NUMERICAL EXPERIMENTS FOR THE ARNOLD-WINTHER MIXED FINITE ELEMENTS FOR THE STOKES PROBLEM*

CARSTEN CARSTENSEN[†], JOSCHA GEDICKE[‡], AND EUN-JAE PARK[§]

ABSTRACT. The stress-velocity formulation of the stationary Stokes problem allows an Arnold–Winther mixed finite element formulation with some superconvergent reconstruction of the velocity. This local postprocessing gives rise to two reliable a posteriori error estimators which recover optimal convergence order for the stress error estimates. The theoretical results are investigated in numerical benchmark examples.

1. INTRODUCTION

The stress-velocity-pressure formulation is the original physical model for incompressible Newtonian flows modeled by the conservation of momentum and the constitutive law. This model involves symmetric strain rates and stress tensors and is recast in a mixed form with symmetric stress tensors. The use of the deviatoric stress tensor $\mathcal{A}\sigma$ leads to the stress-velocity formulation for the Stokes problem

 $\operatorname{div}\boldsymbol{\sigma} = \boldsymbol{f} \quad \text{in } \Omega, \quad \mathcal{A}\boldsymbol{\sigma} - \boldsymbol{\epsilon}(\boldsymbol{u}) = 0 \quad \text{in } \Omega, \quad \text{and} \quad \boldsymbol{u} = \boldsymbol{g} \quad \text{on } \partial \Omega$

for a bounded Lipschitz domain $\Omega \subset \mathbb{R}^2$ and given data $\mathbf{f} \in L^2(\Omega; \mathbb{R}^2)$ and $\mathbf{g} \in H^1(\Omega; \mathbb{R}^2) \cap C(\overline{\Omega}; \mathbb{R}^2)$. The discretization is feasible with the symmetric Arnold–Winther mixed finite element method (MFEM) [AW02] proposed in linear elasticity. Since the Arnold–Winther MFEM has been proven to be stable for any material parameters, it is also a stable method for the Stokes problem as a limit case of linear elasticity.

²⁰⁰⁰ Mathematics Subject Classification. 65K10, 65M12, 65M60.

Key words and phrases. a posteriori error estimates, Arnold–Winther elements, mixed finite element approximations, Stokes problem, stress-velocity formulation.

^{*}This work was supported by the World Class University (WCU) program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science, and Technology (R31-10049).

[†] The research of this author was supported in part by the DFG Research Center MATHEON "Mathematics for Key Technologies" in Berlin.

[‡] The research of this author was supported in part by MATHEON and the Berlin Mathematical School (BMS).

 $^{^{\}S}$ The research of this author was supported in part by the NRF Joint Research Program (NRF-2010-0013374 and NRF-2011-0030934).

Final version published in SIAM J. Sci. Comput., Vol. 34, No. 4, pp. A2267–2287, 2012, DOI 10.1137/100802906.

Suppose $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h)$ is the mixed Arnold–Winther MFEM approximation of order $k \geq 1$ to sufficiently smooth $(\boldsymbol{\sigma}, \boldsymbol{u})$. Then it holds in terms of the standard Sobolev norms $\|\cdot\|_s := \|\cdot\|_{H^s(\Omega)}, m \in \mathbb{N}$,

$$\begin{aligned} \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_0 &\lesssim h^m \|\boldsymbol{\sigma}\|_m, \quad 1 \leq m \leq k+2, \\ \|\mathbf{div}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_0 &\lesssim h^m \|\mathbf{div}\boldsymbol{\sigma}\|_m, \quad 0 \leq m \leq k+1, \\ \|\boldsymbol{u} - \boldsymbol{u}_h\|_0 &\lesssim h^m \|\boldsymbol{u}\|_{m+1}, \quad 1 \leq m \leq k+1. \end{aligned}$$

Compared to the stress errors, the last bound appears suboptimal, but is a consequence of the lower ansatz for the displacement variable reduced by two degrees when compared with the stress variable instead of only one. Here and throughout the paper, for short notation on generic constants C, for any two real numbers or functions or expressions Aand B, $A \leq B$ abbreviates $A \leq CB$. The point is that this multiplicative constant C does not depend on the local or global mesh-sizes but may depend on the domain Ω , the shape regularity of the mesh, and the polynomial degree. Similarly, $A \approx B$ abbreviates $A \leq B \leq A$. The nonstandard finite element method (FEM) for the Stokes problem started with [BW91, DDP95, GR86] and the reader is referred to [AO00, BS01, Ver96] for information on a posteriori error control. The main results of this paper concern the a priori error estimation of the superconvergence of some reconstructed velocity field u_h^* in the sense of

$$\|m{u} - m{u}_h^*\|_0 \lesssim h^{k+3} (\|m{\sigma}\|_{k+2} + \| ext{div}m{\sigma}\|_{k+1}) + h^{m+1} \|m{u}\|_{m+1}$$

for the restricted class of domains Ω with sufficiently smooth boundary $\partial \Omega$. We refer the reader to [Kim07, LM08, Voh10] for some postprocessing for the Poisson problem from Stenberg [Ste88]. Based on the discontinuous postprocessed velocity field \boldsymbol{u}_h^* and a smooth velocity field $\tilde{\boldsymbol{u}}_h \in H^1(\Omega; \mathbb{R}^2)$ we design reliable and efficient a posteriori error estimators μ and η such that reliability or even equivalence holds in the sense of

$$\begin{aligned} \|\mathcal{A}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\|_0 + \|\boldsymbol{\epsilon}(\boldsymbol{u}-\tilde{\boldsymbol{u}}_h)\|_0 + \operatorname{osc}(\boldsymbol{f},\mathcal{T}_h) &\approx \mu + \operatorname{osc}(\boldsymbol{f},\mathcal{T}_h), \\ \|\mathcal{A}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\|_0 + \|\boldsymbol{\epsilon}_{\mathcal{T}_h}(\boldsymbol{u}-\boldsymbol{u}_h^*)\|_0 + \operatorname{osc}(\boldsymbol{f},\mathcal{T}_h) \lesssim \eta + \operatorname{osc}(\boldsymbol{f},\mathcal{T}_h) \end{aligned}$$

up to oscillations

$$\mathrm{osc}^2(oldsymbol{f},\mathcal{T}_h) := \sum_{T\in\mathcal{T}_h} h_T^2 \|oldsymbol{f}-oldsymbol{f}_h\|_{0,T}^2$$

of the right-hand side \boldsymbol{f} and its piecewise L^2 projection \boldsymbol{f}_h onto piecewise polynomials of degree $\leq k$. Note that for the Arnold–Winther FEM, the oscillations $\operatorname{osc}(\boldsymbol{f}, \mathcal{T}_h)$ of the right-hand side \boldsymbol{f} are of the same but not of higher order compared to the stress and strain errors. In principle, the oscillations might dominate the error estimator and therefore lead to an overestimation of the stress error $\|\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_0$, which is empirically confirmed in Section 7 by an academic example

with smooth solution. However, in all practical relevant benchmark examples of this paper, oscillations vanish for the constant f.

The remainder of this paper is organized as follows: Section 2 introduces necessary notation and the stress-velocity formulation. Section 3 recalls the discrete problem and its mixed Arnold–Winther finite element approximation. Section 4 establishes some superconvergent local postprocessing of the velocity for all fixed polynomial degrees. The a posteriori error analysis for the two reliable error estimators η and μ together with the efficiency of μ is presented in Section 5. Section 6 describes some adaptive finite element method (AFEM), and Section 7 presents various numerical examples to verify the theoretical results and to illustrate the performance of the method. It turns out that AFEM is very important to meet optimal convergence rates by proper mesh-design to compensate for corner singularities.

2. Stress Velocity Formulation

For $\boldsymbol{v} = (v_1, v_2)^t \in \mathbb{R}^2$, $\boldsymbol{\tau} = (\tau_{ij})_{2 \times 2}$, and $\boldsymbol{\sigma} = (\sigma_{ij})_{2 \times 2} \in \mathbb{R}^{2 \times 2}$, we define

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

We employ the standard notation for the Sobolev spaces $H^s(\omega)$ for $s \geq 0$. The associated norm is denoted by $\|\cdot\|_{s,\omega}$. For s = 0, we use the notation $L^2(\omega)$ instead of $H^0(\omega)$. In the case $\omega = \Omega$ we simply write $\|\cdot\|_{s,\Omega} = \|\cdot\|_s$. We define $H^{-s}(\omega) := (H^s_0(\omega))^*$ as the dual space of $H^s_0(\omega)$. Extending the definitions to vector- and matrix-valued functions, we let $H^s(\omega; \mathbb{R}^2)$ (simply $H^s(\omega)$) and $H^s(\omega; \mathbb{R}^{2\times 2})$ denote the Sobolev spaces over the set of 2-dimensional vector- and 2×2 matrix-valued functions, respectively. Finally, we define the space

$$oldsymbol{H}(\mathbf{div},\Omega;\mathbb{R}^{2 imes 2}):=\left\{oldsymbol{ au}\inoldsymbol{L}^2(\Omega;\mathbb{R}^{2 imes 2})\;\Big|\;\mathbf{div}oldsymbol{ au}\inoldsymbol{L}^2(\Omega)
ight\}$$

with the norm

$$\|\boldsymbol{\tau}\|_{\boldsymbol{H}(\operatorname{\mathbf{div}},\Omega;\mathbb{R}^{2 imes 2})}^2 := (\boldsymbol{ au}, \boldsymbol{ au}) + (\operatorname{\mathbf{div}} \boldsymbol{ au}, \operatorname{\mathbf{div}} \boldsymbol{ au}).$$

Here and throughout the paper, $(\cdot, \cdot)_{\omega}$ denotes the $L^2(\omega; \mathbb{R}^{2\times 2})$ inner product $\int_{\omega} \boldsymbol{\tau} : \boldsymbol{\tau} \, dx$ as well as the $L^2(\omega)$ inner product $\int_{\omega} \boldsymbol{\tau} \cdot \boldsymbol{\tau} \, dx$. In

the case $\omega = \Omega$ we simply write $(\cdot, \cdot)_{\Omega} = (\cdot, \cdot)$. The extended $L^2(\partial\Omega)$ product along the boundary $\partial\Omega$ is denoted by the duality brackets $\langle \cdot, \cdot \rangle$.

On the domain $\Omega \subset \mathbb{R}^2$ with sufficiently smooth Lipschitz boundary $\partial\Omega$ filled with a fluid of viscosity $\nu > 0$ and given data $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$ and $\boldsymbol{g} \in \boldsymbol{H}^1(\Omega) \cap \boldsymbol{C}(\overline{\Omega})$, the stationary Stokes problem reads

(1)
$$\begin{aligned} -\nu \Delta \boldsymbol{u} + \nabla p &= -\boldsymbol{f} \quad \text{in } \Omega, \\ \text{div } \boldsymbol{u} &= 0 \quad \text{in } \Omega, \\ \boldsymbol{u} &= \boldsymbol{g} \quad \text{on } \partial \Omega \end{aligned}$$

for the unknown velocity \boldsymbol{u} and pressure p. Suppose that the following two compatibility conditions hold:

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

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

$$oldsymbol{\epsilon}(oldsymbol{u}) := rac{1}{2}ig(
abla oldsymbol{u} + (
abla oldsymbol{u})^tig)$$

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

$$div \sigma = f \text{ in } \Omega,$$

$$\sigma + p\delta - 2\nu\epsilon(u) = 0 \text{ in } \Omega,$$

$$div u = 0 \text{ in } \Omega,$$

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

To design the stress-velocity formulation we define the deviatoric operator

 $\mathcal{A}:\mathbb{S}\to\mathbb{S}$

for the symmetric tensors $\mathbb{S} := \{ \boldsymbol{\tau} \in \mathbb{R}^{2 \times 2} | \boldsymbol{\tau} = \boldsymbol{\tau}^t \}$ by

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

Note that $\operatorname{Ker}(\mathcal{A}) = \{q\delta \in \mathbb{S} \mid q \in \mathbb{R}\}\$ and $\mathcal{A}\tau$ is a trace-free tensor called deviatoric. Further, we can easily show that the following properties of the operator \mathcal{A} hold, for all $\tau, \sigma \in \mathbb{S}$:

$$egin{aligned} & (\mathcal{A}m{ au},m{\sigma}) = (m{ au},\mathcal{A}m{\sigma}), \ & (\mathcal{A}m{ au},2
u\mathcal{A}m{\sigma}) = (\mathcal{A}m{\sigma},m{ au}) = rac{1}{2
u}\left((m{\sigma},m{ au}) - rac{1}{2}(\mathrm{tr}\,m{\sigma},\mathrm{tr}\,m{ au})
ight), \ & & \|\mathcal{A}m{ au}\| \leq rac{1}{2
u}\|m{ au}\|. \end{aligned}$$

Using the deviatoric operator, we arrive at the stress-velocity formulation for the Stokes problem (1):

(2)
$$\begin{aligned} \operatorname{div} \boldsymbol{\sigma} &= \boldsymbol{f} \quad \text{in } \Omega, \\ \mathcal{A} \boldsymbol{\sigma} &- \boldsymbol{\epsilon}(\boldsymbol{u}) &= \boldsymbol{0} \quad \text{in } \Omega, \end{aligned}$$

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

The second equation of (2) is obtained from

$$\operatorname{tr} \boldsymbol{\epsilon}(\boldsymbol{u}) = \operatorname{div} \boldsymbol{u} = 0 \quad \text{and} \quad \operatorname{tr} \boldsymbol{\sigma} = -2p$$

and the compatibility condition $\int_{\Omega} p \, dx = 0$ implies

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

We have the following well-known regularity results for sufficiently smooth boundary $\partial\Omega$ or a convex domain. For $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega), \boldsymbol{g} \in \boldsymbol{H}^2(\Omega)$, the solutions to problems (1) and (2) satisfy $\boldsymbol{u} \in \boldsymbol{H}^2(\Omega) \cap \boldsymbol{H}^1(\overline{\Omega})$, $p \in H^1(\Omega)/\mathbb{R}, \boldsymbol{\sigma} \in \boldsymbol{H}^1(\Omega; \mathbb{S})$, and

(3)
$$\|\boldsymbol{u}\|_{2} + \|p\|_{1} + \|\boldsymbol{\sigma}\|_{1} \lesssim \|\boldsymbol{f}\|_{0} + \|\boldsymbol{g}\|_{2}.$$

With $\boldsymbol{V} := \boldsymbol{L}^2(\Omega)$ and

$$\boldsymbol{\Phi} := \boldsymbol{H}(\operatorname{\mathbf{div}}, \Omega, \mathbb{S}) / \mathbb{R} \sim \Big\{ \boldsymbol{\tau} \in \boldsymbol{H}(\operatorname{\mathbf{div}}, \Omega, \mathbb{S}) \mid \int_{\Omega} \operatorname{tr} \boldsymbol{\tau} dx = 0 \Big\},$$

the weak form for the problem (2) reads as follows: Find $\sigma \in \Phi$ and $u \in V$ such that

(4)
$$(\mathcal{A}\boldsymbol{\sigma},\boldsymbol{\tau}) + (\operatorname{div}\boldsymbol{\tau},\boldsymbol{u}) = \langle \boldsymbol{g},\boldsymbol{\tau}\boldsymbol{n} \rangle \text{ for all } \boldsymbol{\tau} \in \boldsymbol{\Phi},$$

(5)
$$(\operatorname{div}\boldsymbol{\sigma}, \boldsymbol{v}) = (\boldsymbol{f}, \boldsymbol{v}) \text{ for all } \boldsymbol{v} \in \boldsymbol{V}.$$

This problem has a unique solution from the well-known inf-sup condition in the mixed formulation and the following lemma [BF91].

Lemma 2.1. For all $\tau \in \Phi$, we have

$$\|oldsymbol{ au}\|_0^2 \lesssim \|\mathcal{A}^{1/2}oldsymbol{ au}\|_0^2 + \| ext{div}oldsymbol{ au}\|_{-1}^2$$

3. Mixed finite element method

Let $\{\mathcal{T}_h\}$ be a family of quasi-uniform triangulations of $\overline{\Omega}$ by triangles T of diameter h_T . For each \mathcal{T}_h , let \mathcal{E}_h denote the set of all edges of \mathcal{T}_h and, given $T \in \mathcal{T}_h$, let $\mathcal{E}(T)$ be the set of its edges. Further, for an edge $E \in \mathcal{E}(T)$, let $\mathbf{t}_E = (-n_2, n_1)^t$ be the unit tangential vector along E for the unit outward normal $\mathbf{n}_E = (n_1, n_2)^t$ to E with the diameter h_E . Moreover, we define the jump $[\mathbf{w}]$ of \mathbf{w} by

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

where \boldsymbol{n}_E points from T_+ into its neighboring element T_- , and $[\boldsymbol{w}]|_E := \boldsymbol{w} - \boldsymbol{g}$ if $E = \overline{T}_+ \cap \partial \Omega$.

We define the finite element spaces associated with the regular triangulation \mathcal{T}_h of Ω into triangles,

$$AW_{k}(T) := \left\{ \boldsymbol{\tau} \in P_{k+2}(T; \mathbb{S}) \, \middle| \, \operatorname{\mathbf{div}} \boldsymbol{\tau} \in P_{k}(T; \mathbb{R}^{2}) \right\}, \\ \boldsymbol{\Phi}_{h} := \left\{ \boldsymbol{\tau} \in \boldsymbol{\Phi} \, \middle| \, \boldsymbol{\tau}|_{T} \in AW_{k}(T) \right\}, \\ \boldsymbol{V}_{h} := \left\{ \boldsymbol{v} \in \boldsymbol{L}^{2}(\Omega) \, \middle| \, v|_{T} \in P_{k}(T; \mathbb{R}^{2}) \right\},$$

where $AW_k(T)$ is the Arnold–Winther element of index $k \geq 1$ of [AW02], and $P_k(T)$ is the set of polynomials of total degree k on the domain T. The space Φ_h consists of all symmetric polynomial matrix fields of degree at most k + 1 together with the divergence-free matrix fields of degree k + 2.

We notice that $\Phi_h \subset \Phi$ and hence if $\tau_h \in \Phi_h$, then τ_h has continuous normal components and the constraint $\int_{\Omega} \operatorname{tr} \tau_h dx = 0$ holds.

The MFEM reads as follows: Find $\sigma_h \in \Phi_h$ and $u_h \in V_h$ such that

(6)
$$(\mathcal{A}\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) + (\operatorname{div}\boldsymbol{\tau}_h, \boldsymbol{u}_h) = \langle \boldsymbol{g}, \boldsymbol{\tau}_h \boldsymbol{n} \rangle$$
 for all $\boldsymbol{\tau}_h \in \boldsymbol{\Phi}_h$,

(7)
$$(\operatorname{div}\boldsymbol{\sigma}_h, \boldsymbol{v}_h) = (\boldsymbol{f}, \boldsymbol{v}_h) \text{ for all } \boldsymbol{v}_h \in \boldsymbol{V}_h.$$

By Lemma 2.1 and the discrete inf-sup condition of the AW_k element space (cf. [BF91]), the discrete problem is well-posed and has a unique solution.

We consider a projection operator over the space Φ . Let Π_h [AW02] denote the Arnold–Winther projection operator associated with the degrees of freedom onto $\Phi_h + span\{\delta\}$. We define $\Pi_h : \Phi \to \Phi_h$ by

(8)
$$\Pi_{h}\boldsymbol{\tau} = \tilde{\Pi}_{h}\boldsymbol{\tau} - \frac{\int_{\Omega} (\operatorname{tr}\Pi_{h}\boldsymbol{\tau}) dx}{2|\Omega|} \boldsymbol{\delta} \quad \text{for all } \boldsymbol{\tau} \in \boldsymbol{\Phi}$$

with the area $|\Omega| = \int_{\Omega} 1 \, dx$ of Ω . We notice that $\int_{\Omega} (\text{tr} \Pi_h \boldsymbol{\tau}) dx = 0$. Let \boldsymbol{P}_h be the \boldsymbol{L}^2 projection onto \boldsymbol{V}_h with the well-known approximation property

(9)
$$\|\boldsymbol{P}_{h}\boldsymbol{v}-\boldsymbol{v}\|_{0} \lesssim h^{k+1}\|\boldsymbol{v}\|_{k+1} \text{ for all } \boldsymbol{v} \in \boldsymbol{H}^{k+1}(\Omega).$$

Then the following two lemmas hold [AW02].

Lemma 3.1. The commutative property $\operatorname{div} \Pi_h = \boldsymbol{P}_h \operatorname{div} holds$. Furthermore, for $\boldsymbol{\tau} \in \boldsymbol{\Phi} \cap \boldsymbol{H}^{k+2}(\Omega; \mathbb{S}), m \in \mathbb{N}$, it holds that

- (10) $\|\boldsymbol{\tau} \Pi_h \boldsymbol{\tau}\|_0 \lesssim h^m \|\boldsymbol{\tau}\|_m, \quad 1 \le m \le k+2,$
- (11) $\|\mathbf{div}\boldsymbol{\tau} \mathbf{div}(\Pi_h\boldsymbol{\tau})\|_0 \lesssim h^m \|\mathbf{div}\boldsymbol{\tau}\|_m \quad 0 \le m \le k+1.$

Lemma 3.2. For the exact solution $(\boldsymbol{\sigma}, \boldsymbol{u}) \in (\boldsymbol{\Phi} \cap \boldsymbol{H}^{k+2}(\Omega; \mathbb{S})) \times \boldsymbol{H}^{k+2}(\Omega)$ of problem (1) and the approximate solution $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h)$ of problem (6)–(7), $m \in \mathbb{N}$, it holds that

$$\begin{aligned} \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_0 &\lesssim h^m \|\boldsymbol{\sigma}\|_m, \quad 1 \leq m \leq k+2, \\ \|\mathbf{div}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_0 &\lesssim h^m \|\mathbf{div}\boldsymbol{\sigma}\|_m, \quad 0 \leq m \leq k+1, \\ \|\boldsymbol{u} - \boldsymbol{u}_h\|_0 &\lesssim h^m \|\boldsymbol{u}\|_{m+1}, \quad 1 \leq m \leq k+1. \end{aligned}$$

Remark 3.3. We note that using the relation $p = -\text{tr} \sigma/2$, we can define the approximation of the pressure by $p_h := -\text{tr} \sigma_h/2$. Then the pressure error estimate holds:

$$\|p - p_h\|_0 = \frac{1}{2} \|\operatorname{tr} \boldsymbol{\sigma} - \operatorname{tr} \boldsymbol{\sigma}_h\|_0 \le \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_0 \lesssim h^m \|\boldsymbol{\sigma}\|_m, \ 1 \le m \le k+2.$$

The estimate for $\|\boldsymbol{P}_{h}\boldsymbol{u} - \boldsymbol{u}_{h}\|_{0}$ presented in the following theorem is used to derive the error estimates of the postprocessed velocity.

Theorem 3.4. With sufficiently smooth boundary $\partial\Omega$, $\boldsymbol{\sigma} \in \boldsymbol{H}^{k+2}(\Omega; \mathbb{S})$ and $\boldsymbol{f} = \operatorname{div} \boldsymbol{\sigma} \in \boldsymbol{H}^{k+1}(\Omega)$, it holds that

(12)
$$\|\boldsymbol{P}_{h}\boldsymbol{u}-\boldsymbol{u}_{h}\|_{0} \lesssim h^{k+3}\left(\|\boldsymbol{\sigma}\|_{k+2}+\|\mathbf{div}\boldsymbol{\sigma}\|_{k+1}\right).$$

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

(13)
$$(\mathcal{A}\boldsymbol{\eta},\boldsymbol{\tau}) + (\operatorname{div}\boldsymbol{\tau},\boldsymbol{z}) = 0 \text{ for all } \boldsymbol{\tau} \in \boldsymbol{\Phi},$$

(14)
$$(\operatorname{div}\boldsymbol{\eta}, \boldsymbol{v}) = (\boldsymbol{P}_h \boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{v}) \text{ for all } \boldsymbol{v} \in \boldsymbol{V}$$

The a priori estimate (3) implies

(15)
$$\|\boldsymbol{z}\|_2 \lesssim \|\boldsymbol{P}_h \boldsymbol{u} - \boldsymbol{u}_h\|_0$$
 and $\|\boldsymbol{\eta}\|_1 \lesssim \|\boldsymbol{P}_h \boldsymbol{u} - \boldsymbol{u}_h\|_0$.

Since (14) and $\mathbf{div}\Pi_h = \mathbf{P}_h \mathbf{div}$, we deduce

(16)
$$\|\boldsymbol{P}_{h}\boldsymbol{u} - \boldsymbol{u}_{h}\|_{0}^{2} = (\boldsymbol{P}_{h}\boldsymbol{u} - \boldsymbol{u}_{h}, \operatorname{div}\boldsymbol{\eta})$$
$$= (\boldsymbol{P}_{h}\boldsymbol{u} - \boldsymbol{u}_{h}, \boldsymbol{P}_{h}\operatorname{div}\boldsymbol{\eta})$$
$$= (\boldsymbol{u} - \boldsymbol{u}_{h}, \operatorname{div}\Pi_{h}\boldsymbol{\eta}).$$

The difference of (4)–(5) and (6)–(7) leads to

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

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

The identities (13), (16)–(18) and the estimates (9)–(10) yield

$$\|oldsymbol{P}_holdsymbol{u}-oldsymbol{u}_h\|_0^2=-(\mathcal{A}(oldsymbol{\sigma}-oldsymbol{\sigma}_h),\Pi_holdsymbol{\eta}-oldsymbol{\eta})-(oldsymbol{\sigma}-oldsymbol{\sigma}_h,\mathcal{A}oldsymbol{\eta})$$

(19)
$$= -(\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h), \Pi_h \boldsymbol{\eta} - \boldsymbol{\eta}) + (\operatorname{div}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h), \boldsymbol{z} - \boldsymbol{P}_h \boldsymbol{z}) \\ \lesssim h \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_0 \|\boldsymbol{\eta}\|_1 + h^2 \|\operatorname{div}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_0 \|\boldsymbol{z}\|_2.$$

Lemma 3.2 and the inequalities (15) and (19) lead to

$$\|\boldsymbol{P}_{h}\boldsymbol{u}-\boldsymbol{u}_{h}\|_{0}\lesssim h^{k+3}(\|\boldsymbol{\sigma}\|_{k+2}+\|\mathbf{div}\boldsymbol{\sigma}\|_{k+1}).$$

4. Postprocessing

Since $\mathcal{A}\boldsymbol{\sigma}_h$ is expected to be a good approximation of $\boldsymbol{\epsilon}(\boldsymbol{u})$, we can obtain an improved approximate solution of the velocity \boldsymbol{u} through local postprocessing in the spirit of Stenberg [Ste88]. For $m \geq k+2$ let

$$\boldsymbol{W}_h^* = \{ \boldsymbol{v} \in \boldsymbol{L}^2(\Omega) \mid \boldsymbol{v}|_T \in P_m(T; \mathbb{R}^2) \text{ for all } T \in \mathcal{T}_h \}.$$

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

$$(20) \boldsymbol{P}_T \boldsymbol{u}_h^* = \boldsymbol{u}_h,$$

(21) $(\boldsymbol{\epsilon}(\boldsymbol{u}_h^*), \boldsymbol{\epsilon}(\boldsymbol{v}))_T = (\mathcal{A}\boldsymbol{\sigma}_h, \boldsymbol{\epsilon}(\boldsymbol{v}))_T \text{ for all } \boldsymbol{v} \in (\boldsymbol{\delta} - \boldsymbol{P}_T)\boldsymbol{W}_h^*|_T.$

In other words, $\boldsymbol{u}_h^*|_T \in P_m(T; \mathbb{R}^2)$ is the Riesz representation of the linear functional $(\mathcal{A}\boldsymbol{\sigma}_h, \boldsymbol{\epsilon}(\cdot))_T$ in the Hilbert space $(\boldsymbol{\delta} - \boldsymbol{P}_T)\boldsymbol{W}_h^*|_T \equiv \{\boldsymbol{v}_m \in P_m(T; \mathbb{R}^2) \mid (\boldsymbol{v}_m, \boldsymbol{w}_k)_T = 0 \text{ for all } \boldsymbol{w}_k \in P_k(T; \mathbb{R}^2) \}$ with scalar product $(\boldsymbol{\epsilon}(\cdot), \boldsymbol{\epsilon}(\cdot))_T$. The postpocessing on each triangle with Lagrange multiplier $\boldsymbol{\lambda}_k \in P_k(T; \mathbb{R}^2)$ can be implemented as the linear system of equations

$$(\boldsymbol{\epsilon}(\boldsymbol{u}_h^*), \boldsymbol{\epsilon}(\boldsymbol{v}_m))_T + (\boldsymbol{\lambda}_k, \boldsymbol{v}_m)_T = (\mathcal{A}\boldsymbol{\sigma}_h, \boldsymbol{\epsilon}(\boldsymbol{v}_m))_T \quad \text{for all } \boldsymbol{v}_m \in P_m(T; \mathbb{R}^2),$$
$$(\boldsymbol{u}_h^*, \boldsymbol{w}_k)_T = (\boldsymbol{u}_h, \boldsymbol{w}_k)_T \quad \text{for all } \boldsymbol{w}_k \in P_k(T; \mathbb{R}^2).$$

The Korn inequality yields positive definiteness of $(\boldsymbol{\epsilon}(\cdot), \boldsymbol{\epsilon}(\cdot))_T$ on $(\boldsymbol{\delta} - \boldsymbol{P}_T) \boldsymbol{W}_h^*|_T$. Since $P_k(T; \mathbb{R}^2) \subset P_m(T; \mathbb{R}^2)$,

$$\sup_{\boldsymbol{v}_m \in P_m(T;\mathbb{R}^2)} (\boldsymbol{v}_m, \boldsymbol{\lambda}_k)_T / \|\boldsymbol{v}_m\|_{0,T} \ge \|\boldsymbol{\lambda}_k\|_{0,T} \quad \text{for all } \boldsymbol{\lambda}_k \in P_k(T;\mathbb{R}^2).$$

Thus, the Brezzi splitting theorem [Bre74] shows that there exists a unique solution on each triangle. The identity $\mathcal{A}\sigma = \epsilon(u)$ and (21) imply the error identity

(22)
$$(\boldsymbol{\epsilon}(\boldsymbol{u}-\boldsymbol{u}_h^*),\boldsymbol{\epsilon}(\boldsymbol{v}))_T = (\mathcal{A}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),\boldsymbol{\epsilon}(\boldsymbol{v}))_T, \forall \boldsymbol{v} \in (\boldsymbol{\delta}-\boldsymbol{P}_T)\boldsymbol{W}_h^*|_T.$$

Theorem 4.1. Let the boundary $\partial\Omega$ be sufficiently smooth, $\boldsymbol{u} \in \boldsymbol{H}^{m+1}(\Omega)$, $\boldsymbol{\sigma} \in \boldsymbol{H}^{k+2}(\Omega; \mathbb{S})$ and $\boldsymbol{f} = \operatorname{div} \boldsymbol{\sigma} \in \boldsymbol{H}^{k+1}(\Omega)$ solve (1). Then it holds that

$$egin{aligned} \|m{u}-m{u}_h^*\|_0 &\lesssim h^{k+3}(\|m{\sigma}\|_{k+2}+\| ext{div}m{\sigma}\|_{k+1})+h^{m+1}\|m{u}\|_{m+1},\ \|
abla_{\mathcal{T}_h}(m{u}-m{u}_h^*)\|_0 &\lesssim h^{k+2}(\|m{\sigma}\|_{k+2}+\| ext{div}m{\sigma}\|_{k+1})+h^m\|m{u}\|_{m+1}, \end{aligned}$$

for the postprocessed velocity $\boldsymbol{u}_h^* \in \boldsymbol{W}_h^*$ and the piecewise gradient $(\nabla_{\mathcal{T}_h} \cdot)|_T := \nabla(\cdot|_T).$

Proof. Let \hat{u} be the L^2 projection of u onto W_h^* . The triangle inequality shows

(23)
$$\|\boldsymbol{u} - \boldsymbol{u}_h^*\|_0 \le \|\boldsymbol{u} - \hat{\boldsymbol{u}}\|_0 + \|\boldsymbol{P}_h(\hat{\boldsymbol{u}} - \boldsymbol{u}_h^*)\|_0 + \|(\boldsymbol{\delta} - \boldsymbol{P}_h)(\hat{\boldsymbol{u}} - \boldsymbol{u}_h^*)\|_0.$$

Since $\hat{\boldsymbol{u}}$ is the L^2 -projection of \boldsymbol{u} onto \boldsymbol{W}_h^* , the a priori estimate (9) shows for the first term on the right-hand side of (23)

$$\|\boldsymbol{u} - \hat{\boldsymbol{u}}\|_0 \lesssim h^{m+1} \|\boldsymbol{u}\|_{m+1}$$
 for all $\boldsymbol{u} \in \boldsymbol{H}^{m+1}(\Omega)$.

For the second term, notice that $\boldsymbol{P}_T \boldsymbol{u}_h^* = \boldsymbol{u}_h$ on each T implies $\|\boldsymbol{P}_T(\hat{\boldsymbol{u}} - \boldsymbol{u}_h^*)\|_{0,T} = \|\boldsymbol{P}_T \hat{\boldsymbol{u}} - \boldsymbol{u}_h\|_{0,T}$. Since $\boldsymbol{V}_h \subset \boldsymbol{W}_h^*$, Theorem 3.4 shows that (24) $\|\boldsymbol{P}_h \hat{\boldsymbol{u}} - \boldsymbol{u}_h\|_0 = \|\boldsymbol{P}_h \boldsymbol{u} - \boldsymbol{u}_h\|_0 \lesssim h^{k+3} (\|\boldsymbol{\sigma}\|_{k+2} + \|\operatorname{div}\boldsymbol{\sigma}\|_{k+1}).$ In order to bound the third term on the right-hand side of (23), define $\boldsymbol{v} \in \boldsymbol{W}_h^*$ by $\boldsymbol{v}|_T = (\boldsymbol{\delta} - \boldsymbol{P}_T)(\hat{\boldsymbol{u}} - \boldsymbol{u}_h^*)$ for each $T \in \mathcal{T}_h$. Since $\boldsymbol{v}|_T \perp P_0(T; \mathbb{R}^2)$, the Poincaré inequality yields

$$\|oldsymbol{v}\|_{0,T} \leq rac{h_T}{\pi} \|
abla oldsymbol{v}\|_{0,T}.$$

Since $(\boldsymbol{\delta} - \boldsymbol{P}_T)\boldsymbol{w} = 0$ for all $\boldsymbol{w} \in P_1(T; \mathbb{R}^2)$, $\boldsymbol{v}|_T \perp \mathcal{RM} := \{\boldsymbol{v} = \boldsymbol{c} + b(x_2, -x_1)^t, \boldsymbol{c} \in \mathbb{R}^2, b \in \mathbb{R}\}$. Thus the second Korn inequality [BS94] leads to

$$\|\boldsymbol{v}\|_{0,T} \lesssim h_T \|\boldsymbol{\epsilon}(\boldsymbol{v})\|_{0,T}.$$

Then (22) and the Cauchy inequality yield

$$egin{aligned} \|oldsymbol{\epsilon}(oldsymbol{v})\|_{0,T}^2 &= (oldsymbol{\epsilon}(\hat{oldsymbol{u}}-oldsymbol{u}_h^*),oldsymbol{\epsilon}(oldsymbol{v}))_T - (oldsymbol{\epsilon}(oldsymbol{u}-oldsymbol{u}_h),oldsymbol{\epsilon}(oldsymbol{v}))_T + (\mathcal{A}(oldsymbol{\sigma}-oldsymbol{\sigma}_h),oldsymbol{\epsilon}(oldsymbol{v}))_T \ &- (oldsymbol{\epsilon}(oldsymbol{P}_T),oldsymbol{\epsilon}(oldsymbol{v}))_T \ &\leq \|oldsymbol{\epsilon}(oldsymbol{u}-oldsymbol{u}_h)\|_{0,T}\|oldsymbol{\epsilon}(oldsymbol{v}))_T + \|\mathcal{A}(oldsymbol{\sigma}-oldsymbol{\sigma}_h)\|_{0,T}\|oldsymbol{\epsilon}(oldsymbol{v}))_T \ &\leq \|oldsymbol{\epsilon}(oldsymbol{u}-oldsymbol{u}_h)\|_{0,T}\|oldsymbol{\epsilon}(oldsymbol{v})\|_{0,T} + \|\mathcal{A}(oldsymbol{\sigma}-oldsymbol{\sigma}_h)\|_{0,T}\|oldsymbol{\epsilon}(oldsymbol{v})\|_{0,T}. \end{aligned}$$

Since $\|\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_{0,T} \leq \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,T}/(2\nu)$, we obtain

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

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

$$\|(\boldsymbol{\delta} - \boldsymbol{P}_T)(\hat{\boldsymbol{u}} - \boldsymbol{u}_h^*)\|_{0,T} \lesssim h_T \|\boldsymbol{\epsilon}(\boldsymbol{v})\|_{0,T} \lesssim \|\boldsymbol{P}_T(\hat{\boldsymbol{u}} - \boldsymbol{u}_h^*)\|_{0,T} \\ + h_T |\hat{\boldsymbol{u}} - \boldsymbol{u}|_{1,T} + h_T \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,T}.$$

Let I_h denote the nodal interpolant $I_h|_T : \boldsymbol{H}^m(T) \to \boldsymbol{W}_h^*|_T$ with the interpolation estimate [BS94]

(25)
$$|\boldsymbol{u} - I_h \boldsymbol{u}|_{\mu,T} \lesssim h_T^{m+1-\mu} \|\boldsymbol{u}\|_{m+1,T}$$
 for all $\boldsymbol{u} \in \boldsymbol{H}^{m+1}(\Omega), \, \mu = 0, 1.$

The triangle inequality and an inverse estimate show

$$egin{aligned} &|\hat{m{u}}-m{u}|_{1,T} \leq |\hat{m{u}}-I_hm{u}|_{1,T} + |m{u}-I_hm{u}|_{1,T} \ &\lesssim h_T^{-1} \|\hat{m{u}}-I_hm{u}\|_{0,T} + |m{u}-I_hm{u}|_{1,T} \ &\leq h_T^{-1} \|\hat{m{u}}-m{u}\|_{0,T} + h_T^{-1} \|m{u}-I_hm{u}\|_{0,T} + |m{u}-I_hm{u}|_{1,T} \end{aligned}$$

The interpolation estimates (25) and the approximation property (9) yield

$$|\hat{\boldsymbol{u}} - \boldsymbol{u}|_{1,T} \lesssim h_T^m \|\boldsymbol{u}\|_{m+1,T}.$$

After squaring and summing over all $T \in \mathcal{T}_h$, $\boldsymbol{P}_T = \boldsymbol{P}_h|_T$, $\|(\boldsymbol{\delta} - \boldsymbol{P}_h)(\hat{\boldsymbol{u}} - \boldsymbol{u}_h^*)\|_0 \lesssim \|\boldsymbol{P}_h(\hat{\boldsymbol{u}} - \boldsymbol{u}_h^*)\|_0 + h^{m+1}\|\boldsymbol{u}\|_{m+1} + h\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_0$. The estimate (24) and Lemma 3.2 lead to

$$\|(\boldsymbol{\delta} - \boldsymbol{P}_h)(\hat{\boldsymbol{u}} - \boldsymbol{u}_h^*)\|_0 \lesssim h^{k+3} \left(\|\boldsymbol{\sigma}\|_{k+2} + \|\mathbf{div}\boldsymbol{\sigma}\|_{k+1}\right) + h^{m+1}\|\boldsymbol{u}\|_{m+1}.$$

For the second assertion, the triangle inequality shows

$$|\boldsymbol{u} - \boldsymbol{u}_h^*|_{1,T} \le |\boldsymbol{u} - \hat{\boldsymbol{u}}|_{1,T} + |(\boldsymbol{\delta} - \boldsymbol{P}_h)(\hat{\boldsymbol{u}} - \boldsymbol{u}_h^*)|_{1,T} + |\boldsymbol{P}_h(\hat{\boldsymbol{u}} - \boldsymbol{u}_h^*)|_{1,T}.$$

Hence, the result follows from an inverse inequality for the last two terms and the presented analysis. $\hfill\square$

5. A posteriori error control

This section concerns some a posteriori error estimation of the stress error. The analysis is based on the unified approach of [Car05]. Let $\boldsymbol{f}_h := \boldsymbol{P}_h \boldsymbol{f}$ denote the piecewise \boldsymbol{L}^2 projection of \boldsymbol{f} onto \boldsymbol{V}_h , i.e.,

$$\int_{T} (\boldsymbol{f} - \boldsymbol{f}_h) \boldsymbol{v}_h \, dx = 0 \quad \text{for all } \boldsymbol{v}_h \in \boldsymbol{V}_h$$

The oscillations of \boldsymbol{f} are defined as

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

Here and below, the notation $\tilde{\boldsymbol{u}}_h \in \boldsymbol{H}_g^1(\Omega) := \{\boldsymbol{v} \in \boldsymbol{H}^1(\Omega) \mid \boldsymbol{v}|_{\partial\Omega} = \boldsymbol{g}\}$ asserts that $\tilde{\boldsymbol{u}}_h$ is not necessarily a discrete function. The subscript of $\tilde{\boldsymbol{u}}_h$ indicates that it is closely related to the discontinuous approximation \boldsymbol{u}_h . Let $\boldsymbol{H} := \boldsymbol{H}_0^1(\Omega)$ and

$$\boldsymbol{L} := \boldsymbol{L}^2(\Omega; \mathbb{S}) / \mathbb{R} = \Big\{ \boldsymbol{\tau} \in \boldsymbol{L}^2(\Omega; \mathbb{S}) \mid \int_{\Omega} \operatorname{tr} \boldsymbol{\tau} \, dx = 0 \Big\}.$$

For given $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$ the primal mixed formulation of the stressvelocity formulation of the Stokes problem (2) in its weak form seeks for $(\boldsymbol{\sigma}, \boldsymbol{u}) \in \boldsymbol{L} \times \boldsymbol{H}_q^1(\Omega)$ such that

$$-(\boldsymbol{\sigma}, \boldsymbol{\epsilon}(\boldsymbol{v})) = (\boldsymbol{f}, \boldsymbol{v}) \quad \text{for all } \boldsymbol{v} \in \boldsymbol{H},$$
$$(2\nu \mathcal{A}\boldsymbol{\sigma}, \mathcal{A}\boldsymbol{\tau}) - (\boldsymbol{\tau}, \boldsymbol{\epsilon}(\boldsymbol{u})) = 0 \quad \text{for all } \boldsymbol{\tau} \in \boldsymbol{L}.$$

Then the following lemma from [Car05] holds.

Lemma 5.1. The operator $A : X \to X^*$, defined for $(\boldsymbol{\sigma}, \boldsymbol{u}) \in X := \boldsymbol{L} \times \boldsymbol{H}$ by

$$(A(\boldsymbol{\sigma}, \boldsymbol{u}))(\boldsymbol{\tau}, \boldsymbol{v}) := (2\nu \mathcal{A}\boldsymbol{\sigma}, \mathcal{A}\boldsymbol{\tau}) - (\boldsymbol{\sigma}, \boldsymbol{\epsilon}(\boldsymbol{v})) - (\boldsymbol{\tau}, \boldsymbol{\epsilon}(\boldsymbol{u}))$$

is linear, bounded and bijective.

One consequence of this lemma is that, for any approximation $(\boldsymbol{\sigma}_h, \tilde{\boldsymbol{u}}_h)$ to $(\boldsymbol{\sigma}, \boldsymbol{u}) \in \boldsymbol{L} \times \boldsymbol{H}_g^1(\Omega)$ such that $(\boldsymbol{g} - \tilde{\boldsymbol{u}}_h)|_{\partial\Omega} = 0$, it holds that

$$\|\mathcal{A}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\|_0+\|\boldsymbol{\epsilon}(\boldsymbol{u}-\tilde{\boldsymbol{u}}_h)\|_0\approx \|\mathrm{Res}_L\|_{\boldsymbol{L}^*}+\|\mathrm{Res}_H\|_{\boldsymbol{H}^*}$$

10

with the residuals defined by

$$egin{aligned} \operatorname{Res}_H(oldsymbol{v}) &:= (oldsymbol{f},oldsymbol{v}) + (oldsymbol{\sigma}_h,oldsymbol{\epsilon}(oldsymbol{v})) & ext{for all }oldsymbol{v} \in oldsymbol{H}, \ \operatorname{Res}_L(oldsymbol{ au}) &:= (2
u\mathcal{A}oldsymbol{\sigma}_h,\mathcal{A}oldsymbol{ au}) - (oldsymbol{ au},oldsymbol{\epsilon}(oldsymbol{ au}_h)) & ext{for all }oldsymbol{ au} \in oldsymbol{L} \end{aligned}$$

The natural error $\|\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_0 + \|\boldsymbol{\epsilon}_{\mathcal{T}_h}(\boldsymbol{u} - \boldsymbol{u}_h)\|_0$ is unbalanced in the sense that the convergence rate of $\|\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_0$ is of order k + 2, while that of $\|\boldsymbol{u} - \boldsymbol{u}_h\|_0$ is of order k + 1. This motivates the error control of $\|\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_0 + \|\boldsymbol{\epsilon}(\boldsymbol{u} - \tilde{\boldsymbol{u}}_h)\|_0$.

Theorem 5.2. Let $(\boldsymbol{\sigma}, \boldsymbol{u}) \in L \times \boldsymbol{H}_{g}^{1}(\Omega)$ be a solution of (4)–(5). Then $(\boldsymbol{\sigma}_{h}, \tilde{\boldsymbol{u}}_{h}) \in L \times \boldsymbol{H}_{g}^{1}(\Omega)$ satisfies

$$egin{aligned} &\|\mathcal{A}(oldsymbol{\sigma}-oldsymbol{\sigma}_h)\|_0+\|oldsymbol{\epsilon}(oldsymbol{u}- ilde{oldsymbol{u}}_h)\|_0+osc(oldsymbol{f},\mathcal{T}_h)\ &pprox\mu:=\|\mathcal{A}oldsymbol{\sigma}_h-oldsymbol{\epsilon}(ilde{oldsymbol{u}}_h)\|_0+osc(oldsymbol{f},\mathcal{T}_h). \end{aligned}$$

Proof. Since $\boldsymbol{v}|_{\partial\Omega} = 0$, Gauss Theorem yields

$$\operatorname{Res}_{H}(\boldsymbol{v}) = \int_{\Omega} (\boldsymbol{f} - \operatorname{div} \boldsymbol{\sigma}_{h}) \boldsymbol{v} \, dx = \int_{\Omega} (\boldsymbol{f} - \boldsymbol{f}_{h}) \boldsymbol{v} \, dx.$$

Let v_h denote the piecewise integral mean value of $v \in H$; then the Poincaré inequality leads to

$$\operatorname{Res}_{H}(\boldsymbol{v}) = \sum_{T \in \mathcal{T}} \int_{T} (\boldsymbol{f} - \boldsymbol{f}_{h}) (\boldsymbol{v} - \boldsymbol{v}_{h}) dx \lesssim \operatorname{osc}(\boldsymbol{f}, \mathcal{T}_{h}) |\boldsymbol{v}|_{1}$$

The second residual Res_L reads

$$\operatorname{Res}_{L}(\boldsymbol{\tau}) = \int_{\Omega} (\mathcal{A}\boldsymbol{\sigma}_{h} - \boldsymbol{\epsilon}(\tilde{\boldsymbol{u}}_{h})) : \boldsymbol{\tau} \, dx \leq \|\mathcal{A}\boldsymbol{\sigma}_{h} - \boldsymbol{\epsilon}(\tilde{\boldsymbol{u}}_{h})\|_{0} \|\boldsymbol{\tau}\|_{0}.$$

Since $\mathcal{A}\boldsymbol{\sigma} = \boldsymbol{\epsilon}(\boldsymbol{u})$, a triangle inequality shows the efficiency

$$\|\mathcal{A}\boldsymbol{\sigma}_h - \boldsymbol{\epsilon}(ilde{oldsymbol{u}}_h)\|_0 \leq \|\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_0 + \|\boldsymbol{\epsilon}(oldsymbol{u} - ilde{oldsymbol{u}}_h)\|_0.$$

Remark 5.3. In the numerical experiments, $\tilde{\boldsymbol{u}}_h$ is approximated in some finite element space of order at least k + 2 such that the boundary condition is not fulfilled. Suppose that \boldsymbol{g} is sufficiently smooth, i.e., $\boldsymbol{g} \in \boldsymbol{C}(\partial\Omega)$ with $\boldsymbol{g}|_E \in \boldsymbol{H}^{k+3}(E)$ for all $E \subset \partial\Omega$. Let $\boldsymbol{w} \in \boldsymbol{H}^1(\Omega)$ denote a harmonic extension of $\boldsymbol{g} - \boldsymbol{g}_h$ to the interior of Ω [BCD04] such that $\boldsymbol{w}|_{\partial\Omega} = \boldsymbol{g} - \boldsymbol{g}_h$ with $\operatorname{supp}(\boldsymbol{w}) \subseteq \{T \in \mathcal{T}_h : T \cap \partial\Omega \neq \emptyset\}$ and the nodal interpolation $\boldsymbol{g}_h|_E \in P_{k+2}(E;\mathbb{R}^2)$ of \boldsymbol{g} , for all $E \subset \partial\Omega$. For $\tilde{\boldsymbol{u}}_h \in \boldsymbol{g}_h + \boldsymbol{H}_0^1(\Omega)$, Theorem 5.2 along with the triangle inequality shows

$$\|\mathcal{A}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\|_0 + \mathrm{osc}(\boldsymbol{f},\mathcal{T}_h) \lesssim \|\mathcal{A}\boldsymbol{\sigma}_h - \boldsymbol{\epsilon}(ilde{oldsymbol{u}}_h)\|_0 + |oldsymbol{w}|_1 + \mathrm{osc}(oldsymbol{f},\mathcal{T}_h).$$

The proof of [BCD04, Theorem 4.2] reads

$$\|\nabla \boldsymbol{w}\|_0^2 \lesssim \sum_{E \in \mathcal{E}_h, E \subseteq \partial \Omega} \left(h_E^{-1} \|\boldsymbol{g} - \boldsymbol{g}_h\|_{L^2(E)}^2 + h_E \|\partial (\boldsymbol{g} - \boldsymbol{g}_h) / \partial s\|_{L^2(E)}^2 \right).$$

By interpolation it holds

$$\begin{aligned} \|\nabla \boldsymbol{w}\|_{0}^{2} &\lesssim h^{2(k+3)-1} \|\partial^{k+3} \boldsymbol{g} / \partial^{k+3} s\|_{L^{2}(\partial\Omega)}^{2} + h^{2(k+2)+1} \|\partial^{k+3} \boldsymbol{g} / \partial^{k+3} s\|_{L^{2}(\partial\Omega)}^{2} \\ &\lesssim h^{2k+5} \|\partial^{k+3} \boldsymbol{g} / \partial^{k+3} s\|_{L^{2}(\partial\Omega)}^{2}. \end{aligned}$$

Therefore, this term is of higher order $\|\nabla \boldsymbol{w}\|_0 \approx \mathcal{O}(h^{k+5/2})$ compared to $\|\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_0 \approx \mathcal{O}(h^{k+2})$.

The following part of this section is devoted to a second error estimator for which the continuity of $\tilde{\boldsymbol{u}}_h$ is not needed. Instead the estimator involves some (possibly) discontinuous function $\boldsymbol{u}_h^* \in \boldsymbol{H}^1(\mathcal{T}_h) := \{\boldsymbol{v} \in$ $\boldsymbol{L}^2(\Omega) \mid \boldsymbol{v} \mid_T \in \boldsymbol{H}^1(T)$ for all $T \in \mathcal{T}_h\}$ from the postprocessing of Section 4. Let ω_E denote the edge patch $\omega_E := \operatorname{int}(T_+ \cup T_-), \nabla_{\mathcal{T}_h}$ the piecewise defined gradient and $\boldsymbol{\epsilon}_{\mathcal{T}_h}$ its piecewise symmetric part.

Theorem 5.4. Let $(\boldsymbol{\sigma}, \boldsymbol{u}) \in \boldsymbol{L} \times \boldsymbol{H}_g^1(\Omega)$ be a solution of (4)–(5). Then any $\boldsymbol{u}_h^* \in \boldsymbol{H}^1(\mathcal{T}_h)$ satisfies

$$egin{aligned} &\|\mathcal{A}(oldsymbol{\sigma}-oldsymbol{\sigma}_h)\|_0+\|oldsymbol{\epsilon}_{\mathcal{T}_h}(oldsymbol{u}-oldsymbol{u}_h^*)\|_0+osc(oldsymbol{f},\mathcal{T}_h)\ &\lesssim\eta:=\|\mathcal{A}oldsymbol{\sigma}_h-oldsymbol{\epsilon}_{\mathcal{T}_h}(oldsymbol{u}_h^*)\|_0+\left(\sum_{E\in\mathcal{E}_h}h_E^{-1}\|[oldsymbol{u}_h^*]\|_{0,E}^2
ight)^{1/2}+osc(oldsymbol{f},\mathcal{T}_h) \end{aligned}$$

Proof. Let $\tilde{\boldsymbol{u}}_h \in \boldsymbol{H}_g^1(\Omega)$ denote the global minimizer of $\|\boldsymbol{\epsilon}_{\mathcal{T}_h}(\boldsymbol{u}_h^*) - \boldsymbol{\epsilon}(\tilde{\boldsymbol{u}}_h)\|_0$. Theorem 5.2 and the triangle inequality yield, for any $\boldsymbol{u}_h^* \in \boldsymbol{H}^1(\mathcal{T}_h)$, that

$$egin{aligned} &\|\mathcal{A}(oldsymbol{\sigma}-oldsymbol{\sigma}_h)\|_0+\|oldsymbol{\epsilon}_{\mathcal{T}_h}(oldsymbol{u}-oldsymbol{u}_h^*)\|_0\ &\leq\|\mathcal{A}(oldsymbol{\sigma}-oldsymbol{\sigma}_h)\|_0+\|oldsymbol{\epsilon}(oldsymbol{u}-oldsymbol{\tilde{u}}_h)\|_0+\|oldsymbol{\epsilon}_{\mathcal{T}_h}(oldsymbol{u}_h^*)-oldsymbol{\epsilon}(oldsymbol{\tilde{u}}_h)-oldsymbol{u}_h^*)\|_0\ &\lesssim\|\mathcal{A}oldsymbol{\sigma}_h-oldsymbol{\epsilon}(oldsymbol{\tilde{u}}_h)\|_0+\|oldsymbol{\epsilon}_{\mathcal{T}_h}(oldsymbol{u}_h^*)-oldsymbol{\epsilon}(oldsymbol{\tilde{u}}_h))\|_0+\mathrm{osc}(oldsymbol{f},\mathcal{T}_h)\ &\lesssim\|\mathcal{A}oldsymbol{\sigma}_h-oldsymbol{\epsilon}_{\mathcal{T}_h}(oldsymbol{u}_h^*)\|_0+\|oldsymbol{\epsilon}_{\mathcal{T}_h}(oldsymbol{u}_h^*)-oldsymbol{\epsilon}(oldsymbol{\tilde{u}}_h))\|_0+\mathrm{osc}(oldsymbol{f},\mathcal{T}_h). \end{aligned}$$

For the set of nodes \mathcal{N}_h let $(\varphi_z), z \in \mathcal{N}_h$, be a Lipschitz continuous partition of unity

$$\sum_{z \in \mathcal{N}_h} \varphi_z = 1 \quad \text{in } \Omega,$$

with $\varphi_z \in L^2(\Omega)$ and $\varphi_z|_T \in P_1(T)$ for all $T \in \mathcal{T}_h$. For any edge $E \in \mathcal{E}_h$ let $\mathcal{N}(E)$ denote the set of all $z \in \mathcal{N}_h$ with $E \in \mathcal{E}_h(z) := \{E \in \mathcal{E}_h | \varphi_z|_E \neq 0\}$. Theorem 3.1 of [CH07] reads

$$\min_{\boldsymbol{v}\in\boldsymbol{H}_g^1(\Omega)} \|\nabla_{\mathcal{T}_h}(\boldsymbol{u}_h^*-\boldsymbol{v})\|_0^2 \lesssim \sum_{E\in\mathcal{E}_h} \sum_{z\in\mathcal{N}(E)} h_E \|\partial[\varphi_z \boldsymbol{u}_h^*]/\partial s\|_{0,E}^2.$$

Thus $\|\boldsymbol{\epsilon}_{\mathcal{T}_h}(\cdot)\|_0 \lesssim \|\nabla_{\mathcal{T}_h}(\cdot)\|_0$ yields

$$egin{aligned} \|oldsymbol{\epsilon}_{\mathcal{T}_h}(oldsymbol{u}_h^*)-oldsymbol{\epsilon}(ilde{oldsymbol{u}}_h)\|_0^2&=\min_{oldsymbol{v}\inoldsymbol{H}_g^1(\Omega)}\|oldsymbol{\epsilon}_{\mathcal{T}_h}(oldsymbol{u}_h^*-oldsymbol{v})\|_0^2\ &\lesssim \sum_{E\in\mathcal{E}_h}\sum_{z\in\mathcal{N}(E)}h_E\|\partial[arphi_zoldsymbol{u}_h^*]/\partial s\|_{0,E}^2. \end{aligned}$$

12

The local inverse inequality [BS94]

$$h_E \|\partial[\varphi_z \boldsymbol{u}_h^*] / \partial s\|_{0,E}^2 \lesssim h_E^{-1} \|[\varphi_z \boldsymbol{u}_h^*]\|_{0,E}^2$$

leads to

$$egin{aligned} \|oldsymbol{\epsilon}_{\mathcal{T}_h}(oldsymbol{u}_h^*- ilde{oldsymbol{u}}_h)\|_0^2 \lesssim \sum_{E\in\mathcal{E}_h}\sum_{z\in\mathcal{N}(E)}h_E^{-1}\|[arphi_zoldsymbol{u}_h^*]\|_{0,E}^2. \end{aligned}$$

For the linear basis functions it holds that $0 \leq \varphi_z \leq 1$ and the cardinality of $|\mathcal{N}(E)|$ is bounded for all $E \in \mathcal{E}_h$. Thus, it follows that

$$\|oldsymbol{\epsilon}_{\mathcal{T}_h}(oldsymbol{u}_h^*- ilde{oldsymbol{u}}_h)\|_0^2\lesssim \sum_{E\in\mathcal{E}_h}h_E^{-1}\|[oldsymbol{u}_h^*]\|_{0,E}^2.$$

Theorem 5.5. Let the boundary $\partial\Omega$ be sufficiently smooth and $(\sigma, u) \in L \times H^1_g(\Omega) \cap H^{k+3}(\Omega)$ be a solution of (4)–(5). Then the postprocessed velocity field $u_h^* \in W_h^*$ from Section 4 satisfies for quasi-uniform meshes

$$\eta \lesssim \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_0 + \|\boldsymbol{\epsilon}_{\mathcal{T}_h}(\boldsymbol{u} - \boldsymbol{u}_h^*)\|_0 + h^m \|\boldsymbol{u}\|_{m+1} + osc(\boldsymbol{f}, \mathcal{T}_h).$$

Proof. The triangle inequality shows that

$$egin{aligned} \|\mathcal{A}m{\sigma}_h - m{\epsilon}_{\mathcal{T}_h}(m{u}_h^*)\|_0 &\leq \|\mathcal{A}(m{\sigma} - m{\sigma}_h)\|_0 + \|m{\epsilon}_{\mathcal{T}_h}(m{u} - m{u}_h^*)\|_0 \ &\lesssim \|m{\sigma} - m{\sigma}_h\|_0 + \|m{\epsilon}_{\mathcal{T}_h}(m{u} - m{u}_h^*)\|_0. \end{aligned}$$

Hence, it remains to bound the jump term. Since \boldsymbol{u} is continuous,

$$\|[m{u}_h^*]\|_{0,E} = \|[m{u}_h^* - m{u}]\|_{0,E}.$$

The trace inequality reads

$$\|[m{u}_h^*-m{u}]\|_{0,E} \lesssim h_E^{-1/2} \|m{u}_h^*-m{u}\|_{0,\omega_E} + h_E^{1/2} \|
abla_{\mathcal{T}_h}(m{u}_h^*-m{u})\|_{0,\omega_E}$$

and leads to

$$h_E^{-1} \| [m{u}_h^*] \|_{0,E}^2 \lesssim h_E^{-2} \| m{u}_h^* - m{u} \|_{0,\omega_E}^2 + \|
abla_{\mathcal{T}_h} (m{u}_h^* - m{u}) \|_{0,\omega_E}^2$$

The summation over all edges with finite overlap of the edge patches ω_E yields

$$\sum_{E \in \mathcal{E}_h} h_E^{-1} \| [\boldsymbol{u}_h^*] \|_{0,E}^2 \lesssim h_{min}^{-2} \| \boldsymbol{u}_h^* - \boldsymbol{u} \|_{0,\omega_E}^2 + \| \nabla_{\mathcal{T}_h} (\boldsymbol{u}_h^* - \boldsymbol{u}) \|_0^2$$

with $h_{min} := \min_{E \in \mathcal{E}_h} h_E$. The proofs of Theorems 3.4 and 4.1 show

$$egin{aligned} \|oldsymbol{u}-oldsymbol{u}_h^*\|_0&\lesssim h^{m+1}\|oldsymbol{u}\|_{m+1}+h\|oldsymbol{\sigma}-oldsymbol{\sigma}_h\|_0+h\operatorname{osc}(oldsymbol{f},\mathcal{T}_h),\ |
abla_{\mathcal{T}_h}(oldsymbol{u}-oldsymbol{u}_h^*)\|_0&\lesssim h^m\|oldsymbol{u}\|_{m+1}+\|oldsymbol{\sigma}-oldsymbol{\sigma}_h\|_0+\operatorname{osc}(oldsymbol{f},\mathcal{T}_h). \end{aligned}$$

The quasi-uniformity of the meshes $h/h_{min} \lesssim 1$ ends the proof.

Remark 5.6. In most practical examples, \boldsymbol{u} is not as smooth as $\boldsymbol{H}^{m+1}(\Omega)$, and adaptive mesh refinement is needed. However, Theorem 5.5 gives at least a hint that η might be not only reliable but also efficient for more general problems and meshes. This is indeed confirmed by the numerical experiments of Section 7.

6. The Adaptive Finite Element Method

The adaptive finite element algorithm computes a sequence of discrete subspaces

$$(\boldsymbol{\Phi}_{h_0}, \boldsymbol{V}_{h_0}) \subsetneq (\boldsymbol{\Phi}_{h_1}, \boldsymbol{V}_{h_1}) \subsetneq \ldots \subsetneq (\boldsymbol{\Phi}_{h_{\ell-1}}, \boldsymbol{V}_{h_{\ell-1}}) \subsetneq (\boldsymbol{\Phi}_{h_\ell}, \boldsymbol{V}_{h_\ell}) \subset (\boldsymbol{\Phi}, \boldsymbol{V})$$

throughout successive local refinement of the domain Ω . The corresponding sequence of meshes $(\mathcal{T}_{h_{\ell}})_{\ell}$ consists of nested regular triangulations. The AFEM consists of the following loop:

Solve
$$\rightarrow$$
 Estimate \rightarrow Mark \rightarrow Refine.

Solve. Given a mesh \mathcal{T}_h the step Solve calculates the solution of the finite-dimensional saddle point problem

$$\begin{pmatrix} A & B^t & c^t \\ B & & \\ c & & \end{pmatrix} \begin{pmatrix} x \\ y \\ \lambda \end{pmatrix} = \begin{pmatrix} b_g \\ b_f \\ 0 \end{pmatrix}.$$

It is assumed throughout the paper that the discrete equations are solved exactly. The system matrices A and B and the right-hand sides b_g and b_g are computed for the bases $span\{\boldsymbol{\tau}_j\} = \boldsymbol{\Phi}_h$ and $span\{\boldsymbol{v}_j\} = \boldsymbol{V}_h$ by

$$A_{jk} := \int_{\Omega} \mathcal{A}\boldsymbol{\tau}_{j} : \boldsymbol{\tau}_{k} \, dx \quad \text{and} \quad B_{jk} := \int_{\Omega} \boldsymbol{v}_{j} \cdot \mathbf{div}\boldsymbol{\tau}_{k} \, dx;$$
$$b_{g,j} := \int_{\partial\Omega} \boldsymbol{g} \cdot (\boldsymbol{\tau}_{j} \boldsymbol{n}_{E}) \, dx \quad \text{and} \quad b_{f,j} := \int_{\Omega} \boldsymbol{f} \cdot \boldsymbol{v}_{j} \, dx.$$

The discrete solutions for the stress σ_h and the velocity u_h are given by

$$\boldsymbol{\sigma}_h = \sum_{k=1}^{\dim(\boldsymbol{\Phi}_h)} x_k \boldsymbol{\tau}_k \text{ and } \boldsymbol{u}_h = \sum_{k=1}^{\dim(\boldsymbol{V}_h)} y_k \boldsymbol{v}_k.$$

The condition $\int_{\Omega} \operatorname{tr} \boldsymbol{\sigma}_h dx = 0$ is incorporated into the system using the Lagrange multiplier λ and

$$c_j := \int_{\Omega} \operatorname{tr} \boldsymbol{\tau}_j \, dx.$$

For more details see [CGRT08].

Estimate. The error $\|\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_0$ is estimated based on the discrete solution $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h)$ of the underlying saddle point problem

$$\|\mathcal{A}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\|_0 \lesssim \mu_h$$
 and $\|\mathcal{A}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\|_0 \lesssim \eta_h$

with

$$\mu_{h}^{2} = \sum_{T \in \mathcal{T}_{h}} h_{T}^{2} \|\boldsymbol{f} - \boldsymbol{f}_{h}\|_{0,T}^{2} + \sum_{T \in \mathcal{T}_{h}} \|\mathcal{A}\boldsymbol{\sigma}_{h} - \boldsymbol{\epsilon}(\tilde{\boldsymbol{u}}_{h})\|_{0,T}^{2} + \sum_{E \in \mathcal{E}_{h}, E \subseteq \partial \Omega} h_{E} \|\partial(\boldsymbol{g} - \boldsymbol{g}_{h})/\partial s\|_{L^{2}(E)}^{2}, \eta_{h}^{2} = \sum_{T \in \mathcal{T}_{h}} h_{T}^{2} \|\boldsymbol{f} - \boldsymbol{f}_{h}\|_{0,T}^{2} + \sum_{T \in \mathcal{T}_{h}} \|\mathcal{A}\boldsymbol{\sigma}_{h} - \boldsymbol{\epsilon}(\boldsymbol{u}_{h}^{*})\|_{0,T}^{2} + \sum_{E \in \mathcal{E}_{h}} h_{E}^{-1} \|[\boldsymbol{u}_{h}^{*}]\|_{0,E}^{2}$$

Here $\boldsymbol{u}_h^* \in \boldsymbol{W}_h^*$ is the solution of the local postprocessing of Section 4. To obtain a conforming approximation $\tilde{\boldsymbol{u}}_h \in \boldsymbol{g}_h + \boldsymbol{H}_0^1(\Omega)$ the (possibly) discontinuous function \boldsymbol{u}_h^* is smoothen by taking the arithmetic mean value

$$\tilde{\boldsymbol{u}}_h(z) := \frac{1}{|\{T \in \mathcal{T}_h : z \in T\}|} \sum_{T \in \mathcal{T}_h : z \in T} \boldsymbol{u}_h^*(z)|_T$$

for each vertex and edge degree of freedom in $z \in \mathbb{R}^2$. The degrees of freedom on the boundary in points $z \in \mathbb{R}^2$ are interpolated $\tilde{\boldsymbol{u}}_h(z) = \boldsymbol{g}(z)$.

Remark 6.1. In the academic case $\boldsymbol{f} \neq \boldsymbol{f}_h$ the oscillations might dominate the other terms in the a posteriori error estimators and therefore lead to a high overestimation of the error $\|\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_0$. For the realistic benchmark problems of Section 7 the oscillation vanish, and Theorem 5.2 shows for $\tilde{\boldsymbol{u}}_h \equiv \boldsymbol{u}$ that the error estimate is sharp. Because of that and $\tilde{\boldsymbol{u}}_h$ being a higher order approximation of \boldsymbol{u} , the efficiency indices $\mu_h/\|\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_0$ are expected to be close to one.

Mark. Based on the refinement indicators, edges and elements are marked for refinement in a bulk criterion such that $\mathcal{M}_h \subseteq \mathcal{T}_h \cup \mathcal{E}_h$ is an (almost) minimal set of marked edges with

$$\theta \eta_h^2 \le \eta_h^2(\mathcal{M}_h), \qquad \eta_h^2(\mathcal{M}_h) := \sum_{T \in \mathcal{M}_h \cap \mathcal{T}_h} \eta_h^2(T) + \sum_{E \in \mathcal{M}_h \cap \mathcal{E}_h} \eta_h^2(E)$$

for a bulk parameter $0 < \theta \leq 1$. This is done in a greedy algorithm which marks edges and elements with larger contributions.

Refine. In this step of the AFEM loop, the mesh is refined locally corresponding to the set \mathcal{M}_h of marked edges and elements. Once an element is selected for refinement, all of its edges will be refined. In order to preserve the quality of the mesh, i.e., the maximal angle condition or its equivalents, additionally edges have to be marked by the closure algorithm before refinement. For each triangle let one edge be the uniquely defined reference edge E(T). The closure algorithm computes a superset $\overline{\mathcal{M}}_h \supset \mathcal{M}_h$ such that

$$\{E(T): T \in \mathcal{T}_h \text{ with } \mathcal{E}(T) \cap \overline{\mathcal{M}}_h \neq \emptyset \text{ or } T \cap \overline{\mathcal{M}}_h \neq \emptyset\} \subseteq \overline{\mathcal{M}}_h$$



FIGURE 1. Refinement rules: Subtriangles with corresponding reference edges depicted with a second edge.

In other words, once an edge of a triangle or itself is marked for refinement, its reference edge E(T) is among them. After the closure algorithm is applied, one of the following refinement rules is applicable: *no* refinement, *red* refinement, *green* refinement, *blue left* refinement, or *blue right* refinement (see Figure 1).

7. Numerical Experiments

This section is devoted to three numerical experiments with known exact solution and two well-known benchmark problems. The experiments compare the convergence behavior for uniform and adaptive meshes using the error estimators of Section 6 for the stress error. The experiments restrict themselves to the lowest order Arnold–Winther finite element (k = 1) [CGRT08] and the parameter $\nu = 1$. Since the experiments show some superconvergence phenomenon of fourth order for the stress error if the right hand side $\mathbf{f} \equiv 0$, the polynomial order for the postprocessing is accordingly increased by one to $\mathbf{u}_h^* \in P_4(\mathcal{T}_h; \mathbb{R}^2)$. In the absence of superconvergence phenomena this is only a minor computational overhead compared to the postprocessing in $P_3(\mathcal{T}_h; \mathbb{R}^2)$.

Smooth example. The first example concerns the model problem (2) in $\Omega = (0, 1) \times (0, 1)$, with source $\mathbf{f} = (4\pi^2 \sin(\pi(x - y)), 0)^t$ and the Dirichlet data \mathbf{g} chosen in such a way that the smooth solution reads

$$\boldsymbol{u} = (\sin(\pi x)\cos(\pi y) - \cos(\pi x)\sin(\pi y), \sin(\pi x)\cos(\pi y) - \cos(\pi x)\sin(\pi y))^t,$$

$$\boldsymbol{\sigma} = 2\boldsymbol{\epsilon}(\boldsymbol{u}) - p\boldsymbol{\delta}, \text{ with } p = -2\pi(\cos(\pi x)\cos(\pi y) + \sin(\pi x)\sin(\pi y)) + 8/\pi.$$

The computed solution is displayed in Figure 2 as streamline plot for a uniform mesh, and the discrete pressure $p_h := -\text{tr } \boldsymbol{\sigma}_h/2$ is visualized on an adaptive mesh. The convergence history in Figure 3 shows empirically optimal convergence rates of $\mathcal{O}(N^{-3/2}), N := \dim(\boldsymbol{\Phi}_h) + \dim(\boldsymbol{V}_h)$, for both uniform and adaptive meshes. Note that for uniform meshes



FIGURE 2. Streamline plot of the discrete solution on a uniform mesh (left) and discrete pressure $p_h = -\text{tr } \boldsymbol{\sigma}_h/2$ on an adaptive mesh with 409 nodes (right) for the smooth example.



FIGURE 3. Convergence history of $\|\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_0$, $\|\boldsymbol{u} - \boldsymbol{u}_h\|_0$, $\|\boldsymbol{u} - \boldsymbol{u}_h^*\|_0$, μ_h and η_h on adaptively and uniformly refined meshes for the smooth example.

it holds that $\mathcal{O}(N^{-3/2}) \approx \mathcal{O}(h^3)$. Due to the H^2 -regularity of the solution, the numerical results for uniform and adaptive refinements do not differ much. Both error estimators are empirically reliable and efficient. In this academically smooth example the efficiency index is not close to one since the oscillations $\operatorname{osc}(\boldsymbol{f}, \mathcal{T}_h)$ dominate both error estimators. Since the solution is smooth, the postprocessed velocity \boldsymbol{u}_h^* is of higher order $\mathcal{O}(N^{-2})$, which confirms the theoretical result.

Colliding flow. As second example consider the model problem (2) in $\Omega = (-1, 1) \times (-1, 1)$ with source $\mathbf{f} \equiv 0$ and the Dirichlet condition



FIGURE 4. Streamline plot of the discrete solution on a uniform mesh (left) and discrete pressure $p_h = -\text{tr } \sigma_h/2$ on an adaptive mesh with 429 nodes (right) for the colliding flow example.



FIGURE 5. Convergence history of $\|\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_0$, $\|\boldsymbol{u} - \boldsymbol{u}_h\|_0$, $\|\boldsymbol{u} - \boldsymbol{u}_h^*\|_0$, μ_h and η_h on adaptively and uniformly refined meshes for the colliding flow example.

 \boldsymbol{g} chosen in such a way that

$$\boldsymbol{u} = (20xy^4 - 4x^5, 20x^4y - 4y^5)^t$$
 and $\boldsymbol{\sigma} = 2\boldsymbol{\epsilon}(\boldsymbol{u}) - p\boldsymbol{\delta}$

with $p = 120x^2y^2 - 20x^4 - 20y^4 - 32/6$ is the solution. Figure 4 shows the streamlines of the approximated velocity field on a uniform refined mesh and the discrete pressure $p_h = -\text{tr }\boldsymbol{\sigma}_h/2$ on an adaptive refined mesh. Both error estimators are numerically reliable and efficient as shown in Figure 5. Since there are no oscillations in this example, the efficiency indices of both estimators are much closer to one. It is remarkable that the order of convergence approaches experimentally



FIGURE 6. Streamline plot of the discrete solution on a uniform mesh (left) and discrete pressure $p_h = -\text{tr } \boldsymbol{\sigma}_h/2$ on an adaptive mesh with 223 nodes (right) for the slit example.



FIGURE 7. Convergence history of $\|\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_0$, $\|\boldsymbol{u} - \boldsymbol{u}_h\|_0$, $\|\boldsymbol{u} - \boldsymbol{u}_h^*\|_0$, μ_h and η_h on adaptively and uniformly refined meshes for the slit example.

 $\mathcal{O}(N^{-2})$. This superconvergence was previously observed in [CGRT08] where it is conjectured that this effect takes place due to $\mathbf{f} \equiv 0$. Consequently, the postprocessed velocity $\mathbf{u}_h^* \in P_4(\mathcal{T}_h; \mathbb{R}^2)$ shows also an increased empirical convergence rate of $\mathcal{O}(N^{-5/2})$. The fact that the adaptive mesh refinement algorithm destroys the symmetry of the mesh might be a reason that the error for adaptively refined meshes is larger than that for uniform meshes.

Slit example. As an example with a nonconvex domain, consider the model problem (2) in $\Omega = (-1, 1)^2 \setminus [0, 1) \times \{0\}$. The right hand side



FIGURE 8. Adaptive meshes obtained using μ_h (left) with 672 nodes and η_h (right) with 761 nodes for the slit example.

 \boldsymbol{f} and the Dirichlet data \boldsymbol{g} are chosen in such a way that the exact solution in polar coordinates reads

$$\boldsymbol{u} = \frac{3\sqrt{r}}{2} \left(\cos\left(\frac{\theta}{2}\right) - \cos\left(\frac{3\theta}{2}\right), 3\sin\left(\frac{\theta}{2}\right) - \sin\left(\frac{3\theta}{2}\right) \right)^t,$$
$$\boldsymbol{\sigma} = 2\boldsymbol{\epsilon}(\boldsymbol{u}) - p\boldsymbol{\delta} \quad \text{with} \quad p = -\frac{6}{\sqrt{r}}\cos\left(\frac{\theta}{2}\right).$$

Due to the re-entrant corner at the origin of the domain, this example allows a singular solution. A discrete approximation of the velocity and the pressure $p_h = -\mathrm{tr} \, \boldsymbol{\sigma}_h/2$ on uniformly and adaptively refined meshes is shown in Figure 6. The convergence history in Figure 7 shows poor convergence for the error in the case of uniform meshes. On the other hand, adaptive refinement results in optimal convergence of the error and in reliable and efficient a posteriori error control, which underlines the importance of adaptivity. It can be observed that η_h is underestimating the error and that the efficiency index is much smaller than that from μ_h , which is close to one. Additionally the error for the adaptive meshes generated with μ_h is significant smaller than that generated with η_h . Figure 8 shows pictures of adaptively refined meshes for μ_h and η_h , which show strong refinement towards the singularity at the origin. The postprocessed velocity \boldsymbol{u}_h^* shows empirical superconvergence with convergence rates of $\mathcal{O}(N^{-2})$ for adaptive meshes for both estimators.

Backward facing step. This example is a well-known benchmark problem for flow problems in the domain Ω of Figure 9. Consider the model problem (2) with $\mathbf{f} \equiv 0$, $\mathbf{g}(x,y) = (0,0)^t$ for -2 < x < 8, $\mathbf{g}(x,y) = (-y(y-1)/10,0)^t$ for x = -2 and $\mathbf{g}(x,y) = (-(y+1)(y-1)/80,0)^t$ for x = 8. The numerical solution of the velocity field and the pressure $p_h = -\text{tr } \boldsymbol{\sigma}_h/2$ on uniformly refined meshes are shown in Figure 9. Note that the pressure is high on the left and low on



FIGURE 9. Approximated velocity field (top) and discrete pressure $p_h = -\text{tr } \boldsymbol{\sigma}_h/2$ (bottom) for the backward facing step example on uniform meshes.



FIGURE 10. Adaptive meshes obtained using μ_h (top) with 760 nodes and η_h (bottom) with 791 nodes for the backward facing step example.



FIGURE 11. Approximated velocity field near the bottom corner for the backward facing step example on uniform meshes.



FIGURE 12. Approximated velocity field (left) and discrete pressure $p_h = -\text{tr } \boldsymbol{\sigma}_h/2$ (right) for the driven cavity example on uniform meshes.



FIGURE 13. Approximated velocity field near the left and right lower corner for the driven cavity example on uniform meshes.

the right. Figure 10 shows two adaptively refined meshes for both error estimators which look quite similar and show strong refinement towards the singularity at the origin. A zoom of the lower left corner in Figure 11 shows not only one eddy, but two streamlines at the left corner indicate a second one. This indicates a high stability of the numerical scheme.

Lid-driven cavity flow. As last example consider the lid-driven cavity flow benchmark problem. Consider the model problem (2) in $\Omega = (-1,1) \times (1,1)$ with $\mathbf{f} \equiv 0$, $\mathbf{g}(x,y) = (0,0)^t$ for y < 1 and $\mathbf{g}(x,y) = (1,0)^t$ for y = 1. Figure 12 displays the numerical solution of the velocity on a uniform mesh with two Moffat eddies at the bottom corners and an approximation of the pressure $p_h = -\text{tr } \boldsymbol{\sigma}_h/2$ on an adaptive mesh. The absolute largest values for the pressure occur in the top corners; in the other areas there seems to be almost no pressure. Figure 13 shows a zoom towards the Moffat eddies in the left and right lower corners.



FIGURE 14. Adaptive meshes obtained using μ_h (left) with 422 nodes and η_h (right) with 355 nodes for the driven cavity example.

Near the bottom corners one or two lines indicate a more detailed resolution of the Moffat eddies, which again illustrates the high stability of the numerical scheme. Both adaptively refined meshes show strong refinement towards the two left and right top corners in Figure 14. It seems that the area away from the corners is refined only to prevent hanging nodes and not due to a high refinement indicator.

Acknowledgement

The authors would like to thank the anonymous referees for their valuable comments and suggestions. The paper was written while the first and second authors enjoyed the kind hospitality of the Department CSE of Yonsei University.

References

- [AO00] M. Ainsworth and J.T. Oden, A posteriori error estimation in finite element analysis, Pure and Applied Mathematics, Wiley-Interscience, New York, 2000.
- [AW02] D.N. Arnold and R. Winther, Mixed finite elements for elasticity, Numer. Math. 92 (2002), no. 3, 401–419.
- [BCD04] S. Bartels, C. Carstensen, and G. Dolzmann, Inhomogeneous Dirichlet conditions in a priori and a posteriori finite element error analysis, Numer. Math. 99 (2004), no. 1, 1–24.
- [BF91] F. Brezzi and M. Fortin, Mixed and hybrid finite element methods, Springer Series in Computational Mathematics, vol. 15, Springer-Verlag, New York, 1991.
- [Bre74] F. Brezzi, On the existence, uniqueness and approximation of saddlepoint problems arising from lagrangian multipliers, Rev. Franc. Automat. Inform. Rech. Operat. R-2 (1974), no. 8, 129–151.
- [BS94] S.C. Brenner and L.R. Scott, The mathematical theory of finite element methods, Texts in Applied Mathematics, vol. 15, Springer-Verlag, New York, 1994.
- [BS01] I. Babuška and T. Strouboulis, The finite element methods and its reliability, Clarendon Press, Oxford, 2001.

- [BW91] R.E. Bank and B.D. Welfert, A posteriori error estimates for the Stokes problem, SIAM J. Numer. Anal. 28 (1991), no. 3, 591–623.
- [Car05] C. Carstensen, A unifying theory of a posteriori finite element error control, Numer. Math. 100 (2005), no. 4, 617–637.
- [CGRT08] C. Carstensen, D. Günther, J. Reininghaus, and J. Thiele, The Arnold-Winther mixed FEM in linear elasticity. I. Implementation and numerical verification, Comput. Methods Appl. Mech. Engrg. 197 (2008), no. 33-40, 3014–3023.
- [CH07] C. Carstensen and J. Hu, A unifying theory of a posteriori error control for nonconforming finite element methods, Numer. Math. 107 (2007), no. 3, 473–502.
- [DDP95] E. Dari, R. Durán, and C. Padra, Error estimators for nonconforming finite element approximations of the Stokes problem, Math. Comp. 64 (1995), no. 211, 1017–1033.
- [GR86] V. Girault and P.-A. Raviart, *Finite element methods for Navier-Stokes equations*, Springer Series in Computational Mathematics, vol. 5, Springer-Verlag, Berlin, 1986.
- [Kim07] K.Y. Kim, A posteriori error analysis for locally conservative mixed methods, Math. Comput. 76 (2007), no. 257, 43–66.
- [LM08] M.G. Larson and A. Målqvist, A posteriori error estimates for mixed finite element approximations of elliptic problems, Numer. Math. 108 (2008), no. 3, 487–500.
- [Ste88] R. Stenberg, A family of mixed finite elements for the elasticity problem, Numer. Math. 53 (1988), no. 5, 513–538.
- [Ver96] R. Verfürth, A review of a posteriori error estimation and adaptive mesh-refinement techniques, Wiley and Teubner, Stuttgart, 1996.
- [Voh10] M. Vohralík, Unified primal formulation-based a priori and a posteriori error analysis of mixed finite element methods, Math. Comput. 79 (2010), no. 272, 2001–2032.

(C. Carstensen) DEPARTMENT OF COMPUTATIONAL SCIENCE AND ENGINEER-ING, YONSEI UNIVERSITY, SEOUL 120-749, KOREA AND INSTITUT FÜR MATHE-MATIK, HUMBOLDT-UNIVERSITÄT ZU BERLIN, D-10099 BERLIN, GERMANY *E-mail address*: cc@mathematik.hu-berlin.de

(J. Gedicke) Institut für Mathematik, Humboldt-Universität zu Berlin, D-10099 Berlin, Germany

E-mail address: gedicke@mathematik.hu-berlin.de

(E.-J. Park) Department of Computational Science and Engineering, Yonsei University, Seoul 120-749, Korea

E-mail address: ejpark@yonsei.ac.kr