for the coupled Stokes and Darcy problem in [38]. Then, a high order SDG method for general second order elliptic problems is developed in [39]. The SDG method designed therein has many salient features, which can be summarized as follows: It is locally conservative over each dual element; superconvergence can be obtained; it works on fairly general meshes; it handles hanging nodes naturally; it is stable without numerical flux or penalty term.

The stress-velocity-pressure formulation [26] is the original physical equations for incompressible Newtonian flows induced by the conservation of momentum and the constitutive law. The construction of symmetric stress is tricky and it usually requires cumbersome spaces with relatively large degrees of freedom (cf. [41, 30, 2]). To overcome this issue, the pseudostress formulation is initially proposed for the elasticity problem in [3]. This technique, which has become very popular, in particular in fluid mechanics, has gained considerable attention in recent years due to its easy applications to various problems; see, for example, [6, 7, 8, 11, 25, 10, 12, 13]. Moreover, the pressure appearing in the Stokes equations can be obtained by taking the trace of pseudostress tensor while preserving optimal convergence order.

In this study, we introduce an arbitrarily high order SDG method on polygonal meshes to solve the pseudostress-velocity formulation of the Stokes problem and derive its equivalent hybridized formulation. The continuity of the pseudostress and velocity approximations is staggered on the interelement boundaries, which naturally gives an interelement flux term free of stabilization parameters. However, due to the subdivision of the primal mesh, the resulting system has relatively large degrees of freedom, and the construction of the basis functions for the tensor function is cumbersome. To overcome this issue, we introduce a hybrid variable for the pseudostress over the dual edges, which enforces the normal continuity of the pseudostress. With implicit normal continuity of pseudostress, we can use a standard nodal basis function within each subtriangle. In addition, we use the idea of HDG methods (cf. [1, 20, 27, 28, 29]) to define the local problem over each dual element, and the final global system is only coupled via the hybrid variable. Note that the local problem can be solved in parallel, and thus very efficient computation can be done compared to [32] and the resulting system is smaller than that of [24]. Finally, we derive two types of local postprocessing for the velocity field, and both schemes can lead to one order higher convergence. The postprocessed velocity obtained from the first type of postprocessing is not $H(\text{div}; \Omega)$ -conforming and is divergence-free on each element only when piecewise constant approximations are exploited. On the other hand, the second type of postprocessing is relatively complex to construct, but it offers an $H(\text{div}; \Omega)$ -conforming and divergence-free postprocessed velocity for arbitrarily polynomial orders of the approximated velocity. We emphasize that the second type of postprocessing is inspired by the postprocessing given in [14], where the authors consider the postprocessing in each primal mesh, which cannot naturally extend to the polygonal mesh considered in this paper. Thus, we modify the postprocessing scheme given there and make it applicable to general meshes. A rigorous error analysis is carried out for L^2 errors of pseudostress and velocity; in addition, we also prove the superconvergence for velocity. Guided by the superconvergence, the postprocessing approximations can be proved to have one order higher convergence.

The rest of the paper is organized as follows. In section 2, an arbitrary high order SDG method is proposed on general polygonal meshes and its hybridized formulation is derived. In section 3, the well-posedness of local and global problems is discussed. A priori error analysis is established for the proposed method in section 4. Two types

of postprocessing are introduced in section 5 along with the properties of the postprocessed velocity. Numerical experiments are conducted to verify the theoretical results in the preceding sections. Also, convergence behaviors with various mesh shapes are investigated to demonstrate the robustness of our method to mesh distortion. Finally, some concluding remarks are given at the end of the paper.

2. Staggered discontinuous Galerkin method. Let $\Omega \in \mathbb{R}^2$ be a bounded simply connected polygonal domain with Lipschitz boundary $\partial\Omega$. For given data $\mathbf{f} \in [L^2(\Omega)]^2$ and $\mathbf{g} \in [H^{1/2}(\partial\Omega)]^2$, the Stokes problem for the unknown velocity \mathbf{u} and the pressure p reads

$$(2.1) \quad \begin{aligned} \Delta\mathbf{u} + \nabla p &= \mathbf{f} && \text{in } \Omega, \\ \nabla \cdot \mathbf{u} &= 0 && \text{in } \Omega, \\ \mathbf{u} &= \mathbf{g} && \text{on } \partial\Omega. \end{aligned}$$

Introducing a pseudostress $\boldsymbol{\sigma}$ defined by $\boldsymbol{\sigma} = \nabla\mathbf{u} + p\mathbf{I}$ in the framework of [7] allows us to recast the model problem (2.1) into the pseudostress-velocity formulation

$$(2.2) \quad \begin{aligned} \mathcal{A}\boldsymbol{\sigma} - \nabla\mathbf{u} &= 0 && \text{in } \Omega, \\ \nabla \cdot \boldsymbol{\sigma} &= \mathbf{f} && \text{in } \Omega, \\ \mathbf{u} &= \mathbf{g} && \text{on } \partial\Omega, \\ \int_{\Omega} \text{tr}(\boldsymbol{\sigma}) \, d\mathbf{x} &= 0. \end{aligned}$$

Here \mathbf{I} is the identity matrix and $\mathcal{A} : \mathbb{R}^{2 \times 2} \rightarrow \mathbb{R}^{2 \times 2}$ is the deviatoric operator defined by

$$\mathcal{A}\boldsymbol{\sigma} = \boldsymbol{\sigma} - \frac{1}{2}\text{tr}(\boldsymbol{\sigma})\mathbf{I}.$$

The first equation of (2.2) is obtained from

$$\text{tr}(\nabla\mathbf{u}) = \nabla \cdot \mathbf{u} = 0, \quad \text{tr}(\boldsymbol{\sigma}) = \text{tr}(\nabla\mathbf{u}) + 2p = 2p.$$

We now introduce some basic notation regarding staggered meshes that will be exploited in the construction of the SDG method. Let \mathcal{S} be a star-shaped polygonal partition of Ω , and the set of all the edges is called primal edges and is denoted by \mathcal{F}_{pr} . In addition, we use \mathcal{F}_{pr}^i and \mathcal{F}_{pr}^b to stand for the interior edges and boundary edges, respectively. For each primal element $T \in \mathcal{S}$, we select an interior point ν and connect it to all the vertices of T , and thereby subtriangular grids are generated during this process. The resulting triangulation is denoted as \mathcal{T}_h and the edges generated during the subdivision process are called dual edges and are denoted by \mathcal{F}_{dl} . We rename the union of subtriangles sharing the common vertex ν as $S(\nu)$. Here, we choose ν to be an interior point of the kernel of $S(\nu) \in \mathcal{T}_u$ and we use \mathcal{N} to represent the union of all interior points ν . For each interior edge $e \in \mathcal{F}_{pr}^i$, we use $\mathcal{D}(e)$ to stand for the dual mesh, which is the union of two triangles in \mathcal{T}_h sharing the common edge e . For each boundary edge $e \in \mathcal{F}_{pr}^b$, we use $\mathcal{D}(e)$ to denote a triangle in \mathcal{T}_h having the edge e ; see Figure 1 for an illustration.

For later analysis, we employ the general mesh regularity assumption (cf. [9, 37]): For every element $S(\nu) \in \mathcal{S}$ and every edge $e \in \partial S(\nu)$, it satisfies $h_e \geq \rho h_{S(\nu)}$ for a positive constant ρ , where h_e is the length of edge e and $h_{S(\nu)}$ is a diameter of $S(\nu)$. Second, each element $S(\nu)$ in \mathcal{S} is star-shaped with respect to a ball of radius $\geq \rho h_{S(\nu)}$. We remark that the above assumptions ensure that the triangulation \mathcal{T}_h is shape regular. We employ the general mesh regularity assumption just for the sake of

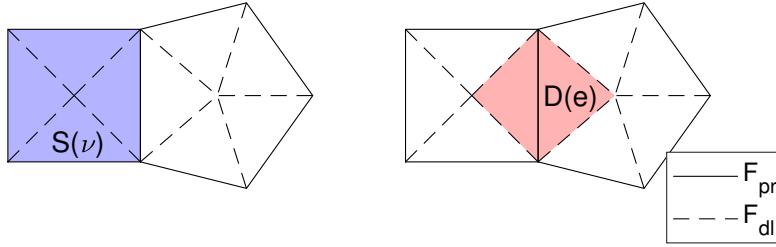


FIG. 1. Schematic of primal and dual meshes. Solid lines are primal edges \mathcal{F}_{pr} and dashed lines are dual edges \mathcal{F}_{dl} . A polygon surrounded by primal edges is called primal element $S(v)$ and triangles surrounded by dual edges are called dual element $D(e)$.

simplicity. In fact, our numerical results given in section 6 indicate that the proposed method actually allows elements with arbitrarily small edges and rigorous analysis for a mesh with small edges will be present in our future work.

The jump $[\![\cdot]\!]$ and the average $\{\!\!\{\cdot\}\!\!\}_e$ is defined by $[\![v]\!]_e = v|_{T_+} - v|_{T_-}$ and $\{\!\!\{v\}\!\!\}_e = (v|_{T_+} + v|_{T_-})/2$, where T_+ and T_- are two elements in \mathcal{T}_h sharing the common edge $e \in \mathcal{F}_{pr}^i \cup \mathcal{F}_{dl}$. For $e \in \mathcal{F}_{pr}^b$, we simply take $[\![v]\!]_e = \{\!\!\{v\}\!\!\}_e = v|_{T_+}$. The subscript e will be omitted when there is no ambiguity. We denote the L^2 -inner product by $(f, g)_\tau = \int_\tau f g \, dx$ for two dimensions and $\langle f, g \rangle_e = \int_e f g \, ds$ for one dimension. When \mathbf{f} and \mathbf{g} are vectors (or tensors), then $(\cdot, \cdot)_\tau$ and $\langle \cdot, \cdot \rangle_e$ are defined by their componentwise sum. That is, when $\mathbf{f}, \mathbf{g} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$,

$$(\mathbf{f}, \mathbf{g})_\tau = \int_\tau \mathbf{f} \cdot \mathbf{g} \, dx, \quad \langle \mathbf{f}, \mathbf{g} \rangle_e = \int_e \mathbf{f} \cdot \mathbf{g} \, ds,$$

and when $\mathbf{f}, \mathbf{g} : \mathbb{R}^2 \rightarrow \mathbb{R}^{2 \times 2}$,

$$(\mathbf{f}, \mathbf{g})_\tau = \int_\tau \mathbf{f} : \mathbf{g} \, dx, \quad \langle \mathbf{f}, \mathbf{g} \rangle_e = \int_e \mathbf{f} : \mathbf{g} \, ds.$$

Here, $A : B = \sum_{ij} A_{ij} B_{ij}$ is the Frobenius inner product. The discrete spaces are defined based on the definition described above. Let (X_h, V_h) be the discrete spaces defined by

$$\begin{aligned} X_h &= \{\psi \in [L^2(\Omega)]^{2 \times 2} : \psi|_\tau \in [\mathbb{P}_k(\tau)]^{2 \times 2} \forall \tau \in \mathcal{T}_h, [\![\psi \mathbf{n}]\!] = \mathbf{0} \text{ on } \mathcal{F}_{dl}, (tr(\psi), 1)_\Omega = 0\}, \\ V_h &= \{\mathbf{v} \in [L^2(\Omega)]^2 : \mathbf{v}|_\tau \in [\mathbb{P}_k(\tau)]^2 \forall \tau \in \mathcal{T}_h, [\![\mathbf{v}]\!] = \mathbf{0} \text{ on } \mathcal{F}_{pr}\}, \end{aligned}$$

where $\mathbb{P}_k(\tau)$ is a complete polynomial space of degree less than or equal to k on a triangle τ . Also, we define V_h^g by

$$V_h^g = \{\mathbf{v} \in [L^2(\Omega)]^2 : \mathbf{v}|_\tau \in \mathbb{P}_k(\tau) \forall \tau \in \mathcal{T}_h, [\![\mathbf{v}]\!] = \mathbf{0} \text{ on } \mathcal{F}_{pr}^i, \mathbf{v}|_e = \Pi_h \mathbf{g} \forall e \in \mathcal{F}_{pr}^b\},$$

where Π_h is the L^2 projection from $L^2(\partial\Omega)$ onto $\mathbb{P}_k(e)$ for each $e \in \mathcal{F}_{pr}$. Here, $\mathbb{P}_k(e)$ is a complete polynomial space of degree less than or equal to k on an edge e . We can also verify the following lemma in the spirit of [17].

LEMMA 2.1 (degrees of freedom). *Any function $\psi \in X_h$ is uniquely determined by the following degrees of freedom:*

(XD1) *For $e \in \mathcal{F}_{pr}$, we have*

$$\Phi_e(\psi) := \int_e \psi \mathbf{n} \cdot \mathbf{p}_k \, ds \quad \forall \mathbf{p}_k \in [\mathbb{P}_k(e)]^2.$$

(XD2) For each $\tau \in \mathcal{T}_h$, we can obtain

$$\Phi_\tau(\psi) := \int_\tau \psi : \mathbf{p}_{k-1} \, d\mathbf{x} \quad \forall \mathbf{p}_{k-1} \in [\mathbb{P}_{k-1}(\tau)]^{2 \times 2}.$$

Similarly, any function $\mathbf{v} \in V_h$ is uniquely determined by the following degrees of freedom:

(VD1) For $e \in \mathcal{F}_{pr}$, we have

$$\phi_e(\mathbf{v}) := \int_e \mathbf{v} \cdot \mathbf{p}_k \, ds \quad \forall \mathbf{p}_k \in [\mathbb{P}_k(e)]^2.$$

(VD2) For each $\tau \in \mathcal{T}_h$, we can obtain

$$\phi_\tau(\mathbf{v}) := \int_\tau \mathbf{v} : \mathbf{p}_{k-1} \, d\mathbf{x} \quad \forall \mathbf{p}_{k-1} \in [\mathbb{P}_{k-1}(\tau)]^2.$$

Multiplying (2.2) by the corresponding test functions and performing integration by parts as in [37, 40], one can obtain the following: Find $(\boldsymbol{\sigma}_h, \mathbf{u}_h) \in X_h \times V_h^g$ such that

$$(2.3) \quad (\mathcal{A}\boldsymbol{\sigma}_h, \psi)_\Omega + b_h^*(\mathbf{u}_h, \psi) = 0 \quad \forall \psi \in X_h,$$

$$(2.4) \quad b_h(\boldsymbol{\sigma}_h, \mathbf{v}) = (\mathbf{f}, \mathbf{v})_\Omega \quad \forall \mathbf{v} \in V_h,$$

where

$$\begin{aligned} b_h^*(\mathbf{u}_h, \psi) &= \sum_{\tau \in \mathcal{T}_h} (\mathbf{u}_h, \nabla \cdot \psi)_\tau - \sum_{e \in \mathcal{F}_{pr}} \langle \mathbf{u}_h, [\![\psi \mathbf{n}]\!] \rangle_e, \\ b_h(\boldsymbol{\sigma}_h, \mathbf{v}) &= - \sum_{\tau \in \mathcal{T}_h} (\boldsymbol{\sigma}_h, \nabla \mathbf{v})_\tau + \sum_{e \in \mathcal{F}_{dl}} \langle \boldsymbol{\sigma}_h \mathbf{n}, [\![\mathbf{v}]\!] \rangle_e. \end{aligned}$$

The following inf-sup condition holds as shown in [17].

$$(2.5) \quad \inf_{\mathbf{v} \in V_h} \sup_{\psi \in X_h} \frac{b_h(\psi, \mathbf{v})}{\|\psi\|_{L^2(\Omega)} \|\mathbf{v}\|_V} \geq C,$$

where

$$\|\mathbf{v}\|_V^2 = \sum_{\tau \in \mathcal{T}_h} \|\nabla \mathbf{v}\|_{L^2(\tau)}^2 + \sum_{e \in \mathcal{F}_{dl}} h_e^{-1} \|[\![\mathbf{v}]\!]\|_{L^2(e)}^2.$$

Performing integration by parts reveals the following discrete adjoint property:

$$(2.6) \quad b_h(\psi, \mathbf{v}) = b_h^*(\mathbf{v}, \psi) \quad \forall (\psi, \mathbf{v}) \in X_h \times V_h.$$

The above formulation can work for our current setting, while the main drawback is more degrees of freedom are introduced due to the subdivision of the primal meshes. Furthermore, the construction of basis functions in particular for high order is also tricky. Therefore, we aim to alleviate the computational burden. To this end, we introduce a Lagrange multiplier $\boldsymbol{\lambda}$ on each dual edge to impose the normal directional continuity of $\boldsymbol{\sigma}_h$. The introduction of the Lagrange multiplier not only relaxes the constraint on the discrete space X_h , but also allows us to consider hybridization of the given formulation in (2.3)–(2.4). In order to hybridize the discrete problem, we

define local spaces \tilde{X}_h^e and V_h^e for each $e \in \mathcal{F}_{pr}$:

$$\begin{aligned}\tilde{X}_h^e &= \{\tilde{\psi} \in [L^2(\mathcal{D}(e))]^{2 \times 2} : \tilde{\psi}|_\tau \in [\mathbb{P}_k(\tau)]^{2 \times 2} \forall \tau \in \mathcal{D}(e), \sum_{\tau \in \mathcal{D}(e)} \int_\tau \operatorname{tr}(\tilde{\psi}) \, d\mathbf{x} = 0\}, \\ V_h^e &= \{\mathbf{v} \in [L^2(\mathcal{D}(e))]^2 : \mathbf{v}|_\tau \in [\mathbb{P}_k(\tau)]^2 \forall \tau \in \mathcal{D}(e), [\mathbf{v}]_e = \mathbf{0}\}.\end{aligned}$$

Note that when $e \subset \partial\Omega$, $[\mathbf{v}_h] = \mathbf{0}$ implies $\mathbf{v}_h|_e = \mathbf{0}$. It is also worth mentioning that $\tilde{X}_h = \bigoplus \tilde{X}_h^e \not\subset X_h$ since $\tilde{\psi} \in \tilde{X}_h$ is not continuous in the normal direction. Similar to V_h^g , we define $V_h^{e,g} \subset V_h^e$ by strongly imposing the Dirichlet boundary condition when $e \subset \partial\Omega$. Finally, the global spaces $P_{\mathcal{D}}$ and M_{dl} are defined by

$$\begin{aligned}P_{\mathcal{D}} &= \{\psi_0 \in L^2(\Omega) : \psi_0|_{\mathcal{D}(e)} \in \mathbb{P}_0(\mathcal{D}(e)) \forall e \in \mathcal{F}_{pr}, \int_\Omega \psi_0 \, d\mathbf{x} = 0\}, \\ M_{dl} &= \{\boldsymbol{\mu} \in [L^2(\mathcal{F}_{dl})]^2 : \boldsymbol{\mu}|_e \in [\mathbb{P}_k(e)]^2 \forall e \in \mathcal{F}_{dl}\}.\end{aligned}$$

Here, $P_{\mathcal{D}}$ accounts for the cell average of $\operatorname{tr}(\boldsymbol{\sigma})$ so that

$$X_h = \{\tilde{\psi} + \psi_0 \mathbf{I} : \tilde{\psi}|_{\mathcal{D}(e)} \in \tilde{X}_h^e \forall e \in \mathcal{F}_{pr}, \psi_0 \in P_{\mathcal{D}}\} \cap [H(\operatorname{div}; \mathcal{S})]^2,$$

where \mathbf{I} is 2×2 matrix and $H(\operatorname{div}; \mathcal{S})$ is a piecewise $H(\operatorname{div})$ space on the primal polygonal partition \mathcal{S} of Ω . The aforementioned discrete spaces are equipped with the following norms:

$$\begin{aligned}\|\tilde{\psi}\|_{X^e}^2 &= \|\tilde{\psi}\|_{L^2(\Omega)}^2 + \sum_{e' \in \mathcal{F}_{dl}^e} h_{e'} \|\tilde{\psi} \mathbf{n}\|_{L^2(e')}^2, \\ \|\mathbf{v}\|_{V^e}^2 &= \sum_{\tau \in \mathcal{D}(e)} \|\nabla \mathbf{v}\|_{L^2(\tau)}^2 + \sum_{e' \in \mathcal{F}_{dl}^e} h_{e'}^{-1} \|\mathbf{v}\|_{L^2(e')}^2, \\ \|\psi_0\|_P^2 &= \sum_{e \in \mathcal{F}_{pr}} h_e \|\psi_0 \mathbf{n}\|_{L^2(e)}^2, \\ \|\boldsymbol{\mu}\|_M^2 &= \sum_{e \in \mathcal{F}_{dl}} h_e^{-1} \|\boldsymbol{\mu}\|_{L^2(e)}^2,\end{aligned}$$

where $\mathcal{F}_{dl}^e \subset \mathcal{F}_{dl}$ is a collection of dual edges adjacent to the dual element $\mathcal{D}(e)$ for given $e \in \mathcal{F}_{pr}$.

For any given $\boldsymbol{\psi} \in X_h$ and for each $e \in \mathcal{F}_{pr}$, define

$$\psi_0|_{\mathcal{D}(e)} = \int_{\mathcal{D}(e)} \frac{1}{2} \operatorname{tr}(\boldsymbol{\psi}) \, d\mathbf{x}, \quad \tilde{\boldsymbol{\psi}} = \boldsymbol{\psi} - \psi_0 \mathbf{I}.$$

Then $\tilde{\boldsymbol{\psi}}|_{\mathcal{D}(e)} \in \tilde{X}_h^e$ and $\psi_0 \in P_{\mathcal{D}}$. Observe that

$$\mathcal{A}\boldsymbol{\psi} = \mathcal{A}\tilde{\boldsymbol{\psi}} + \mathcal{A}(\psi_0 \mathbf{I}) = \mathcal{A}\tilde{\boldsymbol{\psi}}.$$

We define a local problem on $\mathcal{D}(e)$ by, for each $e \in \mathcal{F}_{pr}$, the following: Find $(\tilde{\boldsymbol{\sigma}}_h, \mathbf{u}_h) \in \tilde{X}_h^e \times V_h^{e,g}$ such that

(2.7)

$$\begin{aligned}\sum_{\tau \in \mathcal{D}(e)} (\mathcal{A}\tilde{\boldsymbol{\sigma}}_h, \tilde{\boldsymbol{\psi}})_\tau + \sum_{\tau \in \mathcal{D}(e)} (\mathbf{u}_h, \nabla \cdot \tilde{\boldsymbol{\psi}})_\tau - \langle \mathbf{u}_h, [\tilde{\boldsymbol{\psi}} \mathbf{n}] \rangle_e &= \sum_{\tau \in \mathcal{D}(e)} \langle \boldsymbol{\lambda}, \tilde{\boldsymbol{\psi}} \mathbf{n}_\tau \rangle_{\partial\tau \setminus e} \quad \forall \tilde{\boldsymbol{\psi}} \in \tilde{X}_h^e, \\ (2.8) \quad - \sum_{\tau \in \mathcal{D}(e)} (\tilde{\boldsymbol{\sigma}}_h, \nabla \mathbf{v})_\tau + \sum_{\tau \in \mathcal{D}(e)} \langle \tilde{\boldsymbol{\sigma}}_h \mathbf{n}_\tau, \mathbf{v} \rangle_{\partial\tau \setminus e} &= \sum_{\tau \in \mathcal{D}(e)} (\mathbf{f}, \mathbf{v})_\tau \quad \forall \mathbf{v} \in V_h^e.\end{aligned}$$

Notice that the solution $(\tilde{\boldsymbol{\sigma}}_h, \mathbf{u}_h)$ to (2.7)–(2.8) is determined by the $\boldsymbol{\lambda}$ and \mathbf{f} . This allows us to rewrite the solution by $\tilde{\boldsymbol{\sigma}}_h(\boldsymbol{\lambda}, \mathbf{f})$ and $\mathbf{u}_h(\boldsymbol{\lambda}, \mathbf{f})$. Due to the linearity, we have $\tilde{\boldsymbol{\sigma}}_h(\boldsymbol{\lambda}, \mathbf{f}) = \tilde{\boldsymbol{\sigma}}_h(\boldsymbol{\lambda}, \mathbf{0}) + \tilde{\boldsymbol{\sigma}}_h(\mathbf{0}, \mathbf{f})$ and $\mathbf{u}_h(\boldsymbol{\lambda}, \mathbf{f}) = \mathbf{u}_h(\boldsymbol{\lambda}, \mathbf{0}) + \mathbf{u}_h(\mathbf{0}, \mathbf{f})$. Then $\boldsymbol{\lambda}$ and σ_0 are globally coupled by the normal stress continuity and the locally divergence-free condition:

$$(2.9) \quad \sum_{e \in \mathcal{F}_{dl}} \langle (\llbracket \tilde{\boldsymbol{\sigma}}_h(\boldsymbol{\lambda}, \mathbf{0}) \rrbracket + \llbracket \sigma_0 \mathbf{I} \rrbracket) \mathbf{n}, \boldsymbol{\mu} \rangle_e = - \sum_{e \in \mathcal{F}_{dl}} \langle \llbracket \tilde{\boldsymbol{\sigma}}_h(\mathbf{0}, \mathbf{f}) \mathbf{n} \rrbracket, \boldsymbol{\mu} \rangle_e \quad \forall \boldsymbol{\mu} \in M_{du},$$

$$(2.10) \quad \sum_{\tau \in \mathcal{D}(e)} \int_{\partial \tau \setminus e} \boldsymbol{\lambda} \cdot \mathbf{n}_\tau \, ds = - \int_{e \cap \partial \Omega} \mathbf{g} \cdot \mathbf{n} \, ds \quad \forall e \in \mathcal{F}_{pr}.$$

Remark 2.2. The local discrete problem (2.7)–(2.8) can be viewed as the approximation of $(\tilde{\boldsymbol{\sigma}}, \tilde{\mathbf{u}})$ such that for given $\boldsymbol{\lambda}$ and \mathbf{f} we solve the following local problem on each dual element $\mathcal{D}(e)$:

$$\mathcal{A}\tilde{\boldsymbol{\sigma}} - \nabla \tilde{\mathbf{u}} = \mathbf{C} \quad \text{in } \mathcal{D}(e),$$

$$\nabla \cdot \tilde{\boldsymbol{\sigma}} = \mathbf{f} \quad \text{in } \mathcal{D}(e),$$

$$\tilde{\mathbf{u}} = \boldsymbol{\lambda} \quad \text{on } \mathcal{F}_{dl}^e,$$

$$\tilde{\mathbf{u}} = \mathbf{g} \quad \text{on } \partial \Omega,$$

$$\sum_{\tau \in \mathcal{D}(e)} \int_{\tau} \operatorname{tr}(\tilde{\boldsymbol{\sigma}}) \, d\mathbf{x} = 0,$$

where \mathbf{C} is a piecewise constant on each $\mathcal{D}(e)$ so that the compatibility is satisfied. Also, the local divergence-free condition (2.10) is equivalent to

$$(2.11) \quad \sum_{e \in \mathcal{F}_{dl}} \langle \boldsymbol{\lambda}, \llbracket \psi_0 \mathbf{n} \rrbracket \rangle_e = - \sum_{e \in \mathcal{F}_{pr}^b} \langle \mathbf{g}, \psi_0 \mathbf{n} \rangle_e \quad \forall \psi_0 \in P_{\mathcal{D}}.$$

By taking $\psi_0 = \chi_{\mathcal{D}(e)}$ in (2.11) with the characteristic function χ , we can observe that the compatibility condition is satisfied with $\mathbf{C} \equiv \mathbf{0}$ when $\boldsymbol{\lambda}$ is a solution to (2.9).

Remark 2.3. If $(\tilde{\boldsymbol{\sigma}}_h, \mathbf{u}_h, \boldsymbol{\lambda}, \sigma_0)$ solves (2.7)–(2.11), then $(\tilde{\boldsymbol{\sigma}}_h + \sigma_0 \mathbf{I}, \mathbf{u}_h)$ is a solution to (2.3)–(2.4).

In the remainder of this paper, we assume that $\mathbf{g} = \mathbf{0}$ for simplicity. We denote

$$\begin{aligned} a_h^e(\tilde{\boldsymbol{\sigma}}, \tilde{\boldsymbol{\psi}}) &= \sum_{\tau \in \mathcal{D}(e)} (\mathcal{A}\tilde{\boldsymbol{\sigma}}, \tilde{\boldsymbol{\psi}})_\tau, & b_h^e(\tilde{\boldsymbol{\psi}}, \mathbf{v}) &= - \sum_{\tau \in \mathcal{D}(e)} (\tilde{\boldsymbol{\psi}}, \nabla \mathbf{v})_\tau + \sum_{e' \in \mathcal{F}_{dl}^e} \left\langle \tilde{\boldsymbol{\psi}} \mathbf{n}, \mathbf{v} \right\rangle_{e'}, \\ c_h^e(\tilde{\boldsymbol{\psi}}, \boldsymbol{\lambda}) &= \sum_{e' \in \mathcal{F}_{dl}^e} \left\langle \boldsymbol{\lambda}, \tilde{\boldsymbol{\psi}} \mathbf{n} \right\rangle_{e'}, \\ c_h(\boldsymbol{\lambda}, \boldsymbol{\mu}) &= \sum_{e \in \mathcal{F}_{dl}} \langle \llbracket \tilde{\boldsymbol{\sigma}} \mathbf{n} \rrbracket, \boldsymbol{\mu} \rangle_e, & d_h(\boldsymbol{\mu}, \psi_0) &= \sum_{e \in \mathcal{F}_{dl}} \langle \llbracket \psi_0 \mathbf{n} \rrbracket, \boldsymbol{\mu} \rangle_e. \end{aligned}$$

Here, $\tilde{\boldsymbol{\sigma}}_\lambda = \tilde{\boldsymbol{\sigma}}_h(\boldsymbol{\lambda}, \mathbf{0})$. Also, we define

$$b_h^{e,*}(\mathbf{v}, \tilde{\boldsymbol{\psi}}) = \sum_{\tau \in \mathcal{D}(e)} (\mathbf{v}_h, \nabla \cdot \tilde{\boldsymbol{\psi}})_\tau - \left\langle \mathbf{v}_h, \llbracket \tilde{\boldsymbol{\psi}} \mathbf{n} \rrbracket \right\rangle_e$$

for each $e \in \mathcal{F}_{pr}$. We can observe that the above operators are continuous with respect to the discrete norms by straightforward applications of the Cauchy–Schwarz inequality.

Remark 2.4 (extension to three dimensions). We remark that we restrict our discussion to \mathbb{R}^2 for simplicity and the current approach can also be extended to \mathbb{R}^3 . In this case, we first decompose the computational domain in \mathbb{R}^3 into polyhedrons, then the primal elements and dual elements can be generated by connecting an interior point to all the vertices of the polyhedrons as in \mathbb{R}^2 . Based on the primal elements and dual elements, we can then define our finite dimensional spaces as in \mathbb{R}^2 ; more precisely, we impose the continuity for velocity over the primal faces and enforce the continuity for pseudostress via lagrange multiplier. Then we solve the local problem for each dual element and the resulting solutions are obtained by solving a global interface problem.

3. Well-posedness. In this section, we discuss the well-posedness of the discrete problems.

THEOREM 3.1. *Let $S_e : M_{dl}|_{\partial\mathcal{D}(e)} \times L^2(\mathcal{D}(e)) \rightarrow \tilde{X}_h^e \times V_h^e$ be the solution operator such that*

$$S_e(\boldsymbol{\lambda}, \mathbf{f}) = (\tilde{\boldsymbol{\sigma}}_h, \mathbf{u}_h),$$

where $(\tilde{\boldsymbol{\sigma}}_h, \mathbf{u}_h)$ solves (2.7)–(2.8). Then S_e is well-defined.

Proof. By Lemmas A.3 and A.4 and the continuity of each operator, the discrete problem is well-posed. \square

To ease our analysis, we define a seminorm on M_{dl} by

$$|\boldsymbol{\mu}|_\sigma^2 = \sum_{e \in \mathcal{F}_{pr}} \|\tilde{\boldsymbol{\sigma}}_\mu\|_{X^e}^2 = \sum_{e \in \mathcal{F}_{pr}} \|\tilde{\boldsymbol{\sigma}}_h(\boldsymbol{\mu}, \mathbf{0})\|_{X^e}^2.$$

LEMMA 3.2. *Let $M_{dl}^0 \subset M_{dl}$ defined by*

$$M_{dl}^0 = \{\boldsymbol{\mu} \in M_{dl} : d_h(\boldsymbol{\mu}, \psi_0) = 0 \ \forall \psi_0 \in P_{\mathcal{D}}\}.$$

Then the seminorm $|\cdot|_\sigma$ defines a norm on M_{dl}^0 . Furthermore, there holds

$$\|\boldsymbol{\mu}\|_M \approx |\boldsymbol{\mu}|_\sigma \quad \forall \boldsymbol{\mu} \in M_{dl}^0.$$

Proof. We claim that $|\cdot|_\sigma$ defines a norm on M_{dl}^0 . By the linearity of the solution operator S , it is clear that

$$|a\boldsymbol{\mu}|_\sigma = |a||\boldsymbol{\mu}|_\sigma, \quad |\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2|_\sigma \leq |\boldsymbol{\mu}_1|_\sigma + |\boldsymbol{\mu}_2|_\sigma.$$

Then it is enough to prove that $|\boldsymbol{\mu}|_\sigma = 0$ implies $\boldsymbol{\mu} \equiv \mathbf{0}$. Let $\boldsymbol{\mu} \in M_{dl}^0$ with $|\boldsymbol{\mu}|_\sigma = 0$. Then we have $\tilde{\boldsymbol{\sigma}}_\mu \equiv \mathbf{0}$. Since $(\tilde{\boldsymbol{\sigma}}_\mu, \mathbf{u}_\mu)$ solves (2.7)–(2.8) with $\mathbf{f} \equiv \mathbf{0}$ for each $e \in \mathcal{F}_{pr}$, we have

$$\sum_{\tau \in \mathcal{D}(e)} (\mathbf{u}_\mu, \nabla \cdot \tilde{\boldsymbol{\psi}})_\tau - \left\langle \mathbf{u}_\mu, [\tilde{\boldsymbol{\psi}} \mathbf{n}] \right\rangle_e = \sum_{e' \in \mathcal{F}_{dl}^e} \left\langle \boldsymbol{\mu}, \tilde{\boldsymbol{\psi}} \mathbf{n} \right\rangle_{e'} \quad \forall \tilde{\boldsymbol{\psi}} \in \tilde{X}_h^e.$$

Using integration by parts, we have

$$(3.1) \quad \sum_{\tau \in \mathcal{D}(e)} (\nabla \mathbf{u}_\mu, \tilde{\boldsymbol{\psi}})_\tau + \sum_{e' \in \mathcal{F}_{dl}^e} \left\langle \mathbf{u}_\mu, \tilde{\boldsymbol{\psi}} \mathbf{n} \right\rangle_{e'} = \sum_{e' \in \mathcal{F}_{dl}^e} \left\langle \boldsymbol{\mu}, \tilde{\boldsymbol{\psi}} \mathbf{n} \right\rangle_{e'} \quad \forall \tilde{\boldsymbol{\psi}} \in \tilde{X}_h^e.$$

For any given ξ such that $\xi|_{e'} \in [\mathbb{P}_k(e')]^2$ for all $e' \in \mathcal{F}_{dl}^e$, we can find $\tilde{\psi}$ such that

$$\begin{aligned} (\tilde{\psi}, \mathbf{v})_\tau &= 0 & \forall \mathbf{v} \in [\mathbb{P}_{k-1}(\tau)]^{2 \times 2}, \forall \tau \in \mathcal{D}(e), \\ \langle \tilde{\psi} \mathbf{n}, v \rangle_{e'} &= \langle \xi, v \rangle_{e'} & \forall v \in [\mathbb{P}_k(e')]^2, \forall e' \in \mathcal{F}_{dl}^e. \end{aligned}$$

Then we have

$$\sum_{e' \in \mathcal{F}_{dl}^e} \langle \mathbf{u}_\mu - \boldsymbol{\mu}, \xi \rangle_{e'} = 0.$$

Since both \mathbf{u}_μ and $\boldsymbol{\mu}$ restricted on \mathcal{F}_{dl}^e are piecewise k th order polynomials, this implies $\mathbf{u}_\mu \equiv \boldsymbol{\mu}$ on \mathcal{F}_{dl}^e . By this and the uniqueness of $\boldsymbol{\mu}$ on \mathcal{F}_{dl} , \mathbf{u}_μ is continuous in Ω . On the other hand, $\boldsymbol{\mu} \in M_{dl}^0$ leads to

$$\sum_{\tau \in \mathcal{D}(e)} \int_\tau \operatorname{tr}(\nabla \mathbf{u}_\mu) \, d\mathbf{x} = \sum_{\tau \in \mathcal{D}(e)} \int_\tau \nabla \cdot \mathbf{u}_\mu \, d\mathbf{x} = \sum_{e' \in \mathcal{F}_{dl}^e} \int_{e'} \mathbf{u}_\mu \cdot \mathbf{n} \, ds = \sum_{e' \in \mathcal{F}_{dl}^e} \int_{e'} \boldsymbol{\mu} \cdot \mathbf{n} \, ds = 0 \, ds.$$

Therefore, $\nabla \mathbf{u}_\mu \in \tilde{X}_h^e$. Taking $\tilde{\psi} = \nabla \mathbf{u}_\mu$ in (3.1), we have

$$\sum_{\tau \in \mathcal{D}(e)} \|\nabla \mathbf{u}_\mu\|_{L^2(\tau)}^2 = 0.$$

This implies that \mathbf{u}_μ is some constant vector. Since $\mathbf{u}_\mu = \mathbf{0}$ on $\partial\Omega$, we have $a = b = 0$. Therefore, $\boldsymbol{\mu}|_e = \mathbf{u}_\mu|_e = \mathbf{0}$ for each $e \in \mathcal{F}_{dl}$, which proves that $|\cdot|_\sigma$ defines a norm on M_{dl}^0 . Then the finite dimensionality of M_{dl}^0 and the scaling argument yields the desired result. \square

LEMMA 3.3. *For $\boldsymbol{\mu} \in M_{dl}^0$, there holds*

$$c_h(\boldsymbol{\mu}, \boldsymbol{\mu}) \geq C \|\boldsymbol{\mu}\|_M^2.$$

Remark 2.2 implies that $M_{dl}^0 \equiv \{\boldsymbol{\mu} \in M_{dl} : \sum_{e' \in \mathcal{F}_{dl}^e} \int_{e'} \boldsymbol{\mu} \cdot \mathbf{n} \, ds = 0 \text{ for all } e \in \mathcal{F}_{pr}\}$.

Proof. Taking $\tilde{\psi} = \tilde{\sigma}_\mu$ in (2.7) leads to

$$\sum_{e \in \mathcal{F}_{dl}} \langle [\![\tilde{\sigma}_\mu \mathbf{n}]\!], \boldsymbol{\mu} \rangle_e = \sum_{e \in \mathcal{F}_{pr}} \left[\sum_{\tau \in \mathcal{D}(e)} (\mathcal{A}\tilde{\sigma}_\mu, \tilde{\sigma}_\mu)_\tau + b_h^e(\tilde{\sigma}_\mu, \mathbf{u}_\mu) \right].$$

Taking $\mathbf{v} = \mathbf{u}_\mu$ in (2.8) implies

$$\sum_{e \in \mathcal{F}_{dl}} \langle [\![\tilde{\sigma}_\mu \mathbf{n}]\!], \boldsymbol{\mu} \rangle_e = \sum_{e \in \mathcal{F}_{pr}} \left(\sum_{\tau \in \mathcal{D}(e)} \|\mathcal{A}\tilde{\sigma}_\mu\|_{L^2(\tau)}^2 \right) \geq C \sum_{e \in \mathcal{F}_{pr}} \|\tilde{\sigma}_\mu\|_{X^e}^2 \geq C \|\boldsymbol{\mu}\|_M^2. \quad \square$$

LEMMA 3.4. *There holds*

$$\inf_{\psi_0 \in P_D} \sup_{\boldsymbol{\mu} \in M_{dl}} \frac{d_h(\psi_0, \boldsymbol{\mu})}{\|\psi_0\|_P \|\boldsymbol{\mu}\|_M} = 1.$$

Proof. Let $\psi_0 \in P_D$ be given. Take $\boldsymbol{\mu} = h [\![\psi_0 \mathbf{n}]\!]$. Then $\|\boldsymbol{\mu}\|_M = \|\psi_0\|_P$ and

$$d_h(\psi_0, \boldsymbol{\mu}) = \sum_{e \in \mathcal{F}_{dl}} h_e \|[\![\psi_0 \mathbf{n}]\!]\|_{L^2(e)}^2 = \|\psi_0\|_P^2. \quad \square$$

Lemmas 3.3 and 3.4 prove that the global system (2.9)–(2.11) is well-defined.

4. A priori error estimates. Note that the solution obtained from the hybridization formulation (2.7)–(2.10) is equivalent to the solution obtained from the discrete formulation (2.3)–(2.4). Therefore, the goal of this section is to present the convergence analysis based on the discrete formulation (2.3)–(2.4). To this end, we define two projection operators. We first define the projection operator $I_h^k : [H^1(\Omega)]^2 \rightarrow V_h$ by

$$\begin{aligned}\langle I_h^k \mathbf{v} - \mathbf{v}, \phi \rangle_e &= 0 \quad \forall \phi \in [\mathbb{P}_k(e)]^2, \quad \forall e \in \mathcal{F}_{pr}, \\ (I_h^k \mathbf{v} - \mathbf{v}, \phi)_\tau &= 0 \quad \forall \phi \in [\mathbb{P}_{k-1}(\tau)]^2, \quad \forall \tau \in \mathcal{T}_h.\end{aligned}$$

For simplicity, we write $I_h = I_h^k$ when there is no ambiguity. Then we define $J_h : [H(\text{div}; \Omega)]^2 \rightarrow X_h$ by

$$\begin{aligned}(4.1) \quad \langle (J_h \mathbf{q} - \mathbf{q}) \mathbf{n}, \mathbf{v} \rangle_e &= 0 \quad \forall \mathbf{v} \in [\mathbb{P}_k(e)]^2, \quad \forall e \in \mathcal{F}_{dl}, \\ (J_h \mathbf{q} - \mathbf{q}, \phi)_\tau &= 0 \quad \forall \phi \in [\mathbb{P}_{k-1}(\tau)]^{2 \times 2}, \quad \forall \tau \in \mathcal{T}_h.\end{aligned}$$

By performing integration by parts on (2.3)–(2.4), we obtain

$$\begin{aligned}(4.2) \quad (\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h), \psi)_\Omega &= b_h^*(\mathbf{u}_h - I_h \mathbf{u}, \psi) \quad \forall \psi \in X_h, \\ b_h(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{v}) &= 0 \quad \forall \mathbf{v} \in V_h.\end{aligned}$$

Furthermore, I_h and J_h defined above are polynomial preserving operators and they satisfy the following approximation properties (cf. [17]):

$$(4.3) \quad \|\boldsymbol{\sigma} - J_h \boldsymbol{\sigma}\|_{L^2(\Omega)} \leq Ch^{k+1} \|\boldsymbol{\sigma}\|_{H^{k+1}(\Omega)},$$

$$(4.4) \quad \|\mathbf{u} - I_h \mathbf{u}\|_{L^2(\Omega)} \leq Ch^{k+1} \|\mathbf{u}\|_{H^{k+1}(\Omega)}.$$

LEMMA 4.1. *Let $(\boldsymbol{\sigma}, \mathbf{u})$ be the weak solution of (2.2) and $(\boldsymbol{\sigma}_h, \mathbf{u}_h) \in X_h \times V_h^g$ be the discrete solution of (2.3)–(2.4), and then we have*

$$\|J_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(\Omega)} \approx \|\mathcal{A}(J_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_{L^2(\Omega)} \leq \|\mathcal{A}(J_h \boldsymbol{\sigma} - \boldsymbol{\sigma})\|_{L^2(\Omega)}.$$

In addition, we have

$$\|I_h \mathbf{u} - \mathbf{u}_h\|_V \leq C \|\mathcal{A}(J_h \boldsymbol{\sigma} - \boldsymbol{\sigma})\|_{L^2(\Omega)}.$$

Proof. Let $X_h^\perp = \{\psi \in X_h : b_h(\psi, \mathbf{v}) = 0 \text{ for all } \mathbf{v} \in V_h\}$. Then with minor modification to the proof of Lemma A.2, we have $\|\mathcal{A}(\cdot)\|_{L^2(\Omega)}$ and $\|\cdot\|_{L^2(\Omega)}$ are equivalent. If $J_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h \in X_h^\perp$, then the first equivalence relation is proved. It is clear that $(J_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h)|_\tau \in [\mathbb{P}_k(\tau)]^{2 \times 2}$ for each $\tau \in \mathcal{T}_h$ and $[(J_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h) \mathbf{n}]|_e = \mathbf{0}$ for each $e \in \mathcal{F}_{dl}$. Therefore, it suffices to prove that $\int_\Omega \text{tr}(J_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h) \, d\mathbf{x} = 0$.

When $k \geq 1$, we have $\mathbf{I} \in [\mathbb{P}_0(\tau)]^{2 \times 2} \subset [\mathbb{P}_{k-1}(\tau)]^{2 \times 2}$ for all $\tau \in \mathcal{T}_h$. Therefore, we have

$$\int_\Omega \text{tr}(J_h \boldsymbol{\sigma}) \, d\mathbf{x} = (J_h \boldsymbol{\sigma}, \mathbf{I})_\Omega = (\boldsymbol{\sigma}, \mathbf{I})_\Omega = \int_\Omega \text{tr}(\boldsymbol{\sigma}) \, d\mathbf{x} = 0.$$

Since $\boldsymbol{\sigma}_h \in X_h$, $(J_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h) \in X_h$. Now, let $k = 0$. For simplicity, we define $\Phi_h = J_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h$.

$$\begin{aligned}\int_\Omega \text{tr}(\Phi_h) \, d\mathbf{x} &= (\Phi_h, \mathbf{I})_\Omega \\ &= \sum_{\tau \in \mathcal{T}_h} (\Phi_h, \nabla \mathbf{x})_\tau \\ &= \sum_{\tau \in \mathcal{T}_h} \langle \Phi_h \mathbf{n}, \mathbf{x} \rangle_{\partial\tau}.\end{aligned}$$

Here, the last equality is obtained by using integration by parts and $\Phi_h|_\tau \in [\mathbb{P}_0(\tau)]^{2 \times 2}$ for all $\tau \in \mathcal{T}_h$. By using jump notation and normal continuity of Φ_h , we have

$$\int_{\Omega} \operatorname{tr}(\Phi_h) d\mathbf{x} = \sum_{e \in \mathcal{F}_h} \langle [\![\Phi_h \mathbf{n}]\!], \mathbf{x} \rangle_e = \sum_{e \in \mathcal{F}_{pr}} \langle [\![\Phi_h \mathbf{n}]\!], \mathbf{x} \rangle_e.$$

Observing that $[\![\Phi_h \mathbf{n}]\!] \cdot \mathbf{x} \in \mathbb{P}_1(e)$ for all $e \in \mathcal{F}_{pr}$, one point integration yields

$$(4.5) \quad \int_{\Omega} \operatorname{tr}(\Phi_h) d\mathbf{x} = \sum_{e \in \mathcal{F}_{pr}} \langle [\![\Phi_h \mathbf{n}]\!], \bar{\mathbf{x}}_e \rangle_e,$$

where $\bar{\mathbf{x}}_e$ is the midpoint of $e \in \mathcal{F}_{pr}$. Since $\Phi_h \in [\mathbb{P}_0(\tau)]^{2 \times 2}$, $\nabla \cdot \Phi_h|_\tau = \mathbf{0}$ for each $\tau \in \mathcal{T}_h$. This reads for all $\tau \in \mathcal{T}_h$ and $e \in \mathcal{F}_{pr}$ with nonempty $\partial\tau \cap e$,

$$\int_e \Phi_h \mathbf{n}_\tau ds = - \int_{\partial\tau \setminus e} \Phi_h \mathbf{n}_\tau ds.$$

Using this, we can rewrite (4.5) by

$$(4.6) \quad \int_{\Omega} \operatorname{tr}(\Phi_h) d\mathbf{x} = \sum_{\tau \in \mathcal{T}_h} - \langle \Phi_h \mathbf{n}_\tau, \bar{\mathbf{x}}_\tau \rangle_{\partial\tau \setminus \mathcal{F}_{pr}},$$

where $\bar{\mathbf{x}}_\tau = \bar{\mathbf{x}}_e$ with $e \in \partial\tau \cap \mathcal{F}_{pr}$. Define $\mathbf{x}_h \in V_h$ by $\mathbf{x}_h|_\tau = \bar{\mathbf{x}}_\tau$. By using the definition of \mathbf{x}_h , we rewrite (4.6) by

$$\int_{\Omega} \operatorname{tr}(\Phi_h) d\mathbf{x} = - \sum_{e \in \mathcal{F}_{dl}} \langle \Phi_h \mathbf{n}, [\![\mathbf{x}_h]\!] \rangle_e = -b_h(\Phi_h, \mathbf{x}_h) = -b_h(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{x}_h) = 0,$$

where in the last two equalities we exploit (4.1) and (4.2).

It remains to show that the second inequality holds. Recall that (cf. (4.2))

$$(\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h), \psi)_\Omega = -b_h^*(I_h \mathbf{u} - \mathbf{u}_h, \psi) \quad \forall \psi \in X_h.$$

Then, we can get

$$(4.7) \quad (\mathcal{A}(J_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h), \psi)_\Omega = (\mathcal{A}(J_h \boldsymbol{\sigma} - \boldsymbol{\sigma}), \psi)_\Omega - b_h^*(I_h \mathbf{u} - \mathbf{u}_h, \psi) \quad \forall \psi \in X_h.$$

Taking $\psi = J_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h$, we have

$$\|\mathcal{A}(J_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_{L^2(\Omega)}^2 = (\mathcal{A}(J_h \boldsymbol{\sigma} - \boldsymbol{\sigma}), J_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h)_\Omega - b_h^*(I_h \mathbf{u} - \mathbf{u}_h, J_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h).$$

On the other hand, we have from the discrete adjoint property (2.6), the definition of J_h , and (4.2) that

$$b_h^*(I_h \mathbf{u} - \mathbf{u}_h, J_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h) = b_h(J_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h, I_h \mathbf{u} - \mathbf{u}_h) = b_h(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, I_h \mathbf{u} - \mathbf{u}_h) = 0.$$

Therefore, we have

$$\|\mathcal{A}(J_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_{L^2(\Omega)} \leq \|\mathcal{A}(J_h \boldsymbol{\sigma} - \boldsymbol{\sigma})\|_{L^2(\Omega)}.$$

Finally, we can infer from (2.5), (2.6), and (4.7) that

$$\begin{aligned} \|I_h \mathbf{u} - \mathbf{u}_h\|_V &\leq C \sup_{\psi \in X_h} \frac{b_h(\psi, I_h \mathbf{u} - \mathbf{u}_h)}{\|\psi\|_{L^2(\Omega)}} \\ &= C \sup_{\psi \in X_h} \frac{b_h^*(I_h \mathbf{u} - \mathbf{u}_h, \psi)}{\|\psi\|_{L^2(\Omega)}} \\ &\leq C \|\mathcal{A}(J_h \boldsymbol{\sigma} - \boldsymbol{\sigma})\|_{L^2(\Omega)}. \end{aligned} \quad \square$$

LEMMA 4.2 (superconvergence). *There exists a positive constant C independent of the mesh size such that*

$$\|I_h \mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)} \leq Ch \|\mathcal{A}(J_h \boldsymbol{\sigma} - \boldsymbol{\sigma})\|_{L^2(\Omega)}.$$

Proof. Consider the dual problem

$$\begin{aligned} (4.8) \quad \mathcal{A}\boldsymbol{\eta} - \nabla\varphi &= \mathbf{0} && \text{in } \Omega, \\ \nabla \cdot \boldsymbol{\eta} &= I_h \mathbf{u} - \mathbf{u}_h && \text{in } \Omega, \\ \boldsymbol{\eta} &= \mathbf{0} && \text{on } \partial\Omega, \\ \int_{\Omega} \operatorname{tr}(\boldsymbol{\eta}) &= 0, \end{aligned}$$

which satisfies

$$(4.9) \quad \|\varphi\|_{H^2(\Omega)} + \|\boldsymbol{\eta}\|_{H^1(\Omega)} \leq C \|I_h \mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)}.$$

Integrating by parts and employing (4.8), we can achieve

$$\begin{aligned} \|I_h \mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)}^2 &= (\mathcal{A}\boldsymbol{\eta}, \boldsymbol{\sigma} - \boldsymbol{\sigma}_h)_{\Omega} + \sum_{\tau \in \mathcal{T}_h} \left[(\nabla \cdot \boldsymbol{\eta}, I_h \mathbf{u} - \mathbf{u}_h)_{\tau} - (\nabla\varphi, \boldsymbol{\sigma} - \boldsymbol{\sigma}_h)_{\tau} \right] \\ &= (\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h), \boldsymbol{\eta})_{\Omega} + b_h(\boldsymbol{\eta}, I_h \mathbf{u} - \mathbf{u}_h) + b_h^*(\varphi, \boldsymbol{\sigma} - \boldsymbol{\sigma}_h). \end{aligned}$$

On the other hand, recall that (cf. (4.2))

$$(\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h), J_h \boldsymbol{\eta})_{\Omega} = b_h^*(\mathbf{u}_h - I_h \mathbf{u}, J_h \boldsymbol{\eta}) = -b_h(J_h \boldsymbol{\eta}, I_h \mathbf{u} - \mathbf{u}_h).$$

Also, we infer from the definitions of b_h and J_h

$$b_h(\boldsymbol{\eta} - J_h \boldsymbol{\eta}, \mathbf{v}_h) = 0 \quad \forall \mathbf{v}_h \in V_h.$$

Therefore, we have

$$\begin{aligned} &\|I_h \mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)}^2 \\ &= (\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h), \boldsymbol{\eta} - J_h \boldsymbol{\eta})_{\Omega} + b_h(\boldsymbol{\eta} - J_h \boldsymbol{\eta}, I_h \mathbf{u} - \mathbf{u}_h) + b_h(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \varphi - I_h \varphi) \\ &= (\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h), \boldsymbol{\eta} - J_h \boldsymbol{\eta})_{\Omega} + b_h(\boldsymbol{\sigma} - J_h \boldsymbol{\sigma}, \varphi - I_h \varphi) \\ &\leq Ch \|\boldsymbol{\sigma} - J_h \boldsymbol{\sigma}\|_{L^2(\Omega)} (\|\boldsymbol{\eta}\|_{H^1(\Omega)} + \|\varphi\|_{H^2(\Omega)}) \\ &\leq Ch \|\boldsymbol{\sigma} - J_h \boldsymbol{\sigma}\|_{L^2(\Omega)} \|I_h \mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)}, \end{aligned}$$

where, in the last inequality, we employ (4.9). Dividing both sides by $\|I_h \mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)}$ completes the proof. \square

Combining Lemmas 4.1 and 4.2 with the projection error estimates (4.3)–(4.4), we have the following optimal a priori error estimates.

THEOREM 4.3. *Let $(\mathbf{u}_h, \boldsymbol{\sigma}_h)$ solve (2.3)–(2.4) and $\mathbf{u} \in [H^{k+2}(\Omega)]^2$. Then there holds*

$$\|\mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(\Omega)} \leq Ch^{k+1} \|\mathbf{u}\|_{H^{k+2}(\Omega)}.$$

5. Postprocessing. This section is devoted to the postprocessing of the velocity variable \mathbf{u} . We introduce two kinds of postprocessing: Ritz-projection postprocessing and divergence-free postprocessing. While the first postprocessing is simple and easy to implement, the second postprocessing gives an exact divergence-free velocity field. Also, it is worth mentioning that the first one yields divergence-free postprocessed velocity within each element when $k = 0$.

In the spirit of [1, 11], we propose the following postprocessing:

$$\begin{aligned} I_h^k \mathbf{u}_h^* &= \mathbf{u}_h, \\ (\nabla \mathbf{u}_h^*, \nabla \mathbf{v})_\tau &= (\mathcal{A}\boldsymbol{\sigma}_h, \nabla \mathbf{v})_\tau \quad \forall \mathbf{v} \in (I - I_h^k)W_h^*|_\tau, \forall \tau \in \mathcal{T}_h, \end{aligned}$$

where $W_h^* = \{\mathbf{v} \in [L^2(\Omega)]^2 : \mathbf{v}|_\tau \in [\mathbb{P}_{k+1}(\tau)]^2 \text{ for all } \tau \in \mathcal{T}_h\}$.

The following error equation holds:

$$(5.1) \quad (\nabla(\mathbf{u} - \mathbf{u}_h^*), \nabla \mathbf{v})_\tau = (\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h), \nabla \mathbf{v})_\tau \quad \forall \mathbf{v} \in (I - I_h^k)W_h^*|_\tau, \forall \tau \in \mathcal{T}_h.$$

THEOREM 5.1. *There exists a positive constant C independent of the meshsize such that*

$$\|\mathbf{u} - \mathbf{u}_h^*\|_{L^2(\Omega)} \leq Ch^{k+2} \|\mathbf{u} - \mathbf{u}_h\|_{H^{k+2}(\Omega)}.$$

Proof. Let $\tilde{\mathbf{u}} = I_h^{k+1} \mathbf{u}$, and then $\tilde{\mathbf{u}} \in W_h^*$. Define $\mathbf{v} \in W_h^*$ by $\mathbf{v}|_\tau = (I - I_h^k)(\tilde{\mathbf{u}} - \mathbf{u}_h^*)$ for each $\tau \in \mathcal{T}_h$. Then we have from (5.1) and the Cauchy–Schwarz inequality

$$\begin{aligned} |\mathbf{v}|_{H^1(\tau)}^2 &= (\nabla(\tilde{\mathbf{u}} - \mathbf{u}_h^*), \nabla \mathbf{v})_\tau - (\nabla I_h^k(\tilde{\mathbf{u}} - \mathbf{u}_h^*), \nabla \mathbf{v})_\tau \\ &= (\nabla(\mathbf{u} - \mathbf{u}), \nabla \mathbf{v})_\tau + (\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h), \nabla \mathbf{v})_\tau - (\nabla I_h^k(\tilde{\mathbf{u}} - \mathbf{u}_h^*), \nabla \mathbf{v})_\tau \\ &\leq |\tilde{\mathbf{u}} - \mathbf{u}|_{H^1(\tau)} |\mathbf{v}|_{H^1(\tau)} + \|\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_{L^2(\tau)} |\mathbf{v}|_{H^1(\tau)} \\ &\quad + |I_h^k(\tilde{\mathbf{u}} - \mathbf{u}_h^*)|_{H^1(\tau)} |\mathbf{v}|_{H^1(\tau)}. \end{aligned}$$

Therefore, we have

$$|\mathbf{v}|_{H^1(\tau)} \leq |\tilde{\mathbf{u}} - \mathbf{u}|_{H^1(\tau)} + \|\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_{L^2(\tau)} + |I_h^k(\tilde{\mathbf{u}} - \mathbf{u}_h^*)|_{H^1(\tau)}.$$

Since $(I - I_h^k)\mathbf{w} = 0$ if $\mathbf{w} \in [\mathbb{P}_0(\tau)]^2$, then the Poincaré inequality yields

$$\|\mathbf{v}\|_{L^2(\tau)} \leq Ch_\tau |\mathbf{v}|_{H^1(\tau)}.$$

Noting that $I_h^k(\tilde{\mathbf{u}} - \mathbf{u}_h^*)$ is finite dimensional, the inverse inequality implies

$$|\tilde{\mathbf{u}} - \mathbf{u}_h^*|_{H^1(\tau)} \leq Ch_\tau^{-1} \|I_h^k(\tilde{\mathbf{u}} - \mathbf{u}_h^*)\|_{L^2(\tau)}.$$

By combining these inequalities, we obtain

$$\begin{aligned} \|\mathbf{v}\|_{L^2(\tau)} &\leq Ch_\tau |\mathbf{v}|_{H^1(\tau)} \\ &\leq Ch_\tau \left(|\tilde{\mathbf{u}} - \mathbf{u}|_{H^1(\tau)} + \|\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_{L^2(\tau)} + \|I_h^k(\tilde{\mathbf{u}} - \mathbf{u}_h^*)\|_{L^2(\tau)} \right). \end{aligned}$$

On the other hand, we have from the scaling arguments

$$\begin{aligned} \|I_h^k(\tilde{\mathbf{u}} - \mathbf{u}_h^*)\|_{L^2(\tau)} &\leq Ch_\tau^{1/2} \|I_h^k(\tilde{\mathbf{u}} - \mathbf{u}_h^*)\|_{L^2(\partial\tau \cap \mathcal{F}_{pr})} \\ &= Ch_\tau^{1/2} \|I_h^k \mathbf{u} - \mathbf{u}_h\|_{L^2(\partial\tau \cap \mathcal{F}_{pr})} \\ &\leq C \|I_h^k \mathbf{u} - \mathbf{u}_h\|_{L^2(\tau)}. \end{aligned}$$

Finally, we have

$$\|\mathbf{u} - \mathbf{u}_h^*\|_{L^2(\tau)} \leq \|\mathbf{u} - \tilde{\mathbf{u}}\|_{L^2(\tau)} + \|\mathbf{v}\|_{L^2(\tau)} + \|I_h^k(\tilde{\mathbf{u}} - \mathbf{u}_h^*)\|_{L^2(\tau)}.$$

Combining this with Lemma 4.2 and Theorem 4.3 yields the desired result. \square

When $k = 0$, I_h^k can be defined by ignoring the element constraint. In this case, the postprocessed velocity is divergence-free inside each element.

THEOREM 5.2. *Let \mathbf{u}_h^* be the postprocessed velocity with $k = 0$. Then $\nabla \cdot \mathbf{u}_h^* \equiv 0$ on each $\tau \in \mathcal{T}_h$.*

Proof. Since $\nabla \cdot \mathbf{u}_h^* \in \mathbb{P}_0(\tau)$, it is enough to show that

$$(\nabla \cdot \mathbf{u}_h^*, 1)_\tau = 0 \quad \forall \tau \in \mathcal{T}_h.$$

Applying integration by parts, we have

$$(\nabla \cdot \mathbf{u}_h^*, 1)_\tau = \langle \mathbf{u}_h^* \cdot \mathbf{n}, 1 \rangle_{\partial\tau}.$$

Consider $\mathbf{v} = \mathbf{x} - \bar{\mathbf{x}}$, where $\bar{\mathbf{x}}$ is the midpoint of the primal edge of T . Then there holds

$$I_h^0 \mathbf{v} = \mathbf{0}, \quad \nabla \mathbf{v} = \mathbf{I}, \quad \Delta \mathbf{v} = 0.$$

By the definition of \mathbf{u}_h^* , there holds

$$\begin{aligned} \langle \mathbf{u}_h^* \cdot \mathbf{n}, 1 \rangle_{\partial\tau} &= (\mathbf{u}_h^*, -\Delta \mathbf{v})_\tau + \langle \mathbf{u}_h^*, \nabla \mathbf{v} \mathbf{n} \rangle_{\partial\tau} \\ &= (\nabla \mathbf{u}_h^*, \nabla \mathbf{v})_\tau \\ &= (\mathcal{A} \boldsymbol{\sigma}_h, \mathbf{I})_\tau \\ &= 0, \end{aligned}$$

where we used integration by parts on τ and Lemma A.1. \square

Although the above postprocessing earns one order higher convergence and is easy to implement, the velocity field is not pointwise divergence-free when $k > 0$. Inspired by the postprocessing for HDG methods introduced in [21], the following postprocessing is relatively complicated, but it gives a pointwise divergence-free velocity field. The postprocessed velocity $\mathbf{u}_h^* \in W_h^*$ is determined by the following conditions on each $\tau \in \mathcal{T}_h$: For every edge $e \in \partial\tau$, \mathbf{u}_h^* satisfies

$$(5.2) \quad \begin{aligned} \langle \mathbf{u}_h^* \cdot \mathbf{n}, v \rangle_e &= \langle \hat{\mathbf{u}}_h \cdot \mathbf{n}, v \rangle_e & \forall v \in \mathbb{P}_k(e), \\ \langle (\mathbf{n} \times \nabla)(\mathbf{u}_h^* \cdot \mathbf{n}), (\mathbf{n} \times \nabla)v \rangle_e &= \langle \mathbf{n} \times (\{\{\boldsymbol{\sigma}_h^T\}\} \mathbf{n}), (\mathbf{n} \times \nabla)v \rangle_e & \forall v \in \mathbb{P}_{k+1}^{\perp k}(e), \end{aligned}$$

where $\mathbb{P}_{k+1}^{\perp k}(e) = \{v \in \mathbb{P}_{k+1}(e) : \langle v, w \rangle_e = 0 \text{ for all } w \in \mathbb{P}_k(e)\}$ and

$$\hat{\mathbf{u}}_h = \begin{cases} \mathbf{u}_h & \text{when } e \in \mathcal{F}_{pr}, \\ \boldsymbol{\lambda} & \text{when } e \in \mathcal{F}_{dl}. \end{cases}$$

In addition, when $k \geq 1$, \mathbf{u}_h^* satisfies

$$(5.3) \quad \begin{aligned} (\mathbf{u}_h^*, \nabla v)_\tau &= (\mathbf{u}_h, \nabla v)_\tau & \forall v \in \mathbb{P}_k^{\perp 0}(\tau), \\ (\nabla \times \mathbf{u}_h^*, v \mathcal{B})_\tau &= (\omega_h, v \mathcal{B})_\tau & \forall v \in \mathbb{P}_{k-1}(\tau), \end{aligned}$$

where $\omega_h = (\boldsymbol{\sigma}_h)_{21} - (\boldsymbol{\sigma}_h)_{12}$ is the approximation of the vorticity and \mathcal{B} is the standard bubble function in $\mathbb{P}_3(\tau)$. Here, $\mathbf{n} \times \nabla = \mathbf{n}_x \partial_y - \mathbf{n}_y \partial_x$, $\mathbf{n} \times \mathbf{a} = \mathbf{n}_x \mathbf{a}_2 - \mathbf{n}_y \mathbf{a}_1$, and $\nabla \times \mathbf{u} = \partial_x \mathbf{u}_2 - \partial_y \mathbf{u}_1$. The postprocessing introduced above is quite similar to that of [14], where the postprocessing is based on primal elements (macrotriangles), which fails to work for our current method when general meshes are involved. Therefore, we adapt the method proposed in [14] to meet our purposes by working on each subtriangle.

THEOREM 5.3. Let \mathbf{u}_h^* be the postprocessed velocity defined by (5.2)–(5.3). Then \mathbf{u}_h^* belongs to $H(\operatorname{div}; \Omega)$ and is divergence-free.

Proof. By the definition of \mathbf{u}_h^* (cf. (5.2)), its normal component depends on the single valued function $\hat{\mathbf{u}}_h$ and $\{\!\{\boldsymbol{\sigma}_h\}\!\}$, and thus it is single valued. Therefore, \mathbf{u}_h^* belongs to $H(\operatorname{div}; \Omega)$.

It remains to show that \mathbf{u}_h^* is divergence-free. For each $\tau \in \mathcal{T}_h$, we have $\nabla \cdot \mathbf{u}_h^* \in \mathbb{P}_k(\tau)$. Therefore, if there holds

$$(\nabla \cdot \mathbf{u}_h^*, v)_\tau = 0$$

for all $v \in \mathbb{P}_k(\tau)$, \mathbf{u}_h^* is strongly divergence-free. Let $\tau \in \mathcal{T}_h$ and $v \in \mathbb{P}_k(\tau)$ be given. Then there exists $e \in \mathcal{F}_{pr}$ such that $\tau \in \mathcal{D}(e)$. Take $\tilde{\psi}$ in (2.7) by

$$\begin{aligned} \tilde{\psi} &= (v + c)\mathbf{I} && \text{in } \tau, \\ \tilde{\psi} &= c\mathbf{I} && \text{in } \tau', \end{aligned}$$

where τ' is the adjacent triangle which shares a common edge e with τ and c is a constant properly chosen so that $\tilde{\psi} \in \tilde{X}_h^e$. Since $\ker(\mathcal{A}) = \{v\mathbf{I} : v \text{ is a scalar function}\}$, (2.7) reads

$$(\mathbf{u}_h, \nabla v)_\tau - \langle \mathbf{u}_h \cdot \mathbf{n}, v \rangle_e = \langle \boldsymbol{\lambda} \cdot \mathbf{n}, v \rangle_{\partial\tau \setminus e} + \langle \boldsymbol{\lambda} \cdot \mathbf{n}, c \rangle_{\partial\mathcal{D}(e)}.$$

By Remark 2.2, the last term on the right-hand side is 0. Therefore, we have

$$\begin{aligned} (\nabla \cdot \mathbf{u}_h^*, v)_\tau &= -(\mathbf{u}_h^*, \nabla v)_\tau + \langle \mathbf{u}_h^* \cdot \mathbf{n}, v \rangle_{\partial\tau} \\ &= -(\mathbf{u}_h, \nabla v)_\tau + \langle \mathbf{u}_h \cdot \mathbf{n}, v \rangle_e + \langle \boldsymbol{\lambda} \cdot \mathbf{n}, v \rangle_{\partial\tau \setminus e} \\ &= 0, \end{aligned}$$

where the second identity is from the definition of \mathbf{u}_h^* , (5.2)–(5.3). \square

6. Numerical results. In this section, numerical experiments are performed to verify the theoretical results and performance of proposed methods.

6.1. Example 1: Smooth solution on unit square domain. Consider the unit square domain $\Omega = (0, 1)^2$ and the source function \mathbf{f} is properly chosen so that the solution satisfies

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

where $c = (\cos(1) - 1) \sin(1)$.

The convergence behavior with uniform rectangular mesh is displayed in Figure 2. Both $\boldsymbol{\sigma}_h$ and \mathbf{u}_h converge to the exact solution in $\mathcal{O}(h^{k+1})$, which conforms with the theoretical results in Theorem 4.3. Two types of Postprocessed velocity also converge in $\mathcal{O}(h^{k+2})$. Accuracy between the two postprocessing is negligible. The divergence of the first type of postprocessing converges $\mathcal{O}(h^{k+1})$ except $k = 0$. When $k = 0$, the divergence is compatible with the machine precision. On the other hand, the divergence of the second type of postprocessing is 0 for all k .

6.2. Example 2: L-shaped domain. Consider the L-shaped domain $\Omega = (-1, 1)^2 \setminus [0, 1] \times [-1, 0]$. The source function is set to $\mathbf{f} = \mathbf{0}$. The boundary condition

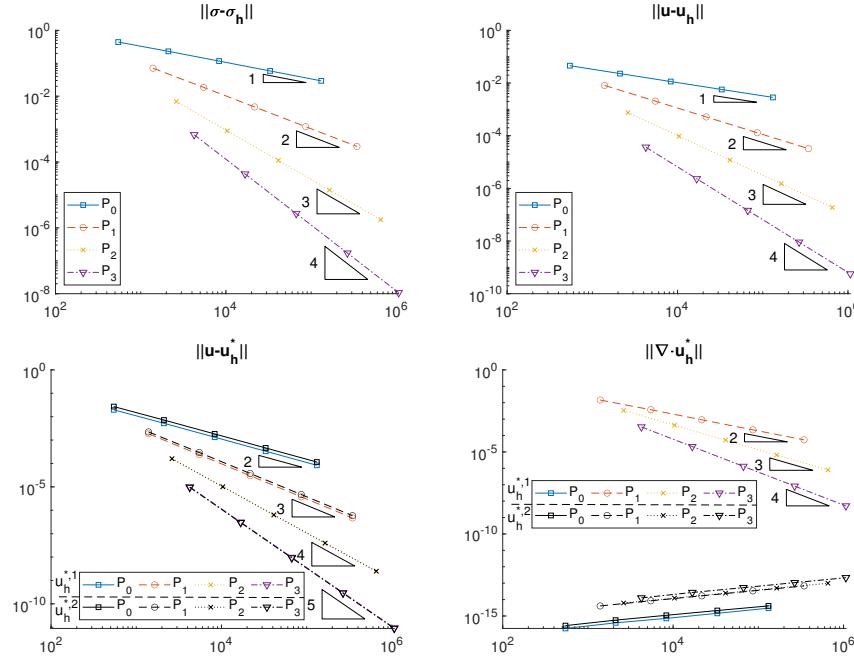


FIG. 2. Convergence history for Example 1 with $k = 0, 1, 2, 3$. The black triangles indicate the convergence rates.

\mathbf{g} is properly chosen so that the solution satisfies

$$\begin{aligned} \mathbf{u}(r, \theta) &= r^\alpha \begin{bmatrix} (1 + \alpha) \sin(\theta) \psi(\theta) + \cos(\theta) \psi'(\theta) \\ -(1 + \alpha) \cos(\theta) \psi(\theta) + \sin(\theta) \psi'(\theta) \end{bmatrix}, \\ p(r, \theta) &= -r^{\alpha-1} ((1 + \alpha)^2 \psi'(\theta) + \psi'''(\theta)) / (1 - \alpha), \end{aligned}$$

where θ is the angle in $(0, 3\pi/2)$ and

$$\begin{aligned} \psi(\theta) &= \sin((1 + \alpha)\theta) \cos(\alpha\omega) / (1 + \alpha) - \cos((1 + \alpha)\theta) \\ &\quad - \sin((1 - \alpha)\theta) \cos(\alpha\omega) / (1 - \alpha) + \cos((1 - \alpha)\theta). \end{aligned}$$

Here, $\omega = 3\pi/2$ and $\alpha \approx 0.54448373678246$ is the smallest positive solution to the equation $\sin(\alpha\omega) + \alpha \sin(\omega) = 0$. The solution (\mathbf{u}, p) is the strongest corner singularity of the Stokes operator in the domain $\Omega = (-1, 1)^2 \setminus [0, 1] \times [-1, 1]$.

The convergence history is depicted in Figure 3. In contrast to Figure 2, the convergence order is bounded by solution regularity, not the polynomial degree except the velocity error with $k = 0$. While the convergence order is not proved for the nonsmooth solution, the results show similar behavior to the theoretical results provided in [37], where $\|\mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)} = \mathcal{O}(h^{2\alpha})$. We note that the convergence order of postprocessed velocity is also bounded by the solution regularity for any choice of the postprocessing or k . While the convergence rate is limited, the postprocessed velocity is pointwise divergence-free when $k = 0$ for the first type and $k \geq 1$ for the second type postprocessing.

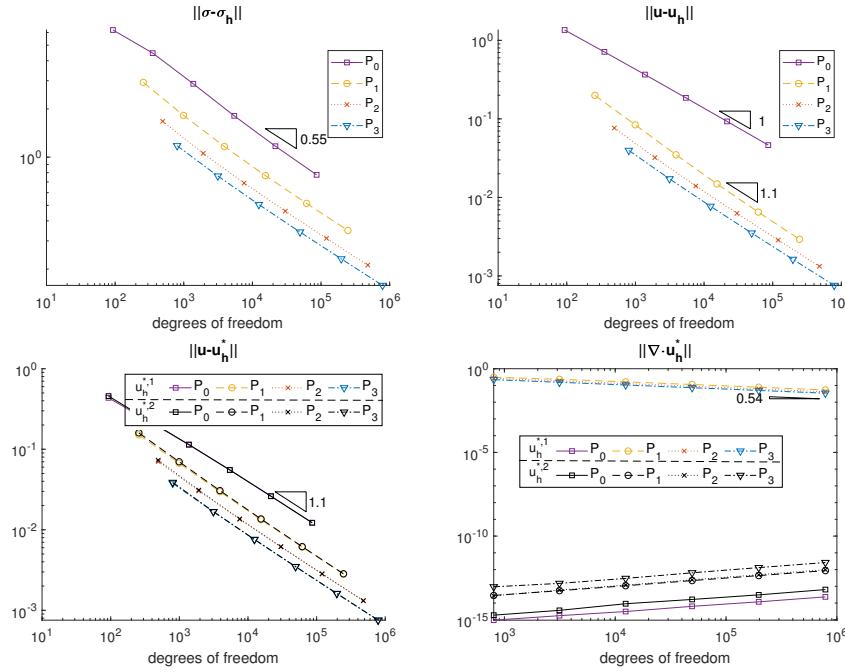


FIG. 3. Convergence history for Example 2 with $k = 0, 1, 2, 3$. The black triangles indicate the convergence rates.

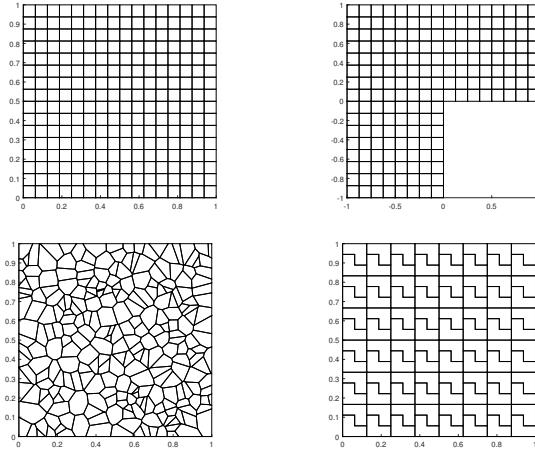


FIG. 4. Uniform rectangular meshes on $(0,1)^2$ (top left) and $(-1,1) \times (-1,1) \setminus [0,1] \times [-1,0]$ (top right). Random Voronoi mesh (bottom left) and L-shaped mesh (bottom right) on $(0,1)^2$.

6.3. Mesh distortion. To demonstrate the robustness of the proposed method to the mesh distortion, we consider two types of meshes: (1) random Voronoi mesh and (2) L-shaped mesh. An example of meshes we used with $h \approx 2^{-4}$ is given in Figure 4. Here, the random Voronoi mesh is obtained by restricting an unbounded Voronoi diagram with random generators from uniform distribution. The proposed method is implemented on the three given meshes with $k = 0, 2$. The convergence history is displayed in Figure 5. All variables converge in optimal order regardless

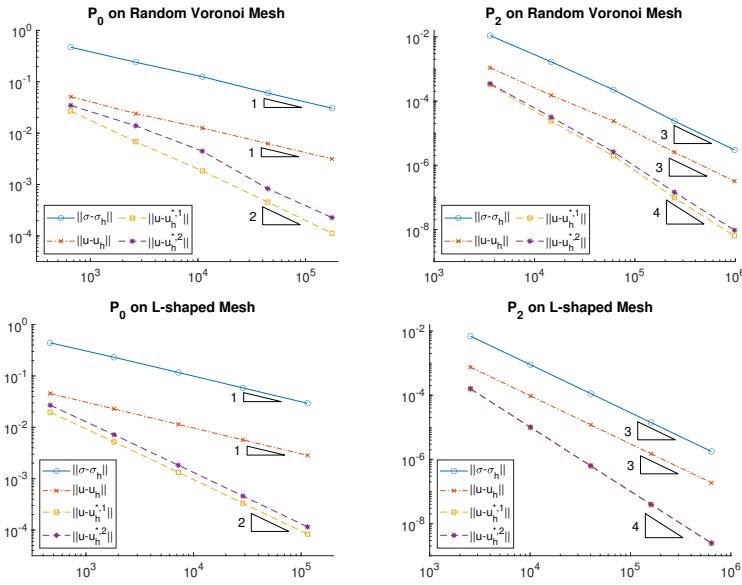


FIG. 5. Convergence history of L^2 -error for Example 1 with $k = 0$ (left) and $k = 2$ (right). Underlying meshes are random Voronoi (top) and L-shaped (bottom) meshes.

of the choice of mesh shape. Also, the difference between the accuracy of numerical solutions on three mesh shapes is not significant. From this result, we infer that our method is reliable even when a mesh is highly distorted or concave.

7. Conclusions. In this paper, we propose staggered discontinuous Galerkin methods for the pseudostress-velocity formulation of the Stokes equations on fairly general polygonal meshes. The hybridization technique is introduced to reduce the global size and ease the construction of the basis functions. The hybrid variable is introduced only on the dual edges. This allows us to consider the local system dual elementwise. A priori analysis is presented to show the optimal convergence for both the pseudostress and velocity. Two postprocessing methods are proposed and both lead to superconvergent velocity fields. The first postprocessing is simple, but the postprocessed velocity field is exact divergence-free only when the piecewise constant approximation is given. The second postprocessing is relatively complex; however, it can yield the exact divergence-free velocity field regardless of the choice of k .

The optimal convergence is verified with a smooth solution on a rectangular domain. The numerical experiments suggest that the optimal convergence can be obtained even when the mesh is highly distorted or concave. Also, the distortion or concavity of the mesh does not have a major influence on the accuracy of the solution. When a solution with a singularity is considered, the convergence rates are limited by the solution regularity as expected (cf. [37]). The postprocessed velocity fields also converge in limited order. However, the divergence-free condition is still satisfied when the second postprocessing is used or the first postprocessing is exploited with $k = 0$.

In the future, we will extend our methods to the stress-velocity formulation of the Stokes equations where the stress tensor is symmetric. The analysis in this study will provide the guidance in getting a numerical method that is optimally convergent and robust to the mesh distortion. Eventually, we will extend our study to the Navier-Stokes equations.

Appendix A. Properties of discrete operators. In this appendix, we investigate properties regarding the finite element spaces and discrete bilinear forms. By the definition of \mathcal{A} , the following lemma is immediate [11].

LEMMA A.1. *For $\psi, \phi \in [L^2(\mathcal{O})]^{2 \times 2}$ with an open set \mathcal{O} , there holds*

$$(\mathcal{A}\psi, \phi)_{\mathcal{O}} = (\psi, \mathcal{A}\phi)_{\mathcal{O}} = (\mathcal{A}\psi, \mathcal{A}\phi)_{\mathcal{O}},$$

and the kernel of \mathcal{A} can be identified by

$$(A.1) \quad \ker(\mathcal{A}) = \{v\mathbf{I} : v \text{ is a scalar function on } \Omega\}.$$

Also, we have the following norm equivalence.

LEMMA A.2. *For all $e \in \mathcal{F}_{pr}$ and $\tilde{\psi} \in \tilde{X}_h^e$, there holds*

$$\sum_{\tau \in \mathcal{D}(e)} \|\tilde{\psi}\|_{L^2(\tau)}^2 \approx \|\tilde{\psi}\|_{X^e}^2 \geq \sum_{\tau \in \mathcal{D}(e)} \|\mathcal{A}\tilde{\psi}\|_{L^2(\tau)}^2.$$

Furthermore, if we have $\tilde{\psi} \in \tilde{X}_h^{e,\perp}$, there holds

$$\|\tilde{\psi}\|_{L^2(\mathcal{D}(e))}^2 \approx \|\tilde{\psi}\|_{X^e}^2 \approx \|\mathcal{A}\tilde{\psi}\|_{L^2(\mathcal{D}(e))}^2,$$

where

$$\tilde{X}_h^{e,\perp} := \{\tilde{\psi} \in \tilde{X}_h^e : b_h^e(\mathbf{v}, \tilde{\psi}) = 0 \forall \mathbf{v} \in V_h^e\}.$$

Here, $A \approx B$ denotes that there exists C_1 and C_2 , independent of h , such that $A \leq C_1 B$ and $B \geq C_2 A$.

Proof. An arithmetic inequality, $a^2 + b^2 \geq (a - b)^2/4$ for any $a, b \in \mathbb{R}$, yields the first inequality. Let $\tilde{\psi} \in \tilde{X}_h^{e,\perp}$ such that $\mathcal{A}\tilde{\psi} = 0$ for given $e \in \mathcal{F}_{pr}$. By (A.1), $\tilde{\psi} = v\mathbf{I}$ for some $v \in \mathbb{P}_k(\tau)$ for each $\tau \in \mathcal{T}_e$. Since $\tilde{\psi} \in \tilde{X}_h^{e,\perp}$, there holds

$$(A.2) \quad \begin{aligned} 0 &= b_h^e(v\mathbf{I}, \mathbf{w}) = - \sum_{\tau \in \mathcal{D}(e)} (v, \nabla \cdot \mathbf{w})_{\tau} + \sum_{e' \in \mathcal{F}_{dl}^e} \langle v, \mathbf{w} \cdot \mathbf{n} \rangle_{e'} \\ &= \sum_{\tau \in \mathcal{D}(e)} (\nabla v, \mathbf{w})_{\tau} - \langle [\![v]\!], \mathbf{w} \cdot \mathbf{n} \rangle_e \end{aligned}$$

for all $\mathbf{w} \in V_h^e$. Since any function $\mathbf{w} \in V_h^e$ is uniquely determined by the degrees of freedom (VD1)–(VD2), we can take \mathbf{w} satisfying

$$\begin{aligned} \mathbf{w} &= \mathbf{0} \quad \text{on } e \in \mathcal{F}_{pr}, \\ (\mathbf{w} - \nabla v, \mathbf{q})_{\tau} &= 0 \quad \forall \mathbf{q} \in [\mathbb{P}_{k-1}(\tau)]^2, \quad \forall \tau \in \mathcal{D}(e). \end{aligned}$$

Noting that $\nabla v|_{\tau} \in \mathbb{P}_{k-1}(\tau)$ for each $\tau \in \mathcal{T}_h$, (A.2) becomes

$$(A.3) \quad \sum_{\tau \in \mathcal{D}(e)} \|\nabla v\|_{L^2(\tau)}^2 = 0.$$

Similarly, taking \mathbf{w} satisfying

$$\begin{aligned} \mathbf{w} &= [\![v]\!] \mathbf{n}, \quad \text{on } e \in \mathcal{F}_{pr}, \\ (\mathbf{w}, \mathbf{q})_{\tau} &= 0 \quad \forall \mathbf{q} \in [\mathbb{P}_{k-1}(\tau)]^2, \quad \forall \tau \in \mathcal{D}(e) \end{aligned}$$

yields

$$(A.4) \quad \|[\![v]\!]\|_{L^2(e)}^2 = 0.$$

Combining (A.3) and (A.4) implies that v is a constant function on $\mathcal{D}(e)$. Also, since $\tilde{\psi} \in \tilde{X}_h^e$, we have

$$2\text{Area}(\mathcal{D}(e))v = \sum_{\tau \in \mathcal{D}(e)} \int_{\tau} 2v \, d\mathbf{x} = \sum_{\tau \in \mathcal{D}(e)} \int_{\tau} \text{tr}(\tilde{\psi}) \, d\mathbf{x} = 0.$$

Therefore, $\tilde{\psi} = v\mathbf{I} \equiv \mathbf{0}$ and this implies that $(\sum_{\tau \in \mathcal{D}(e)} \|\mathcal{A}(\cdot)\|_{L^2(\tau)}^2)^{1/2}$ defines a norm on $\tilde{X}_h^{e,\perp}$. \square

Lemmas A.1 and A.2 imply the following lemma.

LEMMA A.3. *For all $e \in \mathcal{F}_{pr}$ and $\tilde{\psi} \in \tilde{X}_h^{e,\perp}$, there holds*

$$\sum_{\tau \in \mathcal{D}(e)} (\mathcal{A}\tilde{\psi}, \tilde{\psi})_{\tau} \geq C \|\tilde{\psi}\|_{X^e}^2.$$

LEMMA A.4. *For all $e \in \mathcal{F}_{pr}$, there exists a positive constant β independent of h such that*

$$\inf_{\mathbf{v} \in V_h^e} \sup_{\tilde{\psi} \in \tilde{X}_h^e} \frac{b_h^e(\tilde{\psi}, \mathbf{v})}{\|\tilde{\psi}\|_{X^e} \|\mathbf{v}\|_{V^e}} \geq \beta > 0.$$

Proof. First, consider the case with $k = 0$. When $e \subset \partial\Omega$, by the definition of V_h^e , we have $V_h^e = \{\mathbf{0}\}$. Therefore, we only need to consider $e \in \mathcal{F}_{pr}^i$. For given $\mathbf{v} \in V_h^e$, define $\psi \in [\mathbb{P}_k(\mathcal{D}(e))]^{2 \times 2}$ such that

$$\int_{e'} (\psi \mathbf{n}) \cdot \mathbf{q} \, ds = \int_{e'} \frac{1}{h_{e'}} \mathbf{v} \cdot \mathbf{q} \, ds \quad \forall \mathbf{q} \in [\mathbb{P}_0(e')]^2, e' \in \mathcal{F}_{dl}^e.$$

By the scaling argument, we have

$$\|\psi\|_{X^e} \leq C \|\mathbf{v}\|_{V^e}.$$

Now, define $\tilde{\psi} \in \tilde{X}_h^e$ by

$$\tilde{\psi} = \mathcal{A}\psi = \psi - \frac{1}{2} \text{tr}(\psi)I.$$

Then we have

$$(A.5) \quad \|\tilde{\psi}\|_{X^e} = \|\mathcal{A}\psi\|_{X^e} \leq C \|\mathbf{v}\|_{V^e}.$$

Now, observe that

$$\begin{aligned} b_h^e(\tilde{\psi}, \mathbf{v}) &= \sum_{e' \in \mathcal{F}_{dl}^e} \left\langle \tilde{\psi} \mathbf{n}, \mathbf{v} \right\rangle_{e'} \\ &= \sum_{e' \in \mathcal{F}_{dl}^e} \frac{1}{h_{e'}} \left[\langle \mathbf{v}, \mathbf{v} \rangle_{e'} + \langle \mathbf{n}, \mathbf{c} \rangle_{e'} \right], \end{aligned}$$

where $\mathbf{c} = \text{tr}(\psi)\mathbf{v}/2$ is constant on $\partial\mathcal{D}$. Since $\mathcal{D}(e)$ is a closed polygon, the last term is 0. Therefore, we have

$$b_h^e(\tilde{\psi}, \mathbf{v}) = \sum_{e' \in \mathcal{F}_{dl}^e} \frac{1}{h_{e'}} \langle \mathbf{v}, \mathbf{v} \rangle_{e'} = \|\mathbf{v}\|_{V^e}^2.$$

Combining this and (A.5) completes the proof for $k=0$.

Let $k > 0$ and $\mathbf{v} \in V_h^e$ be given. Since $\mathbf{v} \in [\mathbb{P}_k(\tau)]^2$ for all $\tau \in \mathcal{D}(e)$, there exists (cf. Theorem 3.2 in [17]) $\psi \in \mathbb{P}(\tau)$ for all $\tau \in \mathcal{D}(e)$ such that

$$(A.6) \quad \begin{aligned} (\psi, \phi)_\tau &= (\mathcal{A}(-\nabla \mathbf{v}), \phi)_\tau \quad \forall \phi \in [\mathbb{P}_{k-1}(\tau)]^{2 \times 2}, \quad \forall \tau \in \mathcal{D}(e), \\ \langle \psi \mathbf{n}, \mathbf{q} \rangle_{e'} &= \left\langle \frac{\alpha}{h_{e'}} \mathbf{v}, \mathbf{q} \right\rangle_{e'} \quad \forall \mathbf{q} \in [\mathbb{P}_k(e')]^2, \quad \forall e \in \mathcal{F}_{dl}^e, \end{aligned}$$

where α is a constant to be decided later. Then we have from scaling arguments (A.7)

$$\|\psi\|_{X^e}^2 \leq C \left(\sum_{\tau \in \mathcal{D}(e)} \|\mathcal{A}(-\nabla \mathbf{v})\|_{L^2(\tau)}^2 + \sum_{e' \in \mathcal{F}_{dl}^e} \frac{\alpha}{h_{e'}} \|\mathbf{v}\|_{L^2(e')}^2 \right) \leq C(1 + \alpha) \|\mathbf{v}\|_{V^e}^2.$$

By the definition of ψ , we have

$$\sum_{\tau \in \mathcal{D}(e)} (\text{tr}(\psi), 1)_\tau = \sum_{\tau \in \mathcal{D}(e)} (\text{tr}(-\nabla \mathbf{v} + \frac{1}{2} \nabla \cdot \mathbf{v} \mathbf{I}), 1)_\tau = 0,$$

which implies $\psi|_{\mathcal{D}(e)} \in \tilde{X}_h^e$. Then we can obtain

$$(A.8) \quad b_h^e(\psi, \mathbf{v}) = - \sum_{\tau \in \mathcal{D}(e)} (\psi, \nabla \mathbf{v})_\tau + \sum_{e' \in \mathcal{F}_{dl}^e} \langle \psi \mathbf{n}, \mathbf{v} \rangle_{e'}$$

$$(A.9) \quad = \sum_{\tau \in \mathcal{D}(e)} (\nabla \mathbf{v}, \nabla \mathbf{v})_\tau - \sum_{\tau \in \mathcal{D}(e)} \frac{1}{2} (\nabla \cdot \mathbf{v}, \nabla \cdot \mathbf{v})_\tau + \sum_{e' \in \mathcal{F}_{dl}^e} \frac{\alpha}{h_{e'}} \langle \mathbf{v}, \mathbf{v} \rangle_{e'}.$$

Recall the vector Laplacian identity

$$-\Delta \mathbf{v} = \nabla \times (\nabla \times \mathbf{v}) - \nabla(\nabla \cdot \mathbf{v}).$$

Multiplying \mathbf{v} both sides and integration by parts yield

$$(A.10) \quad \begin{aligned} & \sum_{\tau \in \mathcal{D}(e)} (\nabla \mathbf{v}, \nabla \mathbf{v})_\tau - \sum_{e' \in \mathcal{F}_{dl}^e} \langle \partial_n \mathbf{v}, \mathbf{v} \rangle_{e'} \\ &= \sum_{\tau \in \mathcal{D}(e)} (\nabla \times \mathbf{v}, \nabla \times \mathbf{v})_\tau + \sum_{\tau \in \mathcal{D}(e)} (\nabla \cdot \mathbf{v}, \nabla \cdot \mathbf{v})_\tau \\ &+ \sum_{e' \in \mathcal{F}_{dl}^e} \langle \mathbf{v} \times \mathbf{n}, \nabla \times \mathbf{v} \rangle_{e'} - \sum_{e' \in \mathcal{F}_{dl}^e} \langle \mathbf{v} \cdot \mathbf{n}, \nabla \cdot \mathbf{v} \rangle_{e'}. \end{aligned}$$

The trace inequality and Young's inequality yield

$$\begin{aligned} & \sum_{\tau \in \mathcal{D}(e)} (\nabla \mathbf{v}, \nabla \mathbf{v})_{\tau} - \sum_{e' \in \mathcal{F}_{dl}^e} \langle \partial_n \mathbf{v}, \mathbf{v} \rangle_{e'} \\ & \leq \frac{5}{4} \sum_{\tau \in \mathcal{D}(e)} \|\nabla \mathbf{v}\|_{L^2(\tau)}^2 + C_1 \sum_{e' \in \mathcal{F}_{dl}^e} \frac{1}{h_{e'}} \|\mathbf{v}\|_{L^2(e')}^2, \\ & \sum_{\tau \in \mathcal{D}(e)} (\nabla \times \mathbf{v}, \nabla \times \mathbf{v})_{\tau} + \sum_{e' \in \mathcal{F}_{dl}^e} \langle \mathbf{v} \times \mathbf{n}, \nabla \times \mathbf{v} \rangle_{e'} \\ & \geq \frac{3}{4} \sum_{\tau \in \mathcal{D}(e)} \|\nabla \times \mathbf{v}\|_{L^2(\tau)}^2 - C_2 \sum_{e' \in \mathcal{F}_{dl}^e} \frac{1}{h_{e'}} \|\mathbf{v}\|_{L^2(e')}^2, \\ & \sum_{\tau \in \mathcal{D}(e)} (\nabla \cdot \mathbf{v}, \nabla \cdot \mathbf{v})_{\tau} - \sum_{e' \in \mathcal{F}_{dl}^e} \langle \mathbf{v} \cdot \mathbf{n}, \nabla \cdot \mathbf{v} \rangle_{e'} \\ & \geq \frac{3}{4} \sum_{\tau \in \mathcal{D}(e)} \|\nabla \cdot \mathbf{v}\|_{L^2(\tau)}^2 - C_3 \sum_{e' \in \mathcal{F}_{dl}^e} \frac{1}{h_{e'}} \|\mathbf{v}\|_{L^2(e')}^2, \end{aligned}$$

where C_i 's only depend on the mesh regularity and the polynomial degree k but not on h . By combining these inequalities, (A.10) can be rewritten as

$$\begin{aligned} & \frac{5}{4} \sum_{\tau \in \mathcal{D}(e)} \|\nabla \mathbf{v}\|_{L^2(\tau)}^2 + C_4 \sum_{e' \in \mathcal{F}_{dl}^e} \frac{1}{h_{e'}} \|\mathbf{v}\|_{L^2(e')}^2 \\ & \geq \frac{3}{4} \left(\sum_{\tau \in \mathcal{D}(e)} \|\nabla \cdot \mathbf{v}\|_{L^2(\tau)}^2 + \sum_{\tau \in \mathcal{D}(e)} \|\nabla \times \mathbf{v}\|_{L^2(\tau)}^2 \right), \end{aligned}$$

where $C_4 = C_1 + C_2 + C_3$. Taking $\alpha = (4C_4 + 1)/6$ in (A.6), (A.9) reads

$$b_h^e(\psi, \mathbf{v}) \geq \frac{1}{6} \left[\sum_{\tau \in \mathcal{D}(e)} \|\nabla \mathbf{v}\|_{L^2(\tau)}^2 + \sum_{e' \in \mathcal{F}_{dl}^e} \frac{1}{h_{e'}} \|\mathbf{v}\|_{L^2(e')}^2 \right] = \frac{1}{6} \|\mathbf{v}\|_{V^e}^2.$$

This and (A.7) yield the desired result. \square

Acknowledgment. The authors would like to express sincere thanks to the anonymous referees, whose invaluable comments led to an improved version of the paper.

REFERENCES

- [1] D. N. ARNOLD AND F. BREZZI, *Mixed and nonconforming finite element methods: Implementation, postprocessing and error estimates*, ESAIM Math. Model. Numer. Anal., 19 (1985), pp. 7–32.
- [2] D. N. ARNOLD, J. DOUGLAS, JR., AND C. P. GUPTA, *A family of higher order mixed finite element methods for plane elasticity*, Numer. Math., 45 (1984), pp. 1–22.
- [3] D. N. ARNOLD AND R. S. FALK, *A new mixed formulation for elasticity*, Numer. Math., 53 (1988), pp. 13–30.
- [4] L. BEIRÃO DA VEIGA, F. BREZZI, A. CANGIANI, G. MANZINI, L. D. MARINI, AND A. RUSSO, *Basic principles of virtual element methods*, Math. Models Methods Appl. Sci., 23 (2013), pp. 199–214.
- [5] L. BEIRÃO DA VEIGA, F. BREZZI, L. D. MARINI, AND A. RUSSO, *The hitchhiker's guide to the virtual element method*, Math. Models Methods Appl. Sci., 24 (2014), pp. 1541–1573.

- [6] Z. CAI, B. LEE, AND P. WANG, *Least-squares methods for incompressible Newtonian fluid flow: Linear stationary problems*, SIAM J. Numer. Anal., 42 (2004), pp. 843–859.
- [7] Z. CAI AND Y. WANG, *A multigrid method for the pseudostress formulation of Stokes problems*, SIAM J. Sci. Comput., 29 (2007), pp. 2078–2095.
- [8] Z. CAI AND Y. WANG, *Pseudostress-velocity formulation for incompressible Navier-Stokes equations*, Internat. J. Numer. Methods Fluids, 63 (2010), pp. 341–356.
- [9] A. CANGIANI, E. H. GEORGULIS, T. PRYER, AND O. J. SUTTON, *A posteriori error estimates for the virtual element method*, Numer. Math., 137 (2017), pp. 857–893.
- [10] C. CARSTENSEN, D. GALLISTL, AND M. SCHEIDENSACK, *Quasi-optimal adaptive pseudostress approximation of the Stokes equations*, SIAM J. Numer. Anal., 51 (2013), pp. 1715–1734.
- [11] C. CARSTENSEN, D. KIM, AND E.-J. PARK, *A priori and a posteriori pseudostress-velocity mixed finite element error analysis for the Stokes problem*, SIAM J. Numer. Anal., 49 (2011), pp. 2501–2523.
- [12] E. CCERES AND G. N. GATICA, *A mixed virtual element method for the pseudostress-velocity formulation of the Stokes problem*, IMA J. Numer. Anal., 37 (2017), pp. 296–331.
- [13] E. CCERES, G. N. GATICA, AND F. A. SEQUEIRA, *A mixed virtual element method for a pseudostress-based formulation of linear elasticity*, Appl. Numer. Math., 135 (2019), pp. 423–442.
- [14] S. W. CHEUNG, E. CHUNG, H. H. KIM, AND Y. QIAN, *Staggered discontinuous Galerkin methods for the incompressible Navier-Stokes equations*, J. Comput. Phys., 302 (2015), pp. 251–266.
- [15] E. T. CHUNG AND P. CIARLET, JR., *A staggered discontinuous Galerkin method for wave propagation in media with dielectrics and meta-materials*, J. Comput. Appl. Math., 239 (2013), pp. 189–207.
- [16] E. T. CHUNG AND B. ENGQUIST, *Optimal discontinuous Galerkin methods for wave propagation*, SIAM J. Numer. Anal., 44 (2006), pp. 2131–2158.
- [17] E. T. CHUNG AND B. ENGQUIST, *Optimal discontinuous Galerkin methods for the acoustic wave equation in higher dimensions*, SIAM J. Numer. Anal., 47 (2009), pp. 3820–3848.
- [18] E. T. CHUNG, H. H. KIM, AND O. B. WIDLUND, *Two-level overlapping Schwarz algorithms for a staggered discontinuous Galerkin method*, SIAM J. Numer. Anal., 51 (2013), pp. 47–67.
- [19] E. T. CHUNG AND W. QIU, *Analysis of an SDG method for the incompressible Navier-Stokes equations*, SIAM J. Numer. Anal., 55 (2017), pp. 543–569.
- [20] B. COCKBURN, J. GOPALAKRISHNAN, AND R. LAZAROV, *Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems*, SIAM J. Numer. Anal., 47 (2009), pp. 1319–1365.
- [21] B. COCKBURN, J. GOPALAKRISHNAN, N. C. NGUYEN, J. PERAIRE, AND F.-J. SAYAS, *Analysis of HDG methods for Stokes flow*, Math. Comp., 80 (2011), pp. 723–760.
- [22] D. A. DI PIETRO AND J. DRONIQU, *The Hybrid High-Order Method for Polytopal Meshes—Design, Analysis, and Applications*, Model., Simul. and Appl. 19, Springer, New York, 2020.
- [23] D. A. DI PIETRO, A. ERN, AND S. LEMAIRE, *An arbitrary-order and compact-stencil discretization of diffusion on general meshes based on local reconstruction operators*, Comput. Methods Appl. Math., 14 (2014), pp. 461–472.
- [24] J. DU, E. T. CHUNG, M. F. LAM, AND X.-P. WANG, *Discontinuous Galerkin method with staggered hybridization for a class of nonlinear Stokes equations*, J. Sci. Comput., 76 (2018), pp. 1547–1577.
- [25] G. N. GATICA, A. MÁRQUEZ, AND M. A. SÁNCHEZ, *Analysis of a velocity-pressure-pseudostress formulation for the stationary Stokes equation*, Comput. Methods Appl. Mech. Engrg., 199 (2010), pp. 1064–1079.
- [26] M. I. GERRITSMA AND T. N. PHILLIPS, *Compatible spectral approximations for the velocity-stress pressure formulation of the Stokes problem*, SIAM J. Sci. Comput., 20 (1999), pp. 1530–1550.
- [27] Y. JEON AND E.-J. PARK, *A hybrid discontinuous Galerkin method for elliptic problems*, SIAM J. Numer. Anal., 48 (2010), pp. 1968–1983.
- [28] Y. JEON AND E.-J. PARK, *New locally conservative finite element methods on a rectangular mesh*, Numer. Math., 123 (2013), pp. 97–119.
- [29] Y. JEON, E.-J. PARK, AND D. SHEEN, *A hybridized finite element method for the Stokes problem*, Comput. Math. Appl., 68 (2014), pp. 2222–2232.
- [30] C. JOHNSON AND B. MERCIER, *Some equilibrium finite element methods for two-dimensional elasticity problems*, Numer. Math., 30 (1978), pp. 103–116.
- [31] H. H. KIM, E. T. CHUNG, AND C. Y. LAM, *Mortar formulation for a class of staggered discontinuous Galerkin methods*, Comput. Math. Appl., 71 (2016), pp. 1568–1585.

- [32] H. H. KIM, E. T. CHUNG, AND C. S. LEE, *A staggered discontinuous Galerkin method for the Stokes system*, SIAM J. Numer. Anal., 51 (2013), pp. 3327–3350.
- [33] L. MU, J. WANG, G. WEI, X. YE, AND S. ZHAO, *Weak Galerkin methods for second order elliptic interface problems*, J. Comput. Phys., 250 (2013), pp. 106–125.
- [34] N. SUKUMAR AND A. TABARRAEI, *Conforming polygonal finite elements*, Internat. J. Numer. Methods Engrg., 61 (2004), pp. 2045–2066.
- [35] J. WANG AND X. YE, *A weak Galerkin mixed finite element method for second order elliptic problems*, Math. Comp., 241 (2013), pp. 103–115.
- [36] M. WICKE, M. BOTSCHE, AND M. GROSS, *A finite element method on convex polyhedra*, Comput. Graph. Forum, 26 (2007), pp. 355–364.
- [37] L. ZHAO AND E.-J. PARK, *A staggered discontinuous Galerkin method of minimal dimension on quadrilateral and polygonal meshes*, SIAM J. Sci. Comput., 40 (2018), pp. A2543–A2567.
- [38] L. ZHAO AND E.-J. PARK, *A lowest-order staggered DG method for the coupled Stokes-Darcy problem*, IMA J. Numer. Anal., 2020.
- [39] L. ZHAO AND E.-J. PARK, *A new hybrid staggered discontinuous Galerkin method on general meshes*, J. Sci. Comput., 82 (2020), 12 (33 pages).
- [40] L. ZHAO, E.-J. PARK, AND D.-W. SHIN, *A staggered DG method of minimal dimension for the Stokes equations on general meshes*, Comput. Methods Appl. Mech. Engrg., 345 (2019), pp. 854–875.
- [41] O. C. ZIENKIEWICZ, *Displacement and equilibrium models in the finite element method by B. Fraeijs de Veubeke*, Internat. J. Numer. Methods Engrg., 52 (2001), pp. 287–342.