SIAM J. NUMER. ANAL.
Vol. 49, No. 6, pp. 2501–2523

A PRIORI AND A POSTERIORI PSEUDOSTRESS-VELOCITY MIXED FINITE ELEMENT ERROR ANALYSIS FOR THE STOKES PROBLEM*

CARSTEN CARSTENSEN†, DONGHO KIM‡, AND EUN-JAE PARK§

Abstract. The pseudostress-velocity formulation of the stationary Stokes problem allows a Raviart–Thomas mixed finite element formulation with quasi-optimal convergence and some superconvergent reconstruction of the velocity. This local postprocessing gives rise to some averaging a posteriori error estimator with explicit constants for reliable error control. Standard residual-based explicit a posteriori error estimation is shown to be reliable and efficient and motivates adaptive mesh-refining algorithms. Numerical experiments confirm our theoretical findings and illustrate the accuracy of the guaranteed upper error bounds even with reduced regularity.

Key words. mixed finite element approximations, a posteriori error estimates, Stokes problem, pseudostress-velocity formulation

AMS subject classifications. 65K10, 65M12, 65M60

DOI. 10.1137/100816237

1. Introduction. Adaptive mesh-refining plays an important practical role in accurate calculation of the numerical solutions of partial differential equations, especially when the continuous solutions have local singularities or sharp layers. Adaptive mesh-refining algorithms consist of successive loops of SOLVE, ESTIMATE, MARK, and REFINE. A posteriori error estimators provide quantitative estimates for the actual error and motivate local mesh-refinement. Those are computed from the known values such as the given data of the problem and the computed numerical solutions.

Various a posteriori error estimators for finite element methods or mixed finite element methods for second order elliptic problems have already been studied in [2, 4, 12, 22, 23] and for the Stokes problem in [16, 28]. These estimators are of implicit or explicit type and based on the velocity-pressure formulation of the Stokes problem. On the other hand there is a growing interest in a posteriori error estimators which are completely free of unknown constants and lead to guaranteed upper bounds on the numerical error. See, for example, [15, 1, 29] in this direction.

The stress-velocity-pressure formulation [20] is the original physical equations for incompressible Newtonian flows induced by the conservation of momentum and the constitutive law. Arnold and Falk [3] proposed the pseudostress formulation for the equations of linear elasticity which does not require symmetric stress tensors. This allows for an easy discretization via mixed finite elements developed for scalar second order elliptic equations. Cai, Lee, and Wang [9] exploited the pseudostress-velocity formulation to study least-squares methods for the Stokes system.

*Received by the editors November 29, 2010; accepted for publication (in revised form) August 31, 2011; published electronically December 15, 2011. This work was supported by the WCU program through NRF (R31-2008-000-10049-0).
http://www.siam.org/journals/sinum/49-6/81623.html
†Department of Computational Science and Engineering, Yonsei University, Seoul 120-749, Korea and Department of Mathematics, Humboldt-Universität zu Berlin, D-10099 Berlin, Germany (cc@math.hu-berlin.de).
‡University College, Yonsei University, Seoul 120-749, Korea (dhkimm@yonsei.ac.kr).
§Department of Computational Science and Engineering, Yonsei University, Seoul 120-749, Korea (ejpark@yonsei.ac.kr). The research of this author was supported in part by NRF-2010-0013374.

Recently, Cai and Wang [11] used the pseudostress-velocity formulation for the mixed discretization of the Stokes system. The pseudostress is nonsymmetric and the approximation of the pressure, the velocity gradient, or even the stress can be algebraically obtained from the approximate value of the pseudostress. In this paper, we establish a priori and a posteriori error estimates for Raviart–Thomas mixed finite element methods for the pseudostress-velocity formulation of the Stokes problem. We design explicit residual-based reliable and efficient a posteriori error estimators with a possible application to adaptive mesh-refining algorithms. Explicit constants for some averaging estimator make this asymptotically exact.

The motivation for the postprocessing to improve the velocity field is its usage in sharp guaranteed error control; this is the subtle point in the a posteriori error analysis of nonconforming or mixed finite element technologies. The numerical experiments of this paper confirm efficiency indices between one and three. To the best of the authors' knowledge, this is the first result on a posteriori error analysis of the mixed FEMs for the pseudostress-velocity formulation of the Stokes problem.

The remainder of this paper is organized as follows: Section 2 starts with the pseudostress-velocity formulation for the Stokes problem while section 3 introduces its mixed finite element approximation. Section 4 establishes some superconvergent local postprocessing of the velocity for all fixed polynomial degrees. Sections 5 and 6 present an a posteriori error analysis for the stress and the velocity errors. In the final section we present numerical experiments to validate the theoretical results of the previous sections and to explore the accuracy of the guaranteed upper error bounds. The a posteriori stress error estimator requires a finite dimensional approximation for actual computation. Numerical realization is discussed with various adaptive strategies for mesh refinement. Then, numerical examples are given with concluding remarks.

We finish this section with notation and function spaces used in this paper. Let $\mathbb{V}^2$ be the set of two-dimensional vectors and $\mathbb{M}^{2\times 2}$ the set of 2×2 matrices. For $\boldsymbol{v} = (v_1, v_2)^t$, $\boldsymbol{\tau} = (\tau_{ij})_{2\times 2}$, and $\boldsymbol{\sigma} = (\sigma_{ij})_{2\times 2}$, define

$$
\begin{aligned}
\mathbf{curl}\boldsymbol{v} &:= \begin{pmatrix} -\frac{\partial v_1}{\partial y} & \frac{\partial v_1}{\partial x} \\ -\frac{\partial v_2}{\partial y} & \frac{\partial v_2}{\partial x} \end{pmatrix}, &\qquad \nabla \boldsymbol{v} &:= \begin{pmatrix} \frac{\partial v_1}{\partial x} & \frac{\partial v_1}{\partial y} \\ \frac{\partial v_2}{\partial x} & \frac{\partial v_2}{\partial y} \end{pmatrix}, \\
\mathbf{curl}\boldsymbol{\tau} &:= \begin{pmatrix} \frac{\partial \tau_{12}}{\partial x} - \frac{\partial \tau_{11}}{\partial y} \\ \frac{\partial \tau_{22}}{\partial x} - \frac{\partial \tau_{21}}{\partial y} \end{pmatrix}, & \mathbf{div}\boldsymbol{\tau} &:= \begin{pmatrix} \frac{\partial \tau_{11}}{\partial x} + \frac{\partial \tau_{12}}{\partial y} \\ \frac{\partial \tau_{21}}{\partial x} + \frac{\partial \tau_{22}}{\partial y} \end{pmatrix}, \\
\operatorname{tr}\boldsymbol{\tau} &:= \tau_{11} + \tau_{22}, & \boldsymbol{\tau}\boldsymbol{v} &:= \begin{pmatrix} \tau_{11}v_1 + \tau_{12}v_2 \\ \tau_{21}v_1 + \tau_{22}v_2 \end{pmatrix}, \\
\boldsymbol{\tau} : \boldsymbol{\sigma} &:= \sum_{i,j} \tau_{ij}\sigma_{ij}, & \boldsymbol{\delta} &:= 2\times 2 \text{ unit matrix.}
\end{aligned}
$$

Standard Sobolev spaces $H^s(\omega)$ and $H_0^s(\omega)$ for $s \geq 0$ with associated norm $\|\cdot\|_{s,\omega}$ are employed throughout this paper and $(\cdot,\cdot)_\omega$ denotes the $L^2(\omega)$ inner product. In the case $\omega = \Omega$ the lower index is dropped, e.g., $\|\cdot\|_{s,\Omega} = \|\cdot\|_s$ and $(\cdot,\cdot)_\Omega = (\cdot,\cdot)$. We define $H^{-s}(\omega) := (H_0^s(\omega))^*$ as the dual space of $H_0^s(\omega)$. Extending the definitions to vector- and matrix-valued functions, we let $\boldsymbol{H}^s(\omega, \mathbb{V}^2)$ (simply $\boldsymbol{H}^s(\omega)$) and $\boldsymbol{H}^s(\omega, \mathbb{M}^{2\times 2})$ denote the Sobolev spaces over the set of two-dimensional vector- and 2×2 matrix-valued functions, respectively. Finally, we define the space

$$
\boldsymbol{H}(\mathbf{div}, \Omega, \mathbb{M}^{2\times 2}) := \left\{ \boldsymbol{\tau} \in \boldsymbol{L}^2(\Omega, \mathbb{M}^{2\times 2}) \;\middle|\; \mathbf{div}\boldsymbol{\tau} \in \boldsymbol{L}^2(\Omega) \right\}
$$

with the norm $\|\boldsymbol{\tau}\|^2_{\boldsymbol{H}(\mathbf{div},\Omega,\mathbb{M}^{2\times 2})} := (\boldsymbol{\tau},\boldsymbol{\tau}) + (\mathbf{div}\boldsymbol{\tau},\mathbf{div}\boldsymbol{\tau})$. Here, $(\cdot,\cdot)$ denotes the $\boldsymbol{L}^2(\Omega,\mathbb{M}^{2\times 2})$ inner product $\int_\Omega \boldsymbol{\tau} : \boldsymbol{\tau} dx$ and the $\boldsymbol{L}^2(\Omega)$ inner product $\int_\Omega \mathbf{div}\boldsymbol{\tau}\cdot\mathbf{div}\boldsymbol{\tau} dx$ in turn. The extended $L^2(\Gamma)$ product along the boundary Γ is denoted by the duality brackets $\langle\cdot,\cdot\rangle$.

For short notation on generic constants C, for any two real numbers or functions or expressions A and B,

$$A \lesssim B \quad \text{abbreviates} \quad A \leq C\, B.$$

The point is that this multiplicative constant C does not depend on the local or global mesh-sizes but may solely depend on domain Ω. Similarly, $A \approx B$ abbreviates $A \lesssim B \lesssim A$.

2. Pseudostress-velocity formulation. Given a bounded simply connected polygonal domain $\Omega \subset \mathbb{R}^2$ with (connected) Lipschitz boundary Γ filled with a fluid of viscosity $\nu > 0$ and given data $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$ and $\boldsymbol{g} \in \boldsymbol{H}^1(\Omega)$, the stationary Stokes problem for the unknown velocity $\boldsymbol{u}$, and the pressure p reads

$$\begin{aligned} -\nu\triangle\boldsymbol{u} + \nabla p &= -\boldsymbol{f} && \text{in } \Omega,\\ \operatorname{div}\boldsymbol{u} &= 0 && \text{in } \Omega,\\ \boldsymbol{u} &= \boldsymbol{g} && \text{on } \Gamma. \end{aligned} \tag{1}$$

The compatibility conditions read

$$\int_{\partial\Omega} \boldsymbol{g}\cdot\boldsymbol{n}\, ds = 0 \quad \text{and} \quad \int_\Omega p\, dx = 0.$$

Let $\tilde{\boldsymbol{\sigma}} = (\tilde{\sigma}_{ij})_{2\times 2}$ be the stress tensor and let

$$\boldsymbol{\epsilon}(\boldsymbol{u}) := \frac{1}{2}\big(\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^t\big)$$

be the deformation rate tensor. The aforementioned Stokes problem is derived from the stress-velocity-pressure formulation, the original physical equations for incompressible Newtonian flow, i.e.,

$$\begin{aligned} \mathbf{div}\,\tilde{\boldsymbol{\sigma}} &= \boldsymbol{f} && \text{in } \Omega,\\ \tilde{\boldsymbol{\sigma}} + p\boldsymbol{\delta} - 2\nu\boldsymbol{\epsilon}(\boldsymbol{u}) &= \mathbf{0} && \text{in } \Omega,\\ \operatorname{div}\boldsymbol{u} &= 0 && \text{in } \Omega,\\ \boldsymbol{u} &= \boldsymbol{g} && \text{on } \Gamma. \end{aligned} \tag{2}$$

The elimination of the stress in the above system yields the problem (1). To avoid the difficulties caused by the symmetry constraint of the stress tensor we use the (nonsymmetric) pseudostress $\boldsymbol{\sigma} := -p\boldsymbol{\delta} + \nu\nabla\boldsymbol{u}$ of [9]. Direct algebraic calculations recover the velocity gradient, stress, and pressure; the two formulations are equivalent.

For simplicity, we assume that $\nu = 1$ in the Stokes problem (1).

The framework of Cai and Wang [11] enables the design of the pseudostress-velocity formulation as follows. Let $\mathcal{A} : \mathbb{M}^{2\times 2} \to \mathbb{M}^{2\times 2}$ be the deviatoric operator

$$\mathcal{A}\boldsymbol{\tau} := \boldsymbol{\tau} - \frac{1}{2}(\operatorname{tr}\boldsymbol{\tau})\boldsymbol{\delta} \quad \text{for all } \boldsymbol{\tau} \in \mathbb{M}^{2\times 2}.$$

Note that $Ker(\mathcal{A}) = \{q\boldsymbol{\delta} \,|\, q \text{ is a scalar function}\}$ and $\mathcal{A}\boldsymbol{\tau}$ is a trace-free tensor called deviatoric part. The following properties of the operator $\mathcal{A}$ are immediate:

$$
\begin{aligned}
(\mathcal{A}\boldsymbol{\tau}, \boldsymbol{\sigma}) &= (\boldsymbol{\tau}, \mathcal{A}\boldsymbol{\sigma}),\\
(\mathcal{A}\boldsymbol{\tau}, \mathcal{A}\boldsymbol{\tau}) &= (\mathcal{A}\boldsymbol{\tau}, \boldsymbol{\tau}) = (\boldsymbol{\tau}, \boldsymbol{\tau}) - \frac{1}{2}(\mathrm{tr}\,\boldsymbol{\tau}, \mathrm{tr}\,\boldsymbol{\tau}),\\
\|\mathcal{A}\boldsymbol{\tau}\|_0 &\le \|\boldsymbol{\tau}\|_0.
\end{aligned}
$$

The pseudostress

$$\boldsymbol{\sigma} := -p\boldsymbol{\delta} + \nabla \boldsymbol{u}$$

allows for the pseudostress-velocity formulation for the Stokes problem (1),

$$
\begin{aligned}
\mathbf{div}\boldsymbol{\sigma} &= \boldsymbol{f} && \text{in } \Omega,\\
\mathcal{A}\boldsymbol{\sigma} - \nabla \boldsymbol{u} &= \mathbf{0} && \text{in } \Omega,\\
\boldsymbol{u} &= \boldsymbol{g} && \text{on } \Gamma.
\end{aligned}
\tag{3}
$$

The second equation of (3) is obtained from

$$\mathrm{tr}\,(\nabla \boldsymbol{u}) = \mathrm{div}\boldsymbol{u} = 0 \quad \text{and} \quad \mathrm{tr}\,\boldsymbol{\sigma} = -2p.$$

The compatibility condition $\int_\Omega p dx = 0$ implies

$$\int_\Omega \mathrm{tr}\,\boldsymbol{\sigma} dx = 0.$$

We have the following well-known regularity results (see [11, 21]) for sufficiently smooth boundary Γ or for a convex domain. For $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$, the solutions to problems (1) and (3) satisfy $\boldsymbol{u} \in \boldsymbol{H}^2(\Omega) \cap \boldsymbol{H}_0^1(\Omega)$, $p \in H^1(\Omega)/\mathbb{R}$, $\boldsymbol{\sigma} \in \boldsymbol{H}^1(\Omega, \mathbb{M}^{2\times 2})$, and

$$\|\boldsymbol{u}\|_2 + \|p\|_{H^1(\Omega)/\mathbb{R}} + \|\boldsymbol{\sigma}\|_1 \lesssim \|\boldsymbol{f}\|_0 + \|\boldsymbol{g}\|_2. \tag{4}$$

Hereafter, the Stokes problem is said to be H^2-regular if estimate (4) holds; H^{k+2}-regularity is similarly defined via the shift. With $\boldsymbol{V} := \boldsymbol{L}^2(\Omega)$ and

$$\boldsymbol{\Phi} := \boldsymbol{H}(\mathbf{div}, \Omega, \mathbb{M}^{2\times 2})/span\{\boldsymbol{\delta}\} = \left\{ \boldsymbol{\tau} \in \boldsymbol{H}(\mathbf{div}, \Omega, \mathbb{M}^{2\times 2}) \,\middle|\, \int_\Omega \mathrm{tr}\boldsymbol{\tau} dx = 0 \right\},$$

the weak form for the problem (3) reads: *Find* $\boldsymbol{\sigma} \in \boldsymbol{\Phi}$ *and* $\boldsymbol{u} \in \boldsymbol{V}$ *such that*

$$(\mathcal{A}\boldsymbol{\sigma}, \boldsymbol{\tau}) + (\mathbf{div}\boldsymbol{\tau}, \boldsymbol{u}) = <\boldsymbol{g}, \boldsymbol{\tau}\boldsymbol{n}> \quad \text{for all} \quad \boldsymbol{\tau} \in \boldsymbol{\Phi}, \tag{5}$$

$$(\mathbf{div}\boldsymbol{\sigma}, \boldsymbol{v}) = (\boldsymbol{f}, \boldsymbol{v}) \quad \text{for all} \quad \boldsymbol{v} \in \boldsymbol{V}. \tag{6}$$

This problem has a unique solution from the well-known inf-sup condition in the mixed formulation and the following lemma [7, 8].

LEMMA 2.1. *For all* $\boldsymbol{\tau} \in \boldsymbol{\Phi}$*, we have*

$$\|\boldsymbol{\tau}\|_0^2 \lesssim \|\mathcal{A}^{1/2}\boldsymbol{\tau}\|_0^2 + \|\mathbf{div}\boldsymbol{\tau}\|_{-1}^2.$$

3. Mixed finite element method. Let $\{\mathcal{T}_h\}$ be a family of shape-regular triangulations of $\overline{\Omega}$ into triangles T of diameter h_T. For each $\mathcal{T}_h$, denote $\mathcal{E}_h$ to be the set of all edges of $\mathcal{T}_h$. Given $T \in \mathcal{T}_h$, we let $\mathcal{E}(T)$ be the set of its edges. Further, for an edge $E \in \mathcal{E}(T)$, we let $\boldsymbol{t}_E = (-n_2, n_1)^t$ be the unit tangential vector along E for the unit outward normal $\boldsymbol{n}_E = (n_1, n_2)^t$ to E. In what follows, h_E stands for the length of the edge $E \in \mathcal{E}_h$. Moreover, we define the jump $[\boldsymbol{w}]$ of $\boldsymbol{w}$ by

$$[\boldsymbol{w}]\big|_E := (\boldsymbol{w}\big|_{T_+})\big|_E - (\boldsymbol{w}\big|_{T_-})\big|_E \quad \text{if} \quad E = \overline{T}_+ \cap \overline{T}_-,$$

where $\boldsymbol{n}_E$ points from T_+ into its neighbor element T_-. For an edge $E = \overline{T}_+ \cap \Gamma$ on the boundary, the jump reflects boundary conditions with $\boldsymbol{g}$ and hence $[\boldsymbol{w}]\big|_E := \boldsymbol{g} - \boldsymbol{w}$.

We define the finite element spaces associated with the decomposition of Ω,

$$\begin{aligned}
\boldsymbol{\Phi}_T &:= \left\{\boldsymbol{\tau} \in \mathbb{M}^{2\times 2} \,\middle|\, (\tau_{i1}, \tau_{i2}) \in RT_k(T) \text{ for } i = 1, 2\right\},\\
\boldsymbol{\Phi}_h &:= \left\{\boldsymbol{\tau} \in \boldsymbol{\Phi} \,\middle|\, \text{for all } T \in \mathcal{T}_h,\ \boldsymbol{\tau}|_T \in \boldsymbol{\Phi}_T\right\},\\
\boldsymbol{V}_h &:= \left\{\boldsymbol{v} = (v_1, v_2)^t \in \boldsymbol{L}^2(\Omega) \,\middle|\, \text{for all } T \in \mathcal{T}_h,\ v_i|_T \in P_k(T) \text{ for } i = 1, 2\right\},
\end{aligned}$$

where $RT_k(T)$ is the Raviart–Thomas element of index k introduced in [25], and $P_k(T)$ is the set of polynomials of total degree $\le k$ on domain T. We notice that $\boldsymbol{\Phi}_h \subset \boldsymbol{\Phi}$ and hence if $\boldsymbol{\tau}_h \in \boldsymbol{\Phi}_h$, then $\boldsymbol{\tau}_h$ has continuous normal components and the constraint $\int_\Omega \operatorname{tr} \boldsymbol{\tau}_h dx = 0$ holds.

The mixed finite element methods reads: *Find* $\boldsymbol{\sigma}_h \in \boldsymbol{\Phi}_h$ *and* $\boldsymbol{u}_h \in \boldsymbol{V}_h$ *such that*

$$(\mathcal{A}\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) + (\mathbf{div}\boldsymbol{\tau}_h, \boldsymbol{u}_h) = <\boldsymbol{g}, \boldsymbol{\tau}_h \boldsymbol{n}> \quad \text{for all}\, \boldsymbol{\tau}_h \in \boldsymbol{\Phi}_h, \tag{7}$$

$$(\mathbf{div}\boldsymbol{\sigma}_h, \boldsymbol{v}_h) = (\boldsymbol{f}, \boldsymbol{v}_h) \quad \text{for all}\, \boldsymbol{v}_h \in \boldsymbol{V}_h. \tag{8}$$

By Lemma 2.1 and the discrete inf-sup condition of the RT_k element space (cf., [7]), the discrete problem is well posed and has a unique solution.

We consider an interpolation operator over the space $\boldsymbol{\Phi}$. Let $\tilde{\Pi}_h$ denote Raviart–Thomas interpolation operator [7] associated with the degrees of freedom onto $\boldsymbol{\Phi}_h + span\{\boldsymbol{\delta}\}$. We define $\Pi_h : \boldsymbol{\Phi} \to \boldsymbol{\Phi}_h$ by

$$\Pi_h \boldsymbol{\tau} = \tilde{\Pi}_h \boldsymbol{\tau} - \frac{\int_\Omega (\operatorname{tr}\tilde{\Pi}_h \boldsymbol{\tau}) dx}{2|\Omega|} \boldsymbol{\delta} \quad \text{for all}\, \boldsymbol{\tau} \in \boldsymbol{\Phi}, \tag{9}$$

where $|\Omega| = \int_\Omega dx$. We notice that $\int_\Omega (\operatorname{tr}\Pi_h \boldsymbol{\tau}) dx = 0$. Let $\boldsymbol{P}_h$ be the $\boldsymbol{L}^2$ projection onto $\boldsymbol{V}_h$ with the well-known approximation property

$$\|\boldsymbol{P}_h \boldsymbol{v} - \boldsymbol{v}\|_0 \lesssim C h^{k+1} |\boldsymbol{v}|_{k+1} \quad \text{for all}\, \boldsymbol{v} \in \boldsymbol{H}^{k+1}(\Omega). \tag{10}$$

Then the following two lemmas hold: $k = 0$ was treated in [11] and $k \ge 1$ in [10].

LEMMA 3.1. *The commutative property* $\mathbf{div}\Pi_h = \boldsymbol{P}_h \mathbf{div}$ *holds. Furthermore, for* $\boldsymbol{\tau} \in \boldsymbol{\Phi} \cap \boldsymbol{H}^{k+1}(\Omega, \mathbb{M}^{2\times 2})$*, we have*

$$\|\boldsymbol{\tau} - \Pi_h \boldsymbol{\tau}\|_0 \lesssim h^{k+1} |\boldsymbol{\tau}|_{k+1}, \tag{11}$$

$$\|\mathbf{div}\boldsymbol{\tau} - \mathbf{div}(\Pi_h \boldsymbol{\tau})\|_0 \lesssim h^{k+1} \|\mathbf{div}\boldsymbol{\tau}\|_{k+1}. \tag{12}$$

LEMMA 3.2. *For the exact solution* $(\boldsymbol{\sigma}, \boldsymbol{u}) \in (\boldsymbol{\Phi} \cap \boldsymbol{H}^{k+1}(\Omega, \mathbb{M}^{2\times 2})) \times \boldsymbol{H}^{k+1}(\Omega)$ *of problem* (1) *and the approximate solution* $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h)$ *of problem* (7)–(8)*, we have*

$$\begin{aligned}
\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_0 &\lesssim h^{k+1} |\boldsymbol{\sigma}|_{k+1},\\
\|\boldsymbol{u} - \boldsymbol{u}_h\|_0 &\lesssim h^{k+1} (|\boldsymbol{u}|_{k+1} + |\boldsymbol{\sigma}|_{k+1}).
\end{aligned}$$

Remark 3.3. We note that physical quantities such as pressure, gradient velocity, and stress can be expressed as

$$p = -\frac{1}{2}\mathrm{tr}\,\boldsymbol{\sigma}, \quad \nabla \boldsymbol{u} = \mathcal{A}\boldsymbol{\sigma}, \quad \tilde{\boldsymbol{\sigma}} = \boldsymbol{\sigma} + (\nabla \boldsymbol{u})^t = \boldsymbol{\sigma} + (\mathcal{A}\boldsymbol{\sigma})^t = \mathcal{A}\boldsymbol{\sigma} + \boldsymbol{\sigma}^t.$$

From these identities, the approximation of the pressure, gradient velocity, and stress can be defined by

$$p_h := -\frac{1}{2}\mathrm{tr}\boldsymbol{\sigma}_h, \quad (\nabla \boldsymbol{u})_h := \mathcal{A}\boldsymbol{\sigma}_h, \quad \tilde{\boldsymbol{\sigma}}_h := \mathcal{A}\boldsymbol{\sigma}_h + \boldsymbol{\sigma}_h^t.$$

Then the following relations hold:

$$\begin{aligned}
\|p - p_h\|_0 &= \frac{1}{2}\|\mathrm{tr}\boldsymbol{\sigma} - \mathrm{tr}\boldsymbol{\sigma}_h\|_0 \le \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_0,\\
\|\nabla \boldsymbol{u} - (\nabla \boldsymbol{u})_h\|_0 &= \|\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_0 \le \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_0,\\
\|\tilde{\boldsymbol{\sigma}} - \tilde{\boldsymbol{\sigma}}_h\|_0 &\le \|\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_0 + \|(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)^t\|_0 \le 2\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_0.
\end{aligned}$$

The estimate for $\|\boldsymbol{P}_h\boldsymbol{u} - \boldsymbol{u}_h\|_0$ in the following theorem allows for the error estimates of the postprocessed velocity below.

THEOREM 3.4. *Any* $\boldsymbol{\sigma} \in \boldsymbol{H}^{k+1}(\Omega, \mathbb{M}^{2\times 2})$ *and* $\boldsymbol{f} = \mathbf{div}\boldsymbol{\sigma} \in \boldsymbol{H}^{k+1}(\Omega)$ *satisfy*

$$\|\boldsymbol{P}_h\boldsymbol{u} - \boldsymbol{u}_h\|_0 \lesssim h^{k+2}\left(|\boldsymbol{\sigma}|_{k+1} + |\mathbf{div}\boldsymbol{\sigma}|_{k+1}\right). \tag{13}$$

Proof. We start with a duality argument. Let $(\boldsymbol{\eta}, \boldsymbol{z}) \in \boldsymbol{\Phi} \times \boldsymbol{V}$ be the solution to

$$(\mathcal{A}\boldsymbol{\eta}, \boldsymbol{\tau}) + (\mathbf{div}\boldsymbol{\tau}, \boldsymbol{z}) = 0 \quad \text{for all}\, \boldsymbol{\tau} \in \boldsymbol{\Phi}, \tag{14}$$

$$(\mathbf{div}\boldsymbol{\eta}, \boldsymbol{v}) = (\boldsymbol{P}_h\boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{v}) \quad \text{for all}\, \boldsymbol{v} \in \boldsymbol{V}. \tag{15}$$

The convexity of Ω and a priori estimates (4) imply

$$\|\boldsymbol{z}\|_2 \lesssim \|\boldsymbol{P}_h\boldsymbol{u} - \boldsymbol{u}_h\|_0 \quad \text{and} \quad \|\boldsymbol{\eta}\|_1 \lesssim \|\boldsymbol{P}_h\boldsymbol{u} - \boldsymbol{u}_h\|_0. \tag{16}$$

Since (15) and $\mathbf{div}\Pi_h = \boldsymbol{P}_h\mathbf{div}$, we deduce

$$\begin{aligned}
\|\boldsymbol{P}_h\boldsymbol{u} - \boldsymbol{u}_h\|_0^2 &= (\boldsymbol{P}_h\boldsymbol{u} - \boldsymbol{u}_h, \mathbf{div}\boldsymbol{\eta})\\
&= (\boldsymbol{P}_h\boldsymbol{u} - \boldsymbol{u}_h, \mathbf{div}\Pi_h\boldsymbol{\eta})\\
&= (\boldsymbol{u} - \boldsymbol{u}_h, \mathbf{div}\Pi_h\boldsymbol{\eta}).
\end{aligned}$$

Subtracting (7)–(8) from (5)–(6) we obtain the error identities

$$(\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h), \boldsymbol{\tau}_h) + (\mathbf{div}\boldsymbol{\tau}_h, \boldsymbol{u} - \boldsymbol{u}_h) = 0 \quad \text{for all}\, \boldsymbol{\tau}_h \in \boldsymbol{\Phi}_h, \tag{17}$$

$$(\mathbf{div}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h), \boldsymbol{v}_h) = 0 \quad \text{for all}\, \boldsymbol{v}_h \in \boldsymbol{V}_h. \tag{18}$$

The identities (17), (14), and (18) and the estimates (10)–(11) yield

$$\begin{aligned}
\|\boldsymbol{P}_h\boldsymbol{u} - \boldsymbol{u}_h\|_0^2 &= -(\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h), \Pi_h\boldsymbol{\eta} - \boldsymbol{\eta}) - (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathcal{A}\boldsymbol{\eta})\\
&= -(\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h), \Pi_h\boldsymbol{\eta} - \boldsymbol{\eta}) + (\mathbf{div}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h), \boldsymbol{z} - \boldsymbol{P}_h\boldsymbol{z})\\
&\lesssim h(\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_0\|\boldsymbol{\eta}\|_1 + \|\mathbf{div}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_0\|\boldsymbol{z}\|_1).
\end{aligned} \tag{19}$$

To analyze the term $\|\mathbf{div}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_0$, consider $\mathbf{div}\Pi_h = \boldsymbol{P}_h\mathbf{div}$ and (18). For any $\boldsymbol{w}_h \in \boldsymbol{V}_h$, this leads to

$$(\mathbf{div}(\Pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}_h), \boldsymbol{w}_h) = (\mathbf{div}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h), \boldsymbol{w}_h) = 0.$$

Let $\boldsymbol{w}_h = \mathbf{div}(\Pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)$ and deduce $\mathbf{div}(\Pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) = 0$. Hence

$$\|\mathbf{div}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_0 = \|\mathbf{div}(\boldsymbol{\sigma} - \Pi_h\boldsymbol{\sigma})\|_0 = \|\mathbf{div}\boldsymbol{\sigma} - \boldsymbol{P}_h\mathbf{div}\boldsymbol{\sigma}\|_0 \lesssim h^{k+1}|\mathbf{div}\boldsymbol{\sigma}|_{k+1}. \tag{20}$$

Lemma 3.2 and the inequalities (16), (19)–(20) lead to

$$\|\boldsymbol{P}_h\boldsymbol{u} - \boldsymbol{u}_h\|_0 \lesssim h^{k+2}(|\boldsymbol{\sigma}|_{k+1} + |\mathbf{div}\boldsymbol{\sigma}|_{k+1}). \qquad \square$$

4. Postprocessing. From Remark 3.3, we note that $\mathcal{A}\boldsymbol{\sigma}_h = \boldsymbol{\sigma}_h + p_h\boldsymbol{\delta}$ is a good approximation of $\nabla\boldsymbol{u}$. This implies an improved approximate solution of the velocity $\boldsymbol{u}$ through local postprocessing in the spirit of Stenberg [26]. Let

$$\boldsymbol{W}_h^* = \{\boldsymbol{v} \in \boldsymbol{L}^2(\Omega) \;\mid\; \boldsymbol{v}|_T \in P_{k+1}(T)^2 \text{ for all } T \in \mathcal{T}_h\}.$$

We define $\boldsymbol{u}_h^* \in \boldsymbol{W}_h^*$ on each T with $\boldsymbol{P}_T = \boldsymbol{P}_h|_T$ as the solution to the system

$$\boldsymbol{P}_T\boldsymbol{u}_h^* = \boldsymbol{u}_h, \tag{21}$$

$$(\nabla\boldsymbol{u}_h^*, \nabla\boldsymbol{v}_h)_T = (\boldsymbol{\sigma}_h, \nabla\boldsymbol{v}_h)_T + (p_h, \mathrm{div}\boldsymbol{v}_h)_T \quad \text{for all } \boldsymbol{v}_h \in (I - \boldsymbol{P}_T)\boldsymbol{W}_h^*|_T. \tag{22}$$

Note that $\boldsymbol{u}_h^*|_T$ is the Riesz representation of the linear functional $(\boldsymbol{\sigma}_h, \nabla(\cdot))_T + (p_h, \mathrm{div}(\cdot))_T$ in the Hilbert space $(I - \boldsymbol{P}_T)\boldsymbol{W}_h^*|_T \equiv (P_{k+1}(T)/P_k(T))^2$ with scalar product $(\nabla(\cdot), \nabla(\cdot))_T$. The Poincaré inequality yields positive definiteness on each triangle and the well-posedness of (21)–(22) follows. Observe from (22) in case $k = 0$ that $\mathrm{div}\,\boldsymbol{u}_h^* = 0$ on each T. A generalization of this property $\mathrm{div}_h\boldsymbol{u}_h^* = 0$ to $k \geq 1$ appears nonobvious.

From the identity $\boldsymbol{\sigma} = -p\boldsymbol{\delta} + \nabla\boldsymbol{u}$ and (22), we obtain the error identity

$$(\nabla(\boldsymbol{u} - \boldsymbol{u}_h^*), \nabla\boldsymbol{v})_T = (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \nabla\boldsymbol{v})_T + (p - p_h, \mathrm{div}\boldsymbol{v})_T \,\forall\, \boldsymbol{v} \in (I - \boldsymbol{P}_T)\boldsymbol{W}_h^*|_T. \tag{23}$$

THEOREM 4.1. *Let* $\boldsymbol{u} \in \boldsymbol{H}^{k+2}(\Omega)$, $\boldsymbol{\sigma} \in \boldsymbol{H}^{k+1}(\Omega, \mathbb{M}^{2\times2})$, *and* $\boldsymbol{f} = \mathbf{div}\boldsymbol{\sigma} \in \boldsymbol{H}^{k+1}(\Omega)$ *solve* (1) *and let* $\boldsymbol{u}_h^* \in \boldsymbol{W}_h^*$ *satisfy* (21)–(22). *Then we have*

$$\|\boldsymbol{u} - \boldsymbol{u}_h^*\|_0 \lesssim h^{k+2}(|\boldsymbol{u}|_{k+2} + |\boldsymbol{\sigma}|_{k+1} + |\mathrm{div}\boldsymbol{\sigma}|_{k+1}).$$

Proof. Let $\hat{\boldsymbol{u}}$ be the L^2-projection of $\boldsymbol{u}$ onto $\boldsymbol{W}_h^*$. Define $\boldsymbol{v} \in \boldsymbol{W}_h^*$ by $\boldsymbol{v}|_T = (\boldsymbol{\delta} - \boldsymbol{P}_T)(\hat{\boldsymbol{u}} - \boldsymbol{u}_h^*)$ for each T. Then we have from (23) and Schwarz inequality

$$\begin{aligned}
|\boldsymbol{v}|_{1,T}^2 &= (\nabla(\hat{\boldsymbol{u}} - \boldsymbol{u}_h^*), \nabla\boldsymbol{v})_T - (\nabla\boldsymbol{P}_T(\hat{\boldsymbol{u}} - \boldsymbol{u}_h^*), \nabla\boldsymbol{v})_T \\
&= (\nabla(\hat{\boldsymbol{u}} - \boldsymbol{u}), \nabla\boldsymbol{v})_T + (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \nabla\boldsymbol{v})_T + (p - p_h, \mathrm{div}\boldsymbol{v})_T \\
&\quad - (\nabla\boldsymbol{P}_T(\hat{\boldsymbol{u}} - \boldsymbol{u}_h^*), \nabla\boldsymbol{v})_T \\
&\leq |\hat{\boldsymbol{u}} - \boldsymbol{u}|_{1,T}|\boldsymbol{v}|_{1,T} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,T}|\boldsymbol{v}|_{1,T} + \|p - p_h\|_{0,T}\|\mathrm{div}\boldsymbol{v}\|_{0,T} \\
&\quad + |\boldsymbol{P}_T(\hat{\boldsymbol{u}} - \boldsymbol{u}_h^*)|_{1,T}|\boldsymbol{v}|_{1,T}.
\end{aligned} \tag{24}$$

Since $\|\mathrm{div}\boldsymbol{v}\|_{0,T}^2 \leq 2\|\nabla\boldsymbol{v}\|_{0,T}^2$, we obtain

$$|\boldsymbol{v}|_{1,T} \leq |\hat{\boldsymbol{u}} - \boldsymbol{u}|_{1,T} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,T} + \sqrt{2}\|p - p_h\|_{0,T} + |\boldsymbol{P}_T(\hat{\boldsymbol{u}} - \boldsymbol{u}_h^*)|_{1,T}.$$

Since $(\boldsymbol{\delta} - \boldsymbol{P}_T)\boldsymbol{w} = 0$ if $\boldsymbol{w} \in P_0(T)^2$, the Poincaré inequality (with Payne–Weinberger constant $1/\pi$) reads

$$\|\boldsymbol{v}\|_{0,T} \leq h_T/\pi\,|\boldsymbol{v}|_{1,T}.$$

This inequality and the inverse estimate

$$|\boldsymbol{P}_T(\hat{\boldsymbol{u}} - \boldsymbol{u}_h^*)|_{1,T} \lesssim h_T^{-1} \|\boldsymbol{P}_T(\hat{\boldsymbol{u}} - \boldsymbol{u}_h^*)\|_{0,T}$$

yield

$$\begin{aligned} \|(\boldsymbol{\delta} - \boldsymbol{P}_T)(\hat{\boldsymbol{u}} - \boldsymbol{u}_h^*)\|_{0,T} &\lesssim h_T |\boldsymbol{v}|_{1,T} \lesssim \|\boldsymbol{P}_T(\hat{\boldsymbol{u}} - \boldsymbol{u}_h^*)\|_{0,T} \\ &\quad + h_T \left(|\hat{\boldsymbol{u}} - \boldsymbol{u}|_{1,T} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,T} + \|p - p_h\|_{0,T}\right). \end{aligned} \tag{25}$$

After squaring and summing over all $T \in \mathcal{T}_h$, we conclude the proof from Theorem 3.4 and the following estimates:

$$\begin{aligned} &\|\boldsymbol{P}_h(\hat{\boldsymbol{u}} - \boldsymbol{u}_h^*)\|_0 = \|\boldsymbol{P}_h \boldsymbol{u} - \boldsymbol{P}_h \boldsymbol{u}_h^*\|_0 = \|\boldsymbol{P}_h \boldsymbol{u} - \boldsymbol{u}_h\|_0 \lesssim h^{k+2} \left(|\boldsymbol{\sigma}|_{k+1} + |\mathrm{div}\boldsymbol{\sigma}|_{k+1}\right), \\ &\qquad \|p - p_h\|_0 \le \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_0 \lesssim h^{k+1} |\boldsymbol{\sigma}|_{k+1}, \\ &\qquad \|\boldsymbol{u} - \boldsymbol{u}_h^*\|_0 \le \|\boldsymbol{u} - \hat{\boldsymbol{u}}\|_0 + \|(\boldsymbol{\delta} - \boldsymbol{P}_h)(\hat{\boldsymbol{u}} - \boldsymbol{u}_h^*)\|_0 + \|\boldsymbol{P}_h(\hat{\boldsymbol{u}} - \boldsymbol{u}_h^*)\|_0. \qquad \square \end{aligned}$$

5. A posteriori stress error estimation. In its first part, this section follows the unified approach in [13] to obtain reliable and efficient error estimators. The second part analyzes the constants explicitly and leads to asymptotic exactness in Theorem 5.3.

Recall that $\boldsymbol{\varepsilon} := \boldsymbol{\sigma} - \boldsymbol{\sigma}_h$ and $\boldsymbol{e} := \boldsymbol{u} - \boldsymbol{u}_h$ for the unique approximate solution $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h) \in \boldsymbol{\Phi}_h \times \boldsymbol{V}_h$ of the mixed finite element methods (7)–(8). The well posedness of the Stokes system leads to equivalence of errors and residuals. The generic constants, hidden in the equivalence $\approx$ below, represent the norms of some operators on the continuous level (cf. section 3 of [13] for details and proofs) and are independent of $\boldsymbol{\varepsilon}, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{\sigma}_h, \boldsymbol{f}$, and $\boldsymbol{f}_h := \boldsymbol{P}_h \boldsymbol{f}$, etc.

THEOREM 5.1. *Given the exact solution $\boldsymbol{\sigma} \in \boldsymbol{\Phi}$ and $\boldsymbol{u} \in \boldsymbol{g} + \boldsymbol{H}_0^1(\Omega)$ from* (5)–(6) *and the discrete solution $\boldsymbol{\sigma}_h \in \boldsymbol{\Phi}_h$ and $\boldsymbol{u}_h \in \boldsymbol{V}_h$ from* (7)–(8)*, any $\boldsymbol{v} \in \boldsymbol{g} + \boldsymbol{H}_0^1(\Omega)$ satisfies*

$$\|\boldsymbol{\varepsilon}\|_0 + \|\boldsymbol{u} - \boldsymbol{v}\|_1 \approx \|\mathcal{A}\,\boldsymbol{\sigma}_h - \nabla \boldsymbol{v}\|_0 + \|\boldsymbol{f} - \boldsymbol{f}_h\|_{-1}.$$

Proof. The ideas of the proof are contained in [13] and recalled here for the particular situation at hand for convenient reading. The bilinear form from (5)–(6) will be recast into the primal form with the bilinear form

$$\mathcal{B} : \left(\boldsymbol{L}^2(\Omega; \mathbb{M}^{2\times 2})/span\{\boldsymbol{\delta}\} \times \boldsymbol{H}_0^1(\Omega)\right) \times \left(\boldsymbol{L}^2(\Omega; \mathbb{M}^{2\times 2})/span\{\boldsymbol{\delta}\} \times \boldsymbol{H}_0^1(\Omega)\right) \to \mathbb{R}$$

defined for any $(\boldsymbol{\sigma}, \boldsymbol{u}), (\boldsymbol{\tau}, \boldsymbol{v}) \in \boldsymbol{L}^2(\Omega; \mathbb{M}^{2\times 2})/span\{\boldsymbol{\delta}\} \times \boldsymbol{H}_0^1(\Omega)$ by

$$\mathcal{B}\left((\boldsymbol{\sigma}, \boldsymbol{u}), (\boldsymbol{\tau}, \boldsymbol{v})\right) := (\mathcal{A}\boldsymbol{\sigma}, \boldsymbol{\tau}) - (\boldsymbol{\tau}, \nabla \boldsymbol{u}) - (\boldsymbol{\sigma}, \nabla \boldsymbol{v}).$$

This bilinear form is bounded and symmetric and defines an isomorphism. The Brezzi splitting theorem for the proof of the global inf-sup condition [5, 6, 7] requires the following inf-sup condition on the bilinear form:

$$b : \boldsymbol{L}^2(\Omega; \mathbb{M}^{2\times 2})/span\{\boldsymbol{\delta}\} \times \boldsymbol{H}_0^1(\Omega) \to \mathbb{R},\ (\boldsymbol{\tau}, \boldsymbol{v}) \mapsto (\boldsymbol{\tau}, \nabla \boldsymbol{v}).$$

Indeed, any $\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega)$ with gradient $\boldsymbol{\tau} := \nabla \boldsymbol{v} \in \boldsymbol{L}^2(\Omega; \mathbb{M}^{2\times 2})/span\{\boldsymbol{\delta}\}$ satisfies

$$b(\boldsymbol{\tau}, \boldsymbol{v}) = b(\nabla \boldsymbol{v}, \boldsymbol{v}) = \|\nabla \boldsymbol{v}\|_0^2 \approx \|\boldsymbol{v}\|_1^2 \gtrsim \|\boldsymbol{\tau}\|_0 \|\boldsymbol{v}\|_1.$$

The second condition in the Brezzi splitting theorem is the ellipticity of the bilinear form

$$a : \boldsymbol{L}^2(\Omega;\mathbb{M}^{2\times 2}) \times \boldsymbol{L}^2(\Omega;\mathbb{M}^{2\times 2}) \to \mathbb{R},\ (\boldsymbol{\sigma},\boldsymbol{\tau}) \mapsto (\mathcal{A}\boldsymbol{\sigma},\boldsymbol{\tau})$$

on the subspace

$$Z := \{\boldsymbol{\tau} \in \boldsymbol{L}^2(\Omega;\mathbb{M}^{2\times 2})/span\{\boldsymbol{\delta}\} | \text{ for all } \boldsymbol{v} \in \boldsymbol{H}^1_0(\Omega),\ (\boldsymbol{\tau},\nabla\boldsymbol{v}) = 0\}.$$

An integration by parts shows that $\boldsymbol{\tau} \in Z$ is divergence free, and hence Lemma 2.1 leads to

$$\|\boldsymbol{\tau}\|_0^2 \lesssim \|\mathcal{A}^{1/2}\boldsymbol{\tau}\|_0^2 = a(\boldsymbol{\tau},\boldsymbol{\tau}).$$

This completes the proof of the global inf-sup condition on the bilinear form $\mathcal{B}$. Consequently, given the exact solution $\boldsymbol{\sigma} \in \boldsymbol{\Phi}$ and $\boldsymbol{u} \in \boldsymbol{H}^1_0(\Omega)$ and the discrete solution $\boldsymbol{\sigma}_h \in \boldsymbol{\Phi}_h$ and $\boldsymbol{u}_h \in \boldsymbol{V}_h$ plus any $\boldsymbol{v} \in \boldsymbol{H}^1_0(\Omega)$, the norm of $(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{v})$ in $\boldsymbol{L}^2(\Omega;\mathbb{M}^{2\times 2}) \times \boldsymbol{H}^1_0(\Omega)$ is equivalent to the dual norm of $\mathcal{B}((\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{v}),\cdot)$ in the dual space of $\boldsymbol{L}^2(\Omega;\mathbb{M}^{2\times 2}) \times \boldsymbol{H}^1_0(\Omega)$. Since $\mathcal{A}\boldsymbol{\sigma} = \nabla\boldsymbol{u}$, it follows for all $\boldsymbol{\tau} \in \boldsymbol{L}^2(\Omega;\mathbb{M}^{2\times 2})/span\{\boldsymbol{\delta}\}$ and $\boldsymbol{w} \in \boldsymbol{H}^1_0(\Omega)$ that

$$\mathcal{B}((\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{v}),(\boldsymbol{\tau},\boldsymbol{w})) = (\nabla\boldsymbol{v}-\mathcal{A}\boldsymbol{\sigma}_h,\boldsymbol{\tau}) + (\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\nabla\boldsymbol{w}).$$

Since $\mathbf{div}\boldsymbol{\sigma} = \boldsymbol{f}$ and $\mathbf{div}\boldsymbol{\sigma}_h = \boldsymbol{f}_h = \boldsymbol{P}_h\boldsymbol{f}$, an integration by parts leads to

$$(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\nabla\boldsymbol{w}) = -(\mathbf{div}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),\boldsymbol{w}) = -(\boldsymbol{f}-\boldsymbol{f}_h,\boldsymbol{w}).$$

Hence, the square of the dual norm of $\mathcal{B}((\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{v}),\cdot)$ equals

$$\|\mathcal{A}\,\boldsymbol{\sigma}_h - \nabla\boldsymbol{v}\|_0^2 + \|\boldsymbol{f}-\boldsymbol{f}_h\|_{-1}^2.$$

This is equivalent to the error norm $\|\boldsymbol{\varepsilon}\|_0^2 + \|\boldsymbol{u}-\boldsymbol{v}\|_1^2$ and concludes the proof. ∎

The optimal choice of $\boldsymbol{v}$ on the right-hand side leads to the definition of the nonconformity error estimator

$$\mu_h := \min_{\boldsymbol{v}\in\boldsymbol{g}+\boldsymbol{H}^1_0(\Omega)} \|\mathcal{A}\,\boldsymbol{\sigma}_h - \nabla\boldsymbol{v}\|_0.$$

Then, the theorem yields

$$\|\boldsymbol{\varepsilon}\|_0 \lesssim \mu_h + \|\boldsymbol{f}-\boldsymbol{f}_h\|_{-1}. \tag{26}$$

The exact computation of the optimal $\boldsymbol{v} \in \boldsymbol{V}$ in μ_h is certainly too costly, but any approximation of it will lead to a guaranteed upper error bound. Notice that a Poincaré inequality (with the factor $1/\pi$ for convex domains after Payne–Weinberger [24])

$$\|\boldsymbol{f}-\boldsymbol{f}_h\|_{-1} := \sup_{\boldsymbol{v}\in\boldsymbol{H}^1_0(\Omega)\setminus\{0\}} \int (\boldsymbol{f}-\boldsymbol{f}_h)\cdot\boldsymbol{v}\,dx/\|\nabla\boldsymbol{v}\|_0 \le \operatorname{osc}_k(\boldsymbol{f},\mathcal{T}_h)/\pi$$

for the higher-order data oscillation (of order $k+2$ for piecewise H^{k+1} data $\boldsymbol{f}$) leads to

$$\operatorname{osc}_k(\boldsymbol{f},\mathcal{T}_h)^2 := \sum_{T\in\mathcal{T}_h} h_T^2\|\boldsymbol{f}-\boldsymbol{f}_h\|_{0,T}^2.$$

The choice $\boldsymbol{v} = \boldsymbol{u}$ and the relation $\nabla \boldsymbol{u} = \mathcal{A}\boldsymbol{\sigma}$ imply that

$$\mu_h \leq \|\mathcal{A}(\boldsymbol{\sigma}_h - \boldsymbol{\sigma})\|_0 \leq \|\boldsymbol{\varepsilon}\|_0.$$

Therefore, in view of (26), the estimator μ_h is reliable and efficient in the sense of

$$\mu_h \leq \|\boldsymbol{\varepsilon}\|_0 \lesssim \mu_h + \|\boldsymbol{f} - \boldsymbol{f}_h\|_{-1}.$$

In the discussion so far, the generic constants are supressed in the notation behind the signs $\approx$ or $\lesssim$. To exploit those constants in the second part of this section, consider the spaces

$$X := \{\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega) : \operatorname{div} \boldsymbol{v} = 0\}$$

and

$$Y := \left\{\boldsymbol{\tau} \in \boldsymbol{L}^2(\Omega; \mathbb{M}^{2\times 2}) : \text{ for all } \boldsymbol{v} \in X, \int_\Omega \boldsymbol{\tau} : \nabla \boldsymbol{v}\, dx = 0\right\}.$$

The orthogonal split

$$\boldsymbol{L}^2(\Omega; \mathbb{M}^{2\times 2}) = \nabla X \oplus Y$$

is known from another reformulation of the Stokes equations [19].

LEMMA 5.2 (see [19, Lemma 2]). *The positive constant*

$$c_0 := \inf_{q \in L_0^2(\Omega)\setminus\{0\}} \sup_{\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega)\setminus\{0\}} \frac{\int_\Omega q \operatorname{div} \boldsymbol{v}\, dx}{\|\nabla \boldsymbol{v}\|_0 \, \|q\|_0}$$

satisfies

$$\textit{for all } \boldsymbol{\tau} \in Y\, \exists\, \omega \in L_0^2(\Omega) \quad \nabla\omega = \operatorname{div} \boldsymbol{\tau} \textit{ and } c_0\, \|\omega\|_0 \leq \|\boldsymbol{\tau}\|_0.$$

Numerical values for the inf-sup constant c_0 can be found in the literature [17, 27]. Recall that $\boldsymbol{\varepsilon} := \boldsymbol{\sigma} - \boldsymbol{\sigma}_h$ and that $\mathcal{A}\boldsymbol{\varepsilon} = \nabla \boldsymbol{u} - \mathcal{A}\boldsymbol{\sigma}_h$ is its deviatoric part. The second main result of this section is the following estimate which leads to asymptotic exactness of the estimator:

$$\eta_h := \min_{\boldsymbol{v} \in \boldsymbol{g} + \boldsymbol{H}_0^1(\Omega)} (\|\mathcal{A}(\boldsymbol{\sigma}_h - \nabla \boldsymbol{v})\| + c_0^{-1}\|\operatorname{div} \boldsymbol{v}\|). \tag{27}$$

Indeed, the following estimate implies asymptotic exactness in the sense of:

$$\big|\|\mathcal{A}\boldsymbol{\varepsilon}\|_0^2 - \eta_h^2\big| \leq \|\boldsymbol{f} - \boldsymbol{f}_h\|_{-1}^2 \leq \operatorname{osc}_k(\boldsymbol{f}, \mathcal{T}_h)^2/\pi^2,$$

which is of high order for any sufficiently smooth right-hand side $\boldsymbol{f}$.

THEOREM 5.3. *For any $\boldsymbol{v} \in \boldsymbol{g} + \boldsymbol{H}_0^1(\Omega)$, it holds*

$$\eta_h^2 \leq \|\mathcal{A}\boldsymbol{\varepsilon}\|_0^2 \leq \eta_h^2 + \|\boldsymbol{f} - \boldsymbol{f}_h\|_{-1}^2.$$

Proof. Given $\mathcal{A}\boldsymbol{\varepsilon} \in L^2(\Omega; \mathbb{M}^{2\times 2})$, the linear functional

$$(\mathcal{A}\boldsymbol{\varepsilon}, \nabla \cdot) : X \to \mathbb{R}, \boldsymbol{v} \mapsto (\mathcal{A}\boldsymbol{\varepsilon}, \nabla \boldsymbol{v})$$

with the scalar product $(\cdot,\cdot)$ in $\boldsymbol{L}^2(\Omega;\mathbb{M}^{2\times 2})$ allows a unique Riesz representation $\boldsymbol{a}\in X$ in the Hilbert space X with respect to the scalar product $(\nabla\cdot,\nabla\cdot)$ (which is equivalent to the scalar product in $\boldsymbol{H}_0^1(\Omega)$). In other words, $\boldsymbol{a}\in X$ satisfies

$$(\nabla\boldsymbol{a},\nabla\boldsymbol{v})=(\mathcal{A}\,\boldsymbol{\varepsilon},\nabla\boldsymbol{v})\quad\text{for all }\boldsymbol{v}\in X.$$

The remainder $\mathcal{A}\,\boldsymbol{\varepsilon}-\nabla\boldsymbol{a}$ belongs to Y. Hence Lemma 5.2 leads to the representation of its distributional divergence as a distributional gradient of some $\omega\in L_0^2(\Omega)$,

$$\operatorname{div}(\mathcal{A}\,\boldsymbol{\varepsilon}-\nabla\boldsymbol{a})=\nabla\omega\quad\text{and}\quad c_0\,\|\omega\|_0\le\|\mathcal{A}\,\boldsymbol{\varepsilon}-\nabla\boldsymbol{a}\|_0.$$

The orthogonal decomposition $\boldsymbol{L}^2(\Omega;\mathbb{M}^{2\times 2})=\nabla X\oplus Y$, motivates the split

$$\|\mathcal{A}\,\boldsymbol{\varepsilon}\|^2=(\mathcal{A}\,\boldsymbol{\varepsilon},\nabla\boldsymbol{a})+(\mathcal{A}\,\boldsymbol{\varepsilon},\mathcal{A}\,\boldsymbol{\varepsilon}-\nabla\boldsymbol{a}).$$

The analysis of the first term $(\mathcal{A}\,\boldsymbol{\varepsilon},\nabla\boldsymbol{a})$ involves $\operatorname{div}\boldsymbol{a}=\operatorname{tr}\,\nabla\boldsymbol{a}=0$ and an integration by parts,

$$(\mathcal{A}\,\boldsymbol{\varepsilon},\nabla\boldsymbol{a})=(\boldsymbol{\varepsilon},\nabla\boldsymbol{a})=(-\boldsymbol{f}+\operatorname{div}\boldsymbol{\sigma}_h,\boldsymbol{a}).$$

Since $\operatorname{div}\boldsymbol{\sigma}_h=\boldsymbol{f}_h$ is the piecewise polynomial best approximation of $\boldsymbol{f}$,

$$(\mathcal{A}\,\boldsymbol{\varepsilon},\nabla\boldsymbol{a})=-(\boldsymbol{f}-\boldsymbol{f}_h,\boldsymbol{a})\le\|\boldsymbol{f}-\boldsymbol{f}_h\|_{-1}\|\nabla\boldsymbol{a}\|_0.$$

The analysis of the second term involves a test function $\boldsymbol{v}\in\boldsymbol{g}+\boldsymbol{H}_0^1(\Omega)$ and utilizes $\mathcal{A}\,\boldsymbol{\varepsilon}=\nabla\boldsymbol{u}-\mathcal{A}\,\boldsymbol{\sigma}_h$. Then, $\boldsymbol{u}-\boldsymbol{g}\in X$ and so $\mathcal{A}\,\boldsymbol{\varepsilon}-\nabla\boldsymbol{a}$ is orthogonal onto $\nabla(\boldsymbol{u}-\boldsymbol{g})$, i.e.,

$$(\mathcal{A}\,\boldsymbol{\varepsilon},\mathcal{A}\,\boldsymbol{\varepsilon}-\nabla a)=(\mathcal{A}(\nabla\boldsymbol{v}-\boldsymbol{\sigma}_h),\mathcal{A}\,\boldsymbol{\varepsilon}-\nabla\boldsymbol{a})+(\nabla(\boldsymbol{g}-\boldsymbol{v}),\mathcal{A}\,\boldsymbol{\varepsilon}-\nabla\boldsymbol{a}).$$

The last term involves the distributional divergence of $\mathcal{A}\,\boldsymbol{\varepsilon}-\nabla\boldsymbol{a}$ which equals the gradient of the aforementioned function ω. Therefore and with $\operatorname{div}\boldsymbol{g}=0$,

$$(\nabla(\boldsymbol{g}-\boldsymbol{v}),\mathcal{A}\,\boldsymbol{\varepsilon}-\nabla a)=(\omega,\operatorname{div}\boldsymbol{v})\le\|\omega\|_0\|\operatorname{div}\boldsymbol{v}\|_0.$$

The aforementioned estimate of $\|\omega\|_0$ with the inf-sup constant c_0 allows for a control of the last term. The combination of the resulting estimate with the previous inequalities leads to

$$(\mathcal{A}\,\boldsymbol{\varepsilon},\mathcal{A}\,\boldsymbol{\varepsilon}-\nabla\boldsymbol{a})\le\|\mathcal{A}\,\boldsymbol{\varepsilon}-\nabla\boldsymbol{a}\|_0\Big(\|\mathcal{A}(\nabla\boldsymbol{v}-\boldsymbol{\sigma}_h)\|_0+\|\operatorname{div}\boldsymbol{v}\|_0/c_0\Big).$$

The combination of the two estimates with the orthogonal sum

$$\|\mathcal{A}\,\boldsymbol{\varepsilon}\|_0^2=\|\nabla\boldsymbol{a}\|_0^2+\|\mathcal{A}\,\boldsymbol{\varepsilon}-\nabla\boldsymbol{a}\|_0^2$$

concludes the proof of the theorem. ∎

Remark that $\eta_h\approx\mu_h$. This follows from the orthogonality of deviatoric and isotropic part of matrices and Pythagoras' theorem ($|\cdot|$ denotes the Frobenius norm for matrices)

$$|(\mathcal{A}\,\boldsymbol{\sigma}_h-\nabla\boldsymbol{v})(x)|^2=|\mathcal{A}\,(\boldsymbol{\sigma}_h-\nabla\boldsymbol{v})(x)|^2+|\operatorname{tr}(\nabla\boldsymbol{v})|^2/2.$$

The explicit residual-based error estimators can be derived following arguments in the literature [18, 13, 16, 19, 28].

THEOREM 5.4. *With the tangential unit vector* $\boldsymbol{t}_E$ *and the jump of the tangential components of the discrete stress deviator* $[\boldsymbol{t}_E \cdot \mathcal{A}\boldsymbol{\sigma}_h]$, *it holds reliability in the sense*

$$\|\boldsymbol{\varepsilon}\|_0 \lesssim \eta^{Res} := \left(\sum_T \|h_T(\boldsymbol{f} - \boldsymbol{f}_h)\|_{0,T}^2 + \sum_T \|h_T \mathbf{curl}(\mathcal{A}\boldsymbol{\sigma}_h)\|_{0,T}^2 + \sum_E \|h_E^{1/2}[\boldsymbol{t}_E \cdot \mathcal{A}\boldsymbol{\sigma}_h]\|_{0,E}^2 \right)^{1/2}.$$

Local efficiency holds in the sense that, for each $T \in \mathcal{T}_h$,

$$\|h_T \mathbf{curl}(\mathcal{A}\boldsymbol{\sigma}_h)\|_{0,T} \lesssim \|\boldsymbol{\varepsilon}\|_{0,T}\,;$$

for each interior $E \in \mathcal{E}_h$ *with edge-patch* ω_E,

$$h_E^{1/2}\|[\boldsymbol{t}_E \cdot \mathcal{A}\boldsymbol{\sigma}_h]\|_{0,E} \lesssim \|\boldsymbol{\varepsilon}\|_{0,\omega_E}\,;$$

for any edge $E \subset \Gamma$ *on the boundary* (ω_E *is one element domain*),

$$\|h_E^{1/2}[\boldsymbol{t}_E \cdot \mathcal{A}\boldsymbol{\sigma}_h]\|_{0,E} = \|h_E^{1/2}\, \boldsymbol{t}_E \cdot (\nabla \boldsymbol{g} - \mathcal{A}\boldsymbol{\sigma}_h)\|_{0,E} \lesssim \|\boldsymbol{\varepsilon}\|_{0,\omega_E} + osc(\boldsymbol{t}_E \cdot \nabla \boldsymbol{g}, E)$$

with edge oscillations $osc(\gamma, E) := h_E^{1/2}\|\gamma - \gamma_E\|_{0,E}$ *for* $\gamma := \boldsymbol{t}_E \cdot \nabla \boldsymbol{g}$ *and its polynomial* $L^2(E)$ *best approximation* γ_E *in* $P_k(E)^2$.

Proof. For the reliability in view of (26), first note that the Stokes problem with right-hand side $(\operatorname{div}\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h), 0)$ instead of $(\boldsymbol{f}, \boldsymbol{g})$ in (1) leads to some solution $(\boldsymbol{a}, b) \in \boldsymbol{H}_0^1(\Omega) \times L_0^2(\Omega)$ with $\operatorname{div}\boldsymbol{a} = 0$ and

$$\|\boldsymbol{a}\|_{\boldsymbol{H}^1(\Omega)} + \|b\|_{L^2(\Omega)} \lesssim \|\operatorname{div}\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_{\boldsymbol{H}^{-1}(\Omega)} \lesssim \|\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_{L^2(\Omega)}.$$

The Stokes equations imply that $\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) - \nabla \boldsymbol{a} + b\boldsymbol{\delta}$ is divergence free in the simply connected domain Ω and hence equals some rotation $\mathbf{curl}\boldsymbol{r}$ of some function $\boldsymbol{r} \in \boldsymbol{H}^1(\Omega)$. The traces in this Helmholtz decomposition

$$\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) = \nabla \boldsymbol{a} - b\boldsymbol{\delta} + \mathbf{curl}\boldsymbol{r}$$

plus $\operatorname{div}\boldsymbol{a} = 0 = \operatorname{div}\boldsymbol{u}$ and $\mathcal{A}\boldsymbol{\sigma} = \nabla \boldsymbol{u}$ imply

$$\min_{\boldsymbol{v} \in \boldsymbol{g} + \boldsymbol{H}_0^1(\Omega)} \|\mathcal{A}\,\boldsymbol{\sigma}_h - \nabla \boldsymbol{v}\|_0 \le \|\mathcal{A}\boldsymbol{\sigma}_h - \nabla(\boldsymbol{u} - \boldsymbol{a})\|_0 = \|b\boldsymbol{\delta} - \mathbf{curl}\boldsymbol{r}\|_0.$$

On the other hand, the orthogonality in the Helmholtz decomposition up to boundary terms lead to

$$\|b\boldsymbol{\delta} - \mathbf{curl}\boldsymbol{r}\|_0^2 = (b\boldsymbol{\delta} - \mathbf{curl}\boldsymbol{r}, \mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)) = (\mathbf{curl}\boldsymbol{r}, \mathcal{A}\boldsymbol{\sigma}_h) - \langle \boldsymbol{g}, \mathbf{curl}\boldsymbol{r}\boldsymbol{n}\rangle.$$

Given any weak interpolation $\boldsymbol{r}_h$ of $\boldsymbol{r}$ in piecewise affine and globally continuous functions, the divergence of $\mathbf{curl}\boldsymbol{r}_h$ exists in $\boldsymbol{L}^2(\Omega)$ and vanishes. Hence $\boldsymbol{\tau}_h := \mathbf{curl}\boldsymbol{r}_h - s\boldsymbol{\delta} \in \boldsymbol{\Phi}_h$ with the real number $s := \int_\Omega \operatorname{tr} \mathbf{curl}\boldsymbol{r}_h \, dx/2$, leads to

$$(\mathcal{A}\boldsymbol{\varepsilon}, \boldsymbol{\tau}_h) = (\boldsymbol{e}, \operatorname{div}\boldsymbol{\tau}_h) = 0.$$

This implies $< \boldsymbol{g}, \mathbf{curl}\boldsymbol{r}_h\boldsymbol{n} >= (\mathcal{A}\boldsymbol{\sigma}_h, \mathbf{curl}\boldsymbol{r}_h)$. Consequently

$$\|b\boldsymbol{\delta} - \mathbf{curl}\boldsymbol{r}\|_0^2 = (\mathbf{curl}(\boldsymbol{r} - \boldsymbol{r}_h), \mathcal{A}\boldsymbol{\sigma}_h) - \langle \boldsymbol{g}, \mathbf{curl}(\boldsymbol{r} - \boldsymbol{r}_h)\boldsymbol{n}\rangle.$$

An elementwise integration by parts leads to the jumps of tangential components $\boldsymbol{t}\cdot\mathcal{A}\boldsymbol{\sigma}_h$ across interior edges of triangles and to the terms $\boldsymbol{t}\cdot(\nabla\boldsymbol{g}-\mathcal{A}\boldsymbol{\sigma}_h)$ along the outer boundary. Standard stability and approximation results then conclude the proof of the asserted reliability estimate.

The efficiency follows from the discrete test-function technology due to Verfürth and is standard amongst the experts. The function $\boldsymbol{d}:=[\boldsymbol{t}_E\cdot\mathcal{A}\boldsymbol{\sigma}_h]=-[\boldsymbol{t}_E\cdot\mathcal{A}\varepsilon]$ is a polynomial along an edge $E\in\mathcal{E}_h$ with a quadratic function ψ_E defined as the product of the two nodal basis functions of a conforming linear finite element scheme. With $\mathbf{curl}(\mathcal{A}\boldsymbol{\sigma})=\mathbf{0}$, one deduces

$$\|\boldsymbol{d}\|_{0,E}^2\lesssim\|\boldsymbol{\psi}_E^{1/2}\boldsymbol{d}\|_{0,E}^2=-\langle\boldsymbol{\psi}_E\boldsymbol{d},[\boldsymbol{t}_E\cdot\mathcal{A}\varepsilon]\rangle_E.$$

Some extension operator P_{ext} of the polynomial and an integration by parts separately on the two neighboring triangles of the edge patch $\omega_E=\{\psi_E>0\}$ show that the previous integral over the edge E equals

$$\begin{aligned}&(\boldsymbol{\psi}_E P_{\mathrm{ext}}\boldsymbol{d},\mathbf{curl}(\mathcal{A}\boldsymbol{\sigma}_h))_{L^2(\omega_E)}-(\mathcal{A}\boldsymbol{\varepsilon},\mathbf{curl}(\boldsymbol{\psi}_E P_{\mathrm{ext}}\boldsymbol{d}))_{L^2(\omega_E)}\\&\quad\le\|\boldsymbol{\varepsilon}\|_{L^2(\omega_E)}|\boldsymbol{\psi}_E P_{\mathrm{ext}}\boldsymbol{d}|_{1,\omega_E}+\ \|\boldsymbol{\psi}_E P_{\mathrm{ext}}\boldsymbol{d}\|_{L^2(\omega_E)}\|\mathbf{curl}(\mathcal{A}\boldsymbol{\sigma}_h)\|_{L^2(\omega_E)}.\end{aligned}\tag{28}$$

Some stability analysis of the edge-bubble functions as well as of the extension operator leads to

$$\|\boldsymbol{d}\|_{0,E}\lesssim\|\boldsymbol{\varepsilon}\|_{L^2(\omega_E)}+\|\mathbf{curl}(\mathcal{A}\boldsymbol{\sigma}_h)\|_{L^2(\omega_E)}.$$

Together with the first asserted efficiency estimate for $\|\mathbf{curl}(\mathcal{A}\boldsymbol{\sigma}_h)\|_{L^2(T)}$, this proves the second. Since the remaining details are standard and well known for elliptic PDEs, further details are omitted. □

6. A posteriori velocity error estimation. A duality technique allows the control of the velocity error $\boldsymbol{e}=\boldsymbol{u}-\boldsymbol{u}_h$ via the regularity results (4). Recall that $\boldsymbol{u}_h^*$ denotes the postprocessed solution of section 4 and consider $\boldsymbol{e}^*=\boldsymbol{u}-\boldsymbol{u}_h^*$. The exact solution $\boldsymbol{u}$ and its piecewise polynomial L^2 best approximation $\boldsymbol{P}_h\boldsymbol{u}$ in $P_k(\mathcal{T}_h)^2$ satisfy

$$\|\boldsymbol{e}\|_0^2=\|\boldsymbol{u}-\boldsymbol{P}_h\boldsymbol{u}\|_0^2+\|\boldsymbol{P}_h\boldsymbol{e}\|_0^2.$$

While the first term on the right-hand side is expected to be of order $k+1$, the second may be of higher order. This is seen from the a priori error estimates of Theorem 4.1 and $\boldsymbol{P}_h\boldsymbol{u}_h^*=\boldsymbol{P}_h\boldsymbol{u}_h=\boldsymbol{u}_h$ (provided that the Stokes problem is $\boldsymbol{H}^{k+2}$-regular),

$$\|\boldsymbol{P}_h\boldsymbol{e}^*\|_0=\|\boldsymbol{P}_h\boldsymbol{e}\|_0\le\|\boldsymbol{u}-\boldsymbol{u}_h^*\|_0\lesssim h^{k+2}.$$

The following list of a posteriori error estimates reflects this higher order behavior and the leading first-order term. The smooth right-hand side enters in terms of oscillations of order $k+\ell+1\ge k+2$ with $\ell=1$ for $k=0$ while $\ell=2$ for $k\ge1$,

$$\mathrm{osc}_{2,k}(\boldsymbol{f},\mathcal{T}_h)^2:=\sum_{T\in\mathcal{T}_h}\min_{\boldsymbol{f}_h\in P_k(T)^2}h_T^{2\ell}\|\boldsymbol{f}-\boldsymbol{f}_h\|_{0,T}^2.$$

THEOREM 6.1. *Provided that the Stokes problem is $\boldsymbol{H}^2$-regular, it holds that*

$$\text{(a)}\qquad\|\boldsymbol{P}_h\boldsymbol{u}-\boldsymbol{u}_h\|_0\lesssim osc_{2,k}(\boldsymbol{f},\mathcal{T}_h)+\min_{\boldsymbol{v}\in\boldsymbol{g}+\boldsymbol{H}_0^1(\Omega)}\|h(\mathcal{A}\,\boldsymbol{\sigma}_h-\nabla\boldsymbol{v})\|_0.$$

With $C \approx 1$, it also holds that

(b) $\|\boldsymbol{u}-\boldsymbol{u}_h\|_0 \le C\, osc_{2,k}(\boldsymbol{f},\mathcal{T}_h) + \min_{\boldsymbol{v}\in\boldsymbol{g}+\boldsymbol{H}_0^1(\Omega)} \Big(\|\boldsymbol{v}-\boldsymbol{P}_h\boldsymbol{v}\|_0 + C\|h(\mathcal{A}\,\boldsymbol{\sigma}_h - \nabla\boldsymbol{v})\|_0\Big)$ *and*

(c) $\|\boldsymbol{u}-\boldsymbol{u}_h^*\|_0 \le C\, osc_{2,k}(\boldsymbol{f},\mathcal{T}_h)$
$$+ \min_{\boldsymbol{v}\in\boldsymbol{g}+\boldsymbol{H}_0^1(\Omega)} \Big(\|\boldsymbol{v}-\boldsymbol{u}_h^* - \boldsymbol{P}_h(\boldsymbol{v}-\boldsymbol{u}_h^*)\|_0 + C\|h(\mathcal{A}\,\boldsymbol{\sigma}_h - \nabla\boldsymbol{v})\|_0\Big).$$

Proof. In the duality argument, consider the unique solution $(\boldsymbol{\eta},\boldsymbol{z}) \in \boldsymbol{\Phi}\times\boldsymbol{V}\cap\boldsymbol{H}_0^1(\Omega)$ to

$$(\mathcal{A}\boldsymbol{\eta},\boldsymbol{\tau}) + (\mathbf{div}\boldsymbol{\tau},\boldsymbol{z}) = 0 \quad \text{for all } \boldsymbol{\tau}\in\boldsymbol{\Phi}, \tag{29}$$
$$(\mathbf{div}\boldsymbol{\eta},\boldsymbol{v}) = (\boldsymbol{E},\boldsymbol{v}) \quad \text{for all } \boldsymbol{v}\in\boldsymbol{V} \tag{30}$$

of the Stokes equations for three right-hand sides, namely,

$$\boldsymbol{E} = \boldsymbol{P}_h\boldsymbol{u}-\boldsymbol{u}_h, \quad \boldsymbol{E}=\boldsymbol{e}=\boldsymbol{u}-\boldsymbol{u}_h, \quad \text{and} \quad \boldsymbol{E}=\boldsymbol{u}-\boldsymbol{u}_h^* \text{ in } \boldsymbol{L}^2(\Omega)^2,$$

with the regularity properties

$$\|\boldsymbol{z}\|_2 + \|\boldsymbol{\eta}\|_1 \lesssim \|\boldsymbol{E}\|_0. \tag{31}$$

An integration by parts, (3), and the relation $\mathbf{div}\Pi_h = \boldsymbol{P}_h\mathbf{div}$ lead to

$$\begin{aligned}\|\boldsymbol{E}\|_0^2 &= (\boldsymbol{E}-\boldsymbol{e},\mathbf{div}\boldsymbol{\eta}) + (\boldsymbol{e},\mathbf{div}\boldsymbol{\eta})\\ &= (\boldsymbol{E}-\boldsymbol{e},\mathbf{div}\boldsymbol{\eta}) - (\mathcal{A}\boldsymbol{\sigma},\boldsymbol{\eta})\\ &\quad + \langle\boldsymbol{\eta}\boldsymbol{n},\boldsymbol{g}\rangle - (\boldsymbol{u}_h,\mathbf{div}(\Pi_h\boldsymbol{\eta})).\end{aligned}$$

The identities (7)–(8) and (29) allow the subsequent calculations,

$$\begin{aligned}&-(\mathcal{A}\boldsymbol{\sigma},\boldsymbol{\eta}) - (\boldsymbol{u}_h,\mathbf{div}(\Pi_h\boldsymbol{\eta})) + \langle\Pi_h\boldsymbol{\eta}\boldsymbol{n},\boldsymbol{g}\rangle\\ &\quad= -(\mathcal{A}\boldsymbol{\sigma}-\mathcal{A}\boldsymbol{\sigma}_h,\boldsymbol{\eta}) - (\mathcal{A}\boldsymbol{\sigma}_h,\boldsymbol{\eta}) + (\mathcal{A}\boldsymbol{\sigma}_h,\Pi_h\boldsymbol{\eta})\\ &\quad= -(\mathcal{A}\boldsymbol{\eta},\boldsymbol{\sigma}-\boldsymbol{\sigma}_h) + (\mathcal{A}\boldsymbol{\sigma}_h,\Pi_h\boldsymbol{\eta}-\boldsymbol{\eta})\\ &\quad= (\mathbf{div}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),\boldsymbol{z}) + (\mathcal{A}\boldsymbol{\sigma}_h,\Pi_h\boldsymbol{\eta}-\boldsymbol{\eta}).\end{aligned}$$

Given any $\boldsymbol{v}\in\boldsymbol{g}+\boldsymbol{H}_0^1(\Omega)$, an integration by parts leads to

$$(\nabla\boldsymbol{v},\Pi_h\boldsymbol{\eta}-\boldsymbol{\eta}) = \langle(\Pi_h\boldsymbol{\eta}-\boldsymbol{\eta})\boldsymbol{n},\boldsymbol{g}\rangle - (\boldsymbol{v},\mathbf{div}(\Pi_h\boldsymbol{\eta}-\boldsymbol{\eta}))$$

and hence to

$$\begin{aligned}&-(\mathcal{A}\boldsymbol{\sigma},\boldsymbol{\eta}) - (\boldsymbol{u}_h,\mathbf{div}(\Pi_h\boldsymbol{\eta})) + \langle\boldsymbol{\eta}\boldsymbol{n},\boldsymbol{g}\rangle\\ &\quad= (\boldsymbol{f}-\mathbf{div}\boldsymbol{\sigma}_h,\boldsymbol{z}-\boldsymbol{P}_h\boldsymbol{z}) + (\mathcal{A}\boldsymbol{\sigma}_h-\nabla\boldsymbol{v},\Pi_h\boldsymbol{\eta}-\boldsymbol{\eta}) + (\boldsymbol{v},\boldsymbol{E}-\boldsymbol{P}_h\boldsymbol{E}).\end{aligned}$$

The elementwise Cauchy inequality plus the approximation property $\boldsymbol{z}-\boldsymbol{P}_h\boldsymbol{z}$ of order $k+1$ leads to

$$(\boldsymbol{f}-\boldsymbol{f}_h,\boldsymbol{z}-\boldsymbol{P}_h\boldsymbol{z}) \le C\,\mathrm{osc}_{2,k}(\boldsymbol{f},\mathcal{T}_h)\,\|\boldsymbol{z}\|_2.$$

The combination of the established estimates leads to the core estimate of this proof,

$$\begin{aligned}\|\boldsymbol{E}\|_0^2 &\le C\,\mathrm{osc}_{2,k}(\boldsymbol{f},\mathcal{T}_h)\,\|\boldsymbol{z}\|_2 + (\mathcal{A}\boldsymbol{\sigma}_h-\nabla\boldsymbol{v},\Pi_h\boldsymbol{\eta}-\boldsymbol{\eta})\\ &\quad + (\boldsymbol{E}-\boldsymbol{e},\boldsymbol{E}) + (\boldsymbol{v}-\boldsymbol{P}_h\boldsymbol{v},\boldsymbol{E}-\boldsymbol{P}_h\boldsymbol{E}).\end{aligned} \tag{32}$$

In the first case, let $\boldsymbol{E} = \mathbf{div}\boldsymbol{\eta} = \boldsymbol{P}_h\boldsymbol{e} = \mathbf{div}\Pi_h\boldsymbol{\eta}$ be a piecewise polynomial of degree $\leq k$. Then, $\boldsymbol{E} = \boldsymbol{P}_h\boldsymbol{E}$ and $(\boldsymbol{E}-\boldsymbol{e},\boldsymbol{E}) = 0$. Since

$$(\mathcal{A}\boldsymbol{\sigma}_h - \nabla\boldsymbol{v}, \Pi_h\boldsymbol{\eta}-\boldsymbol{\eta}) \lesssim \|h(\mathcal{A}\,\boldsymbol{\sigma}_h - \nabla\boldsymbol{v})\|_0\,\|\boldsymbol{\eta}\|_1,$$

the core estimate (32) leads to

$$\|\boldsymbol{P}_h\boldsymbol{u}-\boldsymbol{u}_h\|_0^2 \lesssim \mathrm{osc}_{2,k}(\boldsymbol{f},\mathcal{T}_h)\,\|\boldsymbol{z}\|_2 + \|h(\mathcal{A}\boldsymbol{\sigma}_h-\nabla\boldsymbol{v})\|_0\|\boldsymbol{\eta}\|_1.$$

This and the regularity estimate (31) prove the assertion (a).

In the second case, let $\boldsymbol{E} = \mathbf{div}\boldsymbol{\eta} = \boldsymbol{e}$ and hence $(\boldsymbol{E}-\boldsymbol{e},\boldsymbol{E}) = 0$ and

$$(\boldsymbol{v},\boldsymbol{E}-\boldsymbol{P}_h\boldsymbol{E}) = (\boldsymbol{v}-\boldsymbol{P}_h\boldsymbol{v},\boldsymbol{e}-\boldsymbol{P}_h\boldsymbol{e}) = (\boldsymbol{v}-\boldsymbol{P}_h\boldsymbol{v},\boldsymbol{e}) \leq \|\boldsymbol{v}-\boldsymbol{P}_h\boldsymbol{v}\|_0\,\|\boldsymbol{e}\|_0.$$

This and the aforementioned arguments with core and regularity estimate (32)–(31) prove the assertion (b).

In the third case, let $\boldsymbol{E} = \mathbf{div}\boldsymbol{\eta} = \boldsymbol{u}-\boldsymbol{u}_h^*$, and notice that

$$\boldsymbol{E}-\boldsymbol{e} = \boldsymbol{u}_h - \boldsymbol{u}_h^* = \boldsymbol{P}_h\boldsymbol{u}_h^* - \boldsymbol{u}_h^*$$

is perpendicular to $P_k(\mathcal{T}_h)^2$. Hence

$$(\boldsymbol{E}-\boldsymbol{e},\boldsymbol{E}) + (\boldsymbol{v}-\boldsymbol{P}_h\boldsymbol{v},\boldsymbol{E}) = (\boldsymbol{v}-\boldsymbol{u}_h^*,\boldsymbol{E}-\boldsymbol{P}_h\boldsymbol{E}) \leq \|(\boldsymbol{\delta}-\boldsymbol{P}_h)(\boldsymbol{v}-\boldsymbol{u}_h^*)\|_0\,\|\boldsymbol{E}\|_0.$$

This and the aforementioned arguments lead to assertion (c). ∎

The explicit residual-based error estimators mainly follow from arguments in the literature [18, 13, 16, 19, 28].

THEOREM 6.2. *Let $\boldsymbol{u}_h^*$ be the postprocessed solution of* (21)–(22) *and let $\boldsymbol{e}^* = \boldsymbol{u}-\boldsymbol{u}_h^*$ for an $\boldsymbol{H}^2$ regular problem. Reliability holds in the sense of*

$$\|\boldsymbol{e}^*\|_0^2 \lesssim osc_{2,k}(\boldsymbol{f},\mathcal{T}_h)^2 + \sum_T \|h_T(\mathcal{A}\boldsymbol{\sigma}_h - \nabla\boldsymbol{u}_h^*)\|_{0,T}^2 + \sum_E \|h_E^{1/2}[\boldsymbol{u}_h^*]\|_{0,E}^2.$$

Local efficiency holds, for each $T\in\mathcal{T}_h$, in the sense that

$$h_T\|\mathcal{A}\boldsymbol{\sigma}_h - \nabla\boldsymbol{u}_h^*\|_{0,T} \lesssim h_T\|\boldsymbol{\varepsilon}\|_{0,T} + \|\boldsymbol{e}^*\|_{0,T}$$

and, for each interior $E\in\mathcal{E}_h$ with edge-patch ω_E,

$$h_E^{1/2}\|[\boldsymbol{u}_h^*]\|_{0,E} \lesssim \|\boldsymbol{e}^*\|_{0,w_E} + h_T\|\boldsymbol{\varepsilon}\|_{0,\omega_E};$$

while, for any edge $E\subset\Gamma$ on the boundary (ω_E is one element domain),

$$h_E^{1/2}\|[\boldsymbol{u}_h^*]\|_{0,E} = h_E^{1/2}\|\,\boldsymbol{g}-\boldsymbol{u}_h^*\|_{0,E} \lesssim \|\boldsymbol{e}^*\|_{0,w_E} + h_T\|\boldsymbol{\varepsilon}\|_{0,\omega_E} + h_E^{1/2}\|\boldsymbol{g}-\boldsymbol{g}_E\|_{0,E}$$

with the $\boldsymbol{L}^2(E)$ best approximation $\boldsymbol{g}_E$ of $\boldsymbol{g}$ in $P_k(E)^2$.

Proof. The proof of reliability follows the lines of case (c) in the duality argument of the proof of Theorem 6.1 but does not involve the variable $\boldsymbol{v}$. Instead, add/subtract the term $(\nabla_h\boldsymbol{u}_h^*,\Pi_h\boldsymbol{\eta}-\boldsymbol{\eta})$. An elementwise integration by parts plus a careful rearrangement of the inner element boundary terms leads—in continuation of the aforementioned part (c)—to

$$(\nabla_h\boldsymbol{u}_h^*,\Pi_h\boldsymbol{\eta}-\boldsymbol{\eta}) + \langle(\boldsymbol{\eta}-\Pi_h\boldsymbol{\eta})\boldsymbol{n},\boldsymbol{g}\rangle = (\boldsymbol{u}_h^*,\mathrm{div}(\boldsymbol{\eta}-\Pi_h\boldsymbol{\eta})) + \sum_E\int_E [\boldsymbol{u}_h^*]\,(\boldsymbol{\eta}-\Pi_h\boldsymbol{\eta})\cdot\boldsymbol{n}_E\,ds.$$

Standard arguments on the jump terms lead to the weighted estimators $\|h_E^{1/2}[\boldsymbol{u}_h^*]\|_{0,E}$. Direct calculations reveal

$$\begin{aligned}(\boldsymbol{u}_h^*, \mathrm{div}(\boldsymbol{\eta} - \Pi_h\boldsymbol{\eta})) &= (\boldsymbol{u}_h^*, \mathrm{div}\boldsymbol{\eta}) - (\boldsymbol{u}_h^*, \mathrm{div}\Pi_h\boldsymbol{\eta})\\ &= (\boldsymbol{u}_h^*, \mathrm{div}\boldsymbol{\eta}) - (\boldsymbol{P}_h\boldsymbol{u}_h^*, \mathrm{div}\boldsymbol{\eta}) = (\boldsymbol{u}_h^* - \boldsymbol{P}_h\boldsymbol{u}_h^*, \boldsymbol{E}).\end{aligned}$$

Since $\boldsymbol{E} = \mathbf{div}\boldsymbol{\eta} = \boldsymbol{u} - \boldsymbol{u}_h^*$ is controlled in this setting, one concludes the proof of reliability by elementary algebra.

To prove the local efficiency on some triangle T, we employ some cubic bubble function $\boldsymbol{\psi}_T$ as the product of the nodal basis function of all the vertices of T. By some inverse inequality and an integration by parts,

$$\begin{aligned}\|\mathcal{A}\boldsymbol{\sigma}_h - \nabla\boldsymbol{u}_h^*\|_{0,T}^2 &\lesssim (\mathcal{A}\boldsymbol{\sigma}_h - \nabla\boldsymbol{u}_h^*, \boldsymbol{\psi}_T(\mathcal{A}\boldsymbol{\sigma}_h - \nabla\boldsymbol{u}_h^*))_T\\ &= -(\mathcal{A}\boldsymbol{\varepsilon}, \boldsymbol{\psi}_T(\mathcal{A}\boldsymbol{\sigma}_h - \nabla\boldsymbol{u}_h^*))_T + (\nabla\boldsymbol{u} - \nabla\boldsymbol{u}_h^*, \boldsymbol{\psi}_T(\mathcal{A}\boldsymbol{\sigma}_h - \nabla\boldsymbol{u}_h^*))_T\\ &= -(\mathcal{A}\boldsymbol{\varepsilon}, \boldsymbol{\psi}_T(\mathcal{A}\boldsymbol{\sigma}_h - \nabla\boldsymbol{u}_h^*))_T - (\boldsymbol{e}^*, \mathbf{div}\boldsymbol{\psi}_T(\mathcal{A}\boldsymbol{\sigma}_h - \nabla\boldsymbol{u}_h^*))_T\\ &\le \|\mathcal{A}\boldsymbol{\varepsilon}\|_{0,T}\|\mathcal{A}\boldsymbol{\sigma}_h - \nabla\boldsymbol{u}_h^*\|_{0,T} + \|\boldsymbol{e}^*\|_{0,T}|\boldsymbol{\psi}_T(\mathcal{A}\boldsymbol{\sigma}_h - \nabla\boldsymbol{u}_h^*)|_{1,T}.\end{aligned}$$

Since $\|\mathcal{A}\boldsymbol{\varepsilon}\|_T \le \|\boldsymbol{\varepsilon}\|_T$, some further inverse inequality

$$|\boldsymbol{\psi}_T(\mathcal{A}\boldsymbol{\sigma}_h - \nabla\boldsymbol{u}_h^*)|_{1,T} \lesssim h_T^{-1}\|\mathcal{A}\boldsymbol{\sigma}_h - \nabla\boldsymbol{u}_h^*\|_{0,T}$$

leads to the asserted efficiency of the volume contribution

$$h_T\|\mathcal{A}\boldsymbol{\sigma}_h - \nabla\boldsymbol{u}_h^*\|_{0,T} \lesssim h_T\|\boldsymbol{\varepsilon}\|_{0,T} + \|\boldsymbol{e}^*\|_{0,T}.$$

The trace inequality leads, with the piecewise gradient ∇_h, to

$$h_E^{1/2}\|[\boldsymbol{u}_h^*]\|_{0,E} \lesssim \|\boldsymbol{e}^*\|_{0,\omega_E} + h_E\|\nabla_h\boldsymbol{e}^*\|_{0,\omega_E}.$$

The triangle inequality for $\nabla_h\boldsymbol{e}^* = \mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) + (\mathcal{A}\boldsymbol{\sigma}_h - \nabla_h\boldsymbol{u}_h^*)$ results in

$$h_E^{1/2}\|[\boldsymbol{u}_h^*]\|_{0,E} \lesssim \|\boldsymbol{e}^*\|_{0,\omega_E} + h_E\|\mathcal{A}\boldsymbol{\varepsilon}\|_{0,\omega_E} + h_E\|\mathcal{A}\boldsymbol{\sigma}_h - \nabla_h\boldsymbol{u}_h^*\|_{0,\omega_E}.$$

This and the previous efficiency on T prove the asserted efficiency of the edge contributions for an interior edge. An exterior edge involves the boundary date $\boldsymbol{g}$ and its approximation $\boldsymbol{g}_E$ utilized in the jump terms $[\boldsymbol{u}_h^*] = \boldsymbol{u} - \boldsymbol{u}_h^*$. With the edge patch ω_T which consists of one triangle $\overline{\omega_T}$ aligned to the edge E, the trace inequality leads to the above second-to-last displayed estimate. The triangle inequality of the last displayed estimate again concludes the proof as in the case of an interior edge. ∎

7. Numerical results. This section presents numerical experiments for the validation of the theoretical results and investigates the accuracy of the guaranteed upper error bounds. We implemented our scheme (7)–(8) with the Raviart–Thomas mixed finite elements for $k = 0$.

The a posteriori stress error estimator η_h of (27) requires a finite dimensional approximation for the actual computation. The numerical realization is addressed in the first subsection and adaptive algorithm follows in subsection 7.2. Numerical examples are given in subsections 7.3 and 7.4 with concluding remarks in the subsection 7.5.

7.1. Numerical realization. The a posteriori stress error estimator η_h of (27) from Theorem 5.3 with $c_0 = 0.4$ [17, 27] utilizes eight different functions $\boldsymbol{v}$ as follows. Let $\mathcal{N}_h$ be the set of nodes of the triangulation $\mathcal{T}_h$ and let φ_z be a nodal basis function of the conforming P_1-element on node z supported on the patch $\bigcup \mathcal{T}_h(z)$ for $\mathcal{T}_h(z) := \{T \in \mathcal{T}_h \,|\, z \in \mathcal{N}_h \cap T\}$. When $\measuredangle(T,z)$ denotes the interior angle of a vertex z of a triangle T, the averaged function $A(\boldsymbol{w}) \in \boldsymbol{g}_h + C_0(\Omega;\mathbb{V}^2)$ of $\boldsymbol{w} \in C(\mathcal{T}_h;\mathbb{V}^2)$ reads

$$A(\boldsymbol{w}) := \sum_{z\in\mathcal{N}_h} \mathcal{I}_z\boldsymbol{w}\varphi_z \quad \text{with } \mathcal{I}_z\boldsymbol{w} := \begin{cases} \displaystyle\sum_{T\in\mathcal{T}_h(z)} \frac{\measuredangle(T,z)}{2\pi}\boldsymbol{w}|_T(z) & \text{for } z \in \mathcal{N}_h \setminus \partial\Omega, \\ \boldsymbol{g}(z) \quad \text{for } z \in \mathcal{N}_h \cap \partial\Omega. \end{cases}$$

Throughout, $\boldsymbol{g}_h$ is a continuous piecewise linear polynomial interpolation of $\boldsymbol{g}$ at the boundary nodes $\mathcal{N}_h \cap \partial\Omega$. Note that in case of nonvanishing $\boldsymbol{g} - \boldsymbol{g}_h$, the difference results in a perturbation term which can be estimated by some explicit extension to the inside of the triangles along the boundary and then controlled as in [14]. For the examples of this paper, the smooth data result in a higher-order perturbation term which is neglected in the numerical outcome displayed below.

For $\boldsymbol{v} = A(\boldsymbol{u}_h^*)$ and for $\boldsymbol{v} = A(\boldsymbol{u}_h)$, let $\eta^{A(\boldsymbol{u}_h^*)}$ and $\eta^{A(\boldsymbol{u}_h)}$ denote the respective error bound

$$\eta^{\boldsymbol{v}} := \left((\|\mathcal{A}(\boldsymbol{\sigma}_h - \nabla\boldsymbol{v})\|_0 + 2.5\|\operatorname{div}\boldsymbol{v}\|_0)^2 + \operatorname{osc}_k(\boldsymbol{f},\mathcal{T}_h)^2/\pi^2\right)^{1/2} \tag{33}$$

from Theorem 5.3. The form of $\eta^{\boldsymbol{v}}$ as a sum of two norms (rather than that of two squares) leads to a difficulty in the minimization for $\boldsymbol{v} \in \boldsymbol{g}_h + \boldsymbol{V}(\mathcal{T}_h)$ with

$$\boldsymbol{V}(\mathcal{T}_h) := \left\{\boldsymbol{v} \in C_0(\Omega;\mathbb{V}^2) \,\middle|\, \text{for all } T \in \mathcal{T}_h,\ \boldsymbol{v}|_T \in P_1(T;\mathbb{V}^2)\right\}.$$

The identity (shown by elementary analysis)

$$(a+b)^2 = \min_{s>0}((1+s)a^2 + (1+1/s)b^2)$$

suggests an iterative procedure to compute the minimum $(\eta^{GM})^2 - \operatorname{osc}_k(\boldsymbol{f},\mathcal{T}_h)^2/\pi^2 :=$

$$\min_{\boldsymbol{v}\in\boldsymbol{g}_h+\boldsymbol{V}(\mathcal{T}_h)} \min_{s>0} \left((1+s)\|\mathcal{A}(\boldsymbol{\sigma}_h - \nabla\boldsymbol{v})\|_0^2 + 6.25(1+1/s)\|\operatorname{div}\boldsymbol{v}\|_0^2\right)$$

as follows. Given any $\boldsymbol{v} \in \boldsymbol{V}(\mathcal{T}_h)$ like $\boldsymbol{v} = A(\boldsymbol{u}_h^*)$, set $a(\boldsymbol{v}) := \|\mathcal{A}(\boldsymbol{\sigma}_h - \nabla\boldsymbol{v})\|_0$ and $b(\boldsymbol{v}) := 2.5\|\operatorname{div}\boldsymbol{v}\|_0$. Set $s = b/a$ and compute $\boldsymbol{v}_h \in \boldsymbol{g}_h + \boldsymbol{V}(\mathcal{T}_h)$ with

$$\int_\Omega \mathcal{A}\nabla\boldsymbol{v}_h : \nabla\boldsymbol{w}_h dx + \frac{6.25}{s}\int_\Omega \operatorname{div}\boldsymbol{v}_h \operatorname{div}\boldsymbol{w}_h dx = \int_\Omega \mathcal{A}\boldsymbol{\sigma}_h : \nabla\boldsymbol{w}_h \quad \text{for all } \boldsymbol{w}_h \in \boldsymbol{V}(\mathcal{T}_h)$$

and restart with that solution $\boldsymbol{v}_h$.

Moreover, $\eta^{GM_{red}}$ is calculated by the same argument as computing η^{GM} on the updated mesh obtained through red refinement of the given mesh $\mathcal{T}_h$. This requires solving a global minimization problem. To reduce computational work for further choices of $\boldsymbol{v}$, local minimization is based on the patch

$$\bar{\omega}_z := \bigcup_{T\in\mathcal{T}_h(z)} \operatorname{conv}\{z, \operatorname{mid}(E_{z_1}), \operatorname{mid}(E_{z_2})\}$$

of Figure 1 with $T = \operatorname{conv}\{z, z_1, z_2\}$ and the midpoint $\operatorname{mid}(E_{z_j})$ of the edge $E_{z_j} = \operatorname{conv}\{z, z_j\}$. The solution of the local minimization problem

$$\min_{\boldsymbol{v}_z\in\boldsymbol{g}_{\bar{\omega}_z}+\boldsymbol{V}(\bar{\omega}_z)} (\|\mathcal{A}(\boldsymbol{\sigma}_h - \nabla\boldsymbol{v}_z)\|_{0,\bar{\omega}_z} + 2.5\|\operatorname{div}\boldsymbol{v}_z\|_{0,\bar{\omega}_z})^2$$

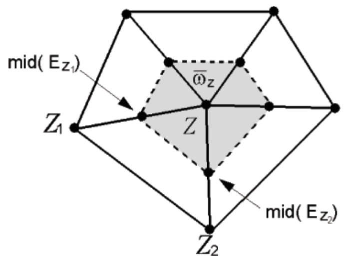


FIG. 1. *Patch of elements $\bar{\omega}_z$.*

with respect to $\boldsymbol{V}(\bar{\omega}_z) := \{\boldsymbol{w}_z \in C_0(\bar{\omega}_z;\mathbb{V}^2) \,|\, \text{for all } T \in \mathcal{T}_h(z),\ \boldsymbol{w}_z|_{T\cap\bar{\omega}_z} \in P_1(T\cap\bar{\omega}_z;\mathbb{V}^2)\}$ gives rise to four versions of $\eta^{LM} = \eta^{\boldsymbol{v}}$ owing to different continuous piecewise linear polynomials $\boldsymbol{g}_{\bar{\omega}_z}$ on the boundary $\partial\bar{\omega}_z$ of the patch $\bar{\omega}_z$. The four choices for the function $\boldsymbol{g}_{\bar{\omega}_z}$ read

$$A(\boldsymbol{u}_h^*)|_{\partial\bar{\omega}_z},\quad A(\boldsymbol{u}_h)|_{\partial\bar{\omega}_z},\quad \frac{1}{2}(\boldsymbol{u}_h^*|_T(\zeta)+\boldsymbol{u}_h^*|_{T'}(\zeta)),\quad \text{and } \frac{1}{2}(\boldsymbol{u}_h|_T(\zeta)+\boldsymbol{u}_h|_{T'}(\zeta))$$

and correspond to

$$\eta^{LM(A(\boldsymbol{u}_h^*))},\quad \eta^{LM(A(\boldsymbol{u}_h))},\quad \eta^{LM(\boldsymbol{u}_h^*)},\quad \text{and } \eta^{LM(\boldsymbol{u}_h)}$$

for a boundary node ζ of $\partial\bar{\omega}_z$ and $T, T' \in \mathcal{T}_h(z)$ with $\zeta \in \bar{T}\cap\bar{T}'$. Given any of those four $\boldsymbol{g}_{\bar{\omega}_z}$, set

$$\boldsymbol{v} := \sum_{z\in\mathcal{N}_h} \boldsymbol{v}_z(z)\varphi_z \quad \text{with } \boldsymbol{v}_z(z) = \boldsymbol{g}(z) \text{ for } z \in \mathcal{N}_h \cap \partial\Omega.$$

7.2. Adaptive mesh refinement. The following adaptive algorithm is motivated from the a posteriori error bounds of this paper and leads to improved convergence rates.

(a) Start with an initial mesh $\mathcal{T}_0$, $\ell = 0$.
(b) Solve the discrete problem (7)–(8) for $(\boldsymbol{u}_\ell, \boldsymbol{\sigma}_\ell)$ with respect to $\mathcal{T}_\ell$.
(c) Compute the following local (η_T) and global (η_ℓ) error estimators on $\mathcal{T}_\ell$, defined as $\eta_h = \eta^{A(u_h^*)}, \eta^{A(u_h)}, \eta^{GM}, \eta^{GM_{red}}, \eta^{LM(A(u_h^*))}, \eta^{LM(A(u_h))}, \eta^{LM(u_h)}, \eta^{LM(u_h^*)}$ with $\boldsymbol{v} \in \boldsymbol{g}_h + \boldsymbol{H}_0^1(\Omega)$ from subsection 7.1. Print respective efficiency indices $\mathrm{EI} := \eta_h/\|\mathcal{A}\boldsymbol{\varepsilon}\|_0$.
(d) Mark elements $T \in \mathcal{T}_\ell$ for refinement using the bulk strategy so that the sum of all η_T^2 over the marked elements exceeds half of the sum over all elements in $\mathcal{T}_\ell$.
(e) Refine marked triangles and compute $\mathcal{T}_{\ell+1}$ using *red* refinement for uniform mesh and *newest vertex bisection* refinement for adaptive mesh. Update ℓ and go to (b).

7.3. Example with smooth solution. The exact solution on the unit square $\Omega := (0,1)^2$

$$\boldsymbol{u}(x,y) = \big(x(1-x)(1-2y),\ -y(1-y)(1-2x)\big), \quad \text{and} \quad \mathrm{p(x,y) = 2(y-x)}$$

specifies data $\boldsymbol{f}$ and $\boldsymbol{g}$ [19]. Figure 2 presents the convergence history of various errors as a function of the number of elements in the case of uniform mesh refinement

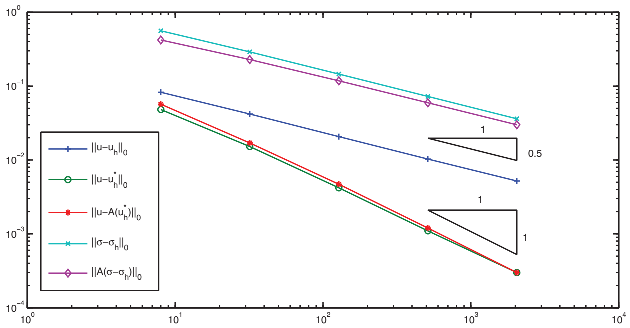
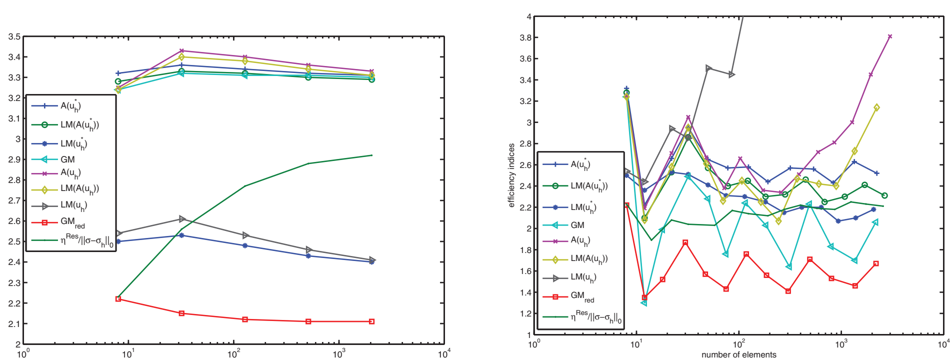


FIG. 2. *History of various errors for uniform meshes as a function of the number of elements in subsection* 7.3.

TABLE 1
Efficiency indices for uniform meshes in subsection 7.3.

NT	E	$\eta^{A(u_h^*)}$	EI	$\eta^{LM(A(u_h^*))}$	EI	$\eta^{LM(u_h^*)}$	EI	η^{GM}	EI
8	4.2031-1	1.3962	3.32	1.3788	3.28	1.0495	2.50	1.3612	3.24
32	2.2837-1	7.6678-1	3.36	7.6154-1	3.33	5.7754-1	2.53	7.5819-1	3.32
128	1.1778-1	3.9294-1	3.34	3.9055-1	3.32	2.9195-1	2.48	3.9022-1	3.31
512	5.9532-2	1.9759-1	3.32	1.9651-1	3.30	1.4479-1	2.43	1.9677-1	3.31
2048	2.9872-2	9.8859-2	3.31	9.8354-2	3.29	7.1747-2	2.40	9.8629-2	3.30
NT	E	$\eta^{A(u_h)}$	EI	$\eta^{LM(A(u_h))}$	EI	$\eta^{LM(u_h)}$	EI	$\eta^{GM_{red}}$	EI
8	4.2031-1	1.3679	3.25	1.3616	3.24	1.0659	2.54	9.3399-1	2.22
32	2.2837-1	7.8341-1	3.43	7.7753-1	3.40	5.9509-1	2.61	4.9045-1	2.15
128	1.1778-1	4.0038-1	3.40	3.9801-1	3.38	2.9854-1	2.53	2.4968-1	2.12
512	5.9532-2	1.9985-1	3.36	1.9881-1	3.34	1.4650-1	2.46	1.2561-1	2.11
2048	2.9872-2	9.9475-2	3.33	9.8987-2	3.31	7.2059-2	2.41	6.2932-2	2.11



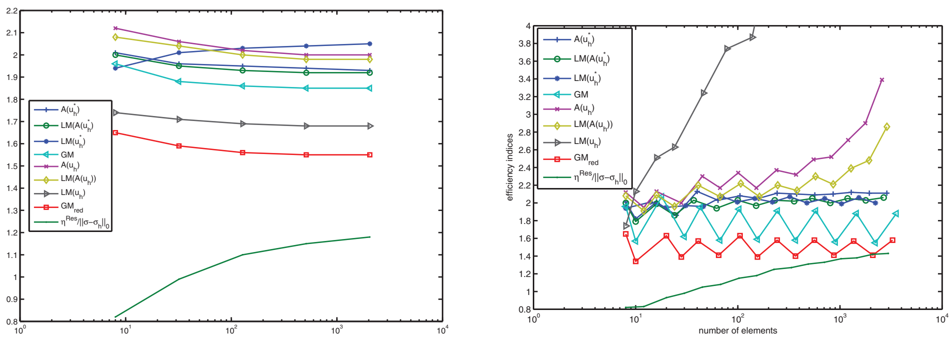
(a) (b)

FIG. 3. *History of efficiency indices* $\eta^{\boldsymbol{v}}/\|\mathcal{A}\boldsymbol{\sigma} - \mathcal{A}\boldsymbol{\sigma}_h\|_0$ *on the uniform mesh (left) and the adaptive mesh (right) for error estimators* $\eta^{\boldsymbol{v}}$ *from* (33) *for various* $\boldsymbol{v}$ *plus efficiency indices* $\eta^{Res}/\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_0$ *in subsection* 7.3.

which confirms the theoretical findings of this paper. Table 1 displays the efficiency indices EI:=$\eta/\|\mathcal{A}\boldsymbol{\sigma} - \mathcal{A}\boldsymbol{\sigma}_h\|_0$ obtained for eight different functions $\boldsymbol{v} \in \boldsymbol{g}_h + \boldsymbol{H}_0^1(\Omega)$ to the error estimator for uniform and adaptive mesh refinements. Figure 3 displays efficiency indices for various error estimators (27) based on various test functions $\boldsymbol{v}$ of

TABLE 2
Efficiency indices for uniform meshes in subsection 7.4.

NT	E	$\eta^{A(u_h^*)}$	EI	$\eta^{LM(A(u_h^*))}$	EI	$\eta^{LM(u_h^*)}$	EI	η^{GM}	EI
8	8.5332-1	1.7179	2.01	1.7068	2.00	1.6568	1.94	1.6683	1.96
32	6.6507-1	1.3063	1.96	1.2973	1.95	1.3395	2.01	1.2527	1.88
128	4.8606-1	9.4576-1	1.95	9.3916-1	1.93	9.8880-1	2.03	9.0455-1	1.86
512	3.4344-1	6.6561-1	1.94	6.6099-1	1.92	7.0110-1	2.04	6.3636-1	1.85
2048	2.3903-1	4.6239-1	1.93	4.5919-1	1.92	4.8883-1	2.05	4.4201-1	1.85
NT	E	$\eta^{A(u_h)}$	EI	$\eta^{LM(A(u_h))}$	EI	$\eta^{LM(u_h)}$	EI	$\eta^{GM_{red}}$	EI
8	8.5332-1	1.8100	2.12	1.7786	2.08	1.4832	1.74	1.4096	1.65
32	6.6507-1	1.3710	2.06	1.3550	2.04	1.1397	1.71	1.0556	1.59
128	4.8606-1	9.8213-1	2.02	9.7109-1	2.00	8.2245-1	1.69	7.5947-1	1.56
512	3.4344-1	6.8859-1	2.00	6.8104-1	1.98	5.7742-1	1.68	5.3318-1	1.55
2048	2.3903-1	4.7778-1	2.00	4.7259-1	1.98	4.0076-1	1.68	3.6996-1	1.55



(a) (b)

FIG. 4. *History of efficiency indices $\eta^{\boldsymbol{v}}/\|\mathcal{A}\boldsymbol{\sigma} - \mathcal{A}\boldsymbol{\sigma}_h\|_0$ on the uniform mesh (left) and the adaptive mesh (right) for error estimators $\eta^{\boldsymbol{v}}$ from* (33) *for various $\boldsymbol{v}$ plus efficiency indices $\eta^{Res}/\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_0$ in subsection* 7.4.

subsection 7.1 plus efficiency indices $\eta^{Res}/\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_0$ obtained from Theorem 5.4. Our other (undisplayed) numerical experiments for more examples convinced us that the efficiency indices $\eta^{Res}/\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_0$ lay in the range of 2.4–3.8 (uniform mesh) and 2–2.7 (adaptive mesh). Table 1 and Figure 3 underline the superiority of postprocessed solution $\boldsymbol{u}_h^*$ in the design of efficient error estimators. The benefits of adaptive versus uniform refinement are not significant in the case of a smooth solution.

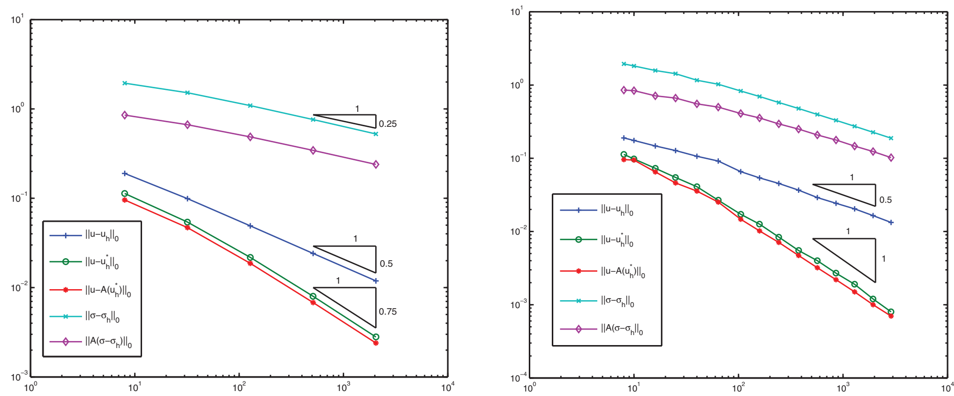
7.4. Example with singularity. The exact solution on $\Omega := (0,1)^2$ with $\boldsymbol{f} \equiv \boldsymbol{0}$, in polar coordinates $(r,\phi) \in (0,\infty)\times(0,\pi/2)$,

$$\boldsymbol{u}(r,\phi) = r^\alpha \left((\alpha+1)\sin(\phi)\psi(\phi) + \cos(\phi)\psi'(\phi), -(\alpha+1)\cos(\phi)\psi(\phi) + \sin(\phi)\psi'(\phi)\right),$$

$$p(r,\phi) = \frac{1}{\alpha-1} r^{\alpha-1}\left((\alpha+1)^2\psi'(\phi) + \psi'''(\phi)\right) \quad \text{with } \alpha = 856399/1572864 = 0.5444,$$

$$\psi(\phi) = \frac{\sin((\alpha+1)\phi)\cos(\alpha\omega) - \cos((\alpha+1)\phi)}{\alpha+1} - \frac{\sin((\alpha-1)\phi)\cos(\alpha\omega) + \cos((\alpha-1)\phi)}{\alpha-1}$$

for $\omega = 3\pi/2$, leads to efficiency indices displayed in Table 2 and Figure 4. Despite the lack of regularity, the postprocessed solution $\boldsymbol{u}_h^*$ leads to more efficient error control.



(a) (b)

FIG. 5. *History of various errors for uniform (left) and adaptive (right) meshes as a function of the number of elements in subsection* 7.4.

TABLE 3
History of various errors for adaptive meshes in subsection 7.4.

NT	$\Vert\boldsymbol{u}-\boldsymbol{u}_h\Vert_0$	$\Vert\boldsymbol{u}-\boldsymbol{u}_h^*\Vert_0$	$\Vert\boldsymbol{u}-A(\boldsymbol{u}_h^*)\Vert_0$	$\Vert\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\Vert_0$	$\Vert\mathcal{A}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\Vert_0$
8	1.8973-1	1.1299-1	9.5809-2	1.9421+0	8.5332-1
10	1.7523-1	9.8042-2	9.4563-2	1.8205+0	8.3778-1
16	1.4739-1	7.2931-2	6.5010-2	1.5774+0	7.1206-1
25	1.2741-1	5.4688-2	4.6016-2	1.4274+0	6.6380-1
40	1.0637-1	4.0912-2	3.5574-2	1.1636+0	5.5374-1
64	9.1551-2	2.6674-2	2.5208-2	1.0262+0	4.9978-1
105	6.5587-2	1.7200-2	1.4748-2	8.2858-1	4.0849-1
159	5.3960-2	1.2610-2	1.0241-2	6.9667-1	3.5563-1
243	4.5269-2	8.4151-3	7.1129-3	5.7729-1	2.9484-1
376	3.6733-2	5.5034-3	4.6795-3	4.7732-1	2.5041-1
568	2.9121-2	3.9621-3	3.1753-3	3.9636-1	2.0832-1
858	2.4312-2	2.7331-3	2.2152-3	3.2971-1	1.7756-1
1291	2.0350-2	1.8624-3	1.5459-3	2.7348-1	1.4671-1
1941	1.6537-2	1.2292-3	1.0193-3	2.2662-1	1.2373-1
2882	1.3251-2	8.2603-4	6.7257-4	1.8817-1	1.0208-1

In this example, the approximation of $\boldsymbol{g}$ by $\boldsymbol{g}_h$ appears problematic in the sense that it may not lead to negligible high-order perturbation. We display the convergence history of various errors as a function of the number of elements for uniform (left) and adaptive (right) meshes in Figure 5 (cf. Table 3). The superiority of adaptive mesh refinement can be clearly seen in Figure 5. Then the convergence order for $\Vert\boldsymbol{u}-\boldsymbol{u}_h^*\Vert_0$ from Theorem 4.1 is recovered even for the singular solution.

Although the convergence of the adaptive algorithm is not the focus of this paper, all the a posteriori error estimators have been employed for refinement indications in adaptive strategies. Figure 5 (right) displays the result based on the test function $\boldsymbol{v}$ from $A(\boldsymbol{u}_h^*)$ as a typical outcome of our other (undisplayed) numerical experiments.

In fact, all adaptive algorithms lead to significant improvements of the convergence rate to optimal order.

7.5. Conclusions. Higher-order convergence of $\Vert\boldsymbol{u}-\boldsymbol{u}_h^*\Vert_0$ proved in Theorem 4.1 is confirmed for the smooth solution in subsection 7.3 and is even observed for adaptive mesh refinement in subsection 7.4.

Guaranteed error control is feasible with efficiency indices between 1 and 3. The postprocessed solution $\boldsymbol{u}_h^*$ allows for accurate error control via $\eta^{LM(u_h^*)}$. Although the method and the error estimation in [19] are different, the efficiency indices in subsection 7.3 (resp., subsection 7.4) are about 3 and 2.5 for uniform and adaptive meshes (resp., 2 for both cases) and compare with those of [19] with 3.3 and 3.5 (resp., 3.0 and 3.8).

Reliability constant for the residual-based explicit a posteriori error estimator η^{Res} from Theorem 5.4 is not controlled and therefore η^{Res} does not lead to guaranteed error control.

The superiority of adaptive over uniform mesh refinement is clearly visible in the improved convergence rate for solutions of limited regularity.

REFERENCES

[1] M. Ainsworth, *A posteriori error estimation for discontinuous Galerkin finite element approximation*, SIAM J. Numer. Anal., 45 (2007), pp. 1777–1798.

[2] M. Ainsworth and J. T. Oden, *A Posteriori Error Estimation in Finite Element Analysis*, John Wiley & Sons, New York, 2000.

[3] D. N. Arnold and R. S. Falk, *A new mixed formulation for elasticity*, Numer. Math., 53 (1988), pp. 13–30.

[4] I. Babuska and T. Strouboulis, *The Finite Element Method and its Reliability*, Numer. Math. Sci. Comput., Oxford University Press, New York, 2001.

[5] D. Braess, *Finite Elements: Theory, Fast Solvers and Applications in Solid Mechanics*, Cambridge University Press, Cambridge, UK, 2007.

[6] S. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, Springer-Verlag, New York, 1994.

[7] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, New York, 1991.

[8] Z. Cai and G. Starke, *First-order system least squares for the stress-displacement formulation: Linear elasticity*, SIAM J. Numer. Anal., 41 (2003), pp. 715–730.

[9] 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.

[10] Z. Cai, C. Tong, P. S. Vassilevski, and C. Wang, *Mixed finite element methods for incompressible flow: Stationary Stokes equations*, Numer. Methods Partial Differential Equations, 26 (2010), pp. 957–978.

[11] Z. Cai and Y. Wang, *A multigrid method for the pseudostress formulation of Stokes problems*, SIAM J. Sci. Comput., 29 (2007), pp. 2078–2095.

[12] C. Carstensen, *A posteriori error estimate for the mixed finite element method*, Math. Comp., 66 (1997), pp. 465–476.

[13] C. Carstensen, *A unifying theory of a posteriori finite element error control*, Numer. Math., 100 (2005), pp. 617–637.

[14] C. Carstensen, G. Dolzmann, and S. Bartels, *Inhomogeneous Dirichlet conditions in a priori and a posteriori finite element error analysis*, Numer. Math., 99 (2004), pp. 1–24.

[15] C. Carstensen and S. A. Funken, *Constants in Clément-interpolation error and residual based a posteriori error estimates in finite element methods*, East-West J. Numer. Math., 8 (2000), pp. 153–175.

[16] C. Carstensen and S. A. Funken, *A posteriori error control in low-order finite element discretizations of incompressible stationary flow problems*, Math. Comp., 70 (2001), pp. 1353–1381.

[17] E. V. Chizhonkov and M. A. Olshanskii, *On the domain geometry dependence of the LBB condition*, RAIRO Model. Math. Anal. Numer., 34 (2000), pp. 935–951.

[18] E. Dari, R. Durán, and C. Padra, *Error estimators for nonconforming finite element approximations of the Stokes problem*, Math. Comp., 64 (1995), pp. 1017–1033.

[19] W. Döfler and M. Ainsworth, *Reliable a posteriori error control for nonconforming finite element approximation of Stokes flow*, Math. Comp., 74 (2005), pp. 1599–1619.

[20] M. Gerritsma and T. Phillips, *Compatible spectral approximations for the velocity-stress-pressure formulation of the Stokes problem*, SIAM J. Sci. Comput., 20 (1999), pp. 1530–1550.

[21] V. Girault and P. A. Raviart, *Finite Element Methods for Navier-Stokes Equations*, Springer-Verlag, New York, 1986.

[22] D. Kim and E.-J. Park, *A posteriori error estimator for expanded mixed hybrid methods*, Numer. Methods Partial Differential Equations, 23 (2007), pp. 330–349.

[23] D. Kim and E.-J. Park, *A posteriori error estimators for the upstream weighting mixed methods for convection diffusion problems*, Comput. Methods Appl. Mech. Engrg., 197 (2008), pp. 806–820.

[24] L. E. Payne and H. F. Weinberger, *An optimal Poincaré inequality for convex domains*, Arch. Rational Mech. Anal., 5 (1960), pp. 286–292.

[25] P. A. Raviart and J. Thomas, *A mixed finite element method for 2nd order elliptic problems*, in Mathematical Aspects of the Finite Elements Method, Lectures Notes in Math. 606, Springer, Berlin, 1977, pp. 292–315.

[26] R. Stenberg, *A family of mixed finite elements for the elasticity problem*, Numer. Math., 53 (1988), pp. 513–538.

[27] G. Stoyan, *Towards discrete Velte decompositions and narrow bounds for inf-sup constants*, Comput. Math. Appl., 38 (1999), pp. 243–261.

[28] R. Verfürth, *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*, Teubner-Wiley, Stuttgart, Germany, 1996.

[29] M. Vohralik, *Two types of guaranteed (and robust) a posteriori estimates for finite volume methods*, in Proceedings of Finite Volumes for Complex Applications V, R. Eymard and J.-M. Herard, eds., ISTE and John Wiley & Sons, London, Hoboken, NJ, 2008, pp. 649–656.