SIAM J. NUMER. ANAL.
Vol. 51, No. 3, pp. 1715–1734




# QUASI-OPTIMAL ADAPTIVE PSEUDOSTRESS APPROXIMATION OF THE STOKES EQUATIONS*

CARSTEN CARSTENSEN†, DIETMAR GALLISTL‡, AND MIRA SCHEDENSACK‡

*Dedicated to Professor Vidar Thomée on the occasion of his 80th birthday*

**Abstract.** The pseudostress-velocity formulation of the stationary Stokes problem allows some quasi-optimal Raviart–Thomas mixed finite element formulation for any polynomial degree. The adaptive algorithm employs standard residual-based explicit a posteriori error estimation from Carstensen, Kim, and Park [*SIAM J. Numer. Anal.*, 49 (2011), pp. 2501–2523] for the lowest-order Raviart–Thomas finite element functions in a simply connected Lipschitz domain. This paper proves optimal convergence rates in terms of the number of unknowns of the adaptive mesh-refining algorithm based on the concept of approximation classes. The proofs use some novel equivalence to first-order nonconforming Crouzeix–Raviart discretization plus a particular Helmholtz decomposition of deviatoric tensors.

**Key words.** mixed finite element approximations, a posteriori error estimates, Stokes problem, pseudostress formulation, adaptivity, optimality

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

**DOI.** 10.1137/110852346

## 1. Introduction.

The pseudostress-velocity formulation of the stationary Stokes equations

$$-\Delta u + \nabla p = f \quad \text{and} \quad \operatorname{div} u = 0 \quad \text{in } \Omega \tag{1.1}$$

with Dirichlet boundary conditions along the polygonal boundary $\partial\Omega$ has attracted recent investigation. The early paper [27] introduces the pseudostress method for symmetric stress tensors in $H(\operatorname{div}, \Omega; \mathbb{R}^{2\times 2})$ while the version in this paper is more recently introduced in [10] and [8, 9, 11, 14, 21, 22, 23].

The explicit residual-based a posteriori error estimates from [14] are utilized to drive a novel adaptive mesh-refining algorithm as a sequence of successive loops with

$$\text{SOLVE}, \text{ESTIMATE}, \text{MARK}, \text{REFINE}.$$

In the context of elliptic PDEs, it has recently become clear how to prove optimal convergence rates [20, 31, 18]. The analysis for the lowest-order adaptive pseudostress method (APSFEM) follows ideas of the analysis of nonconforming and mixed adaptive algorithms [2, 12, 13, 17, 19, 29] and enables the key properties quasi orthogonality and discrete reliability. In the context of the Stokes equations, recent progress is documented in [1, 25, 26] for the nonconforming Crouzeix–Raviart finite element method. However, the recent work [1] is disputable (the estimate in line 23 on page 983 in the last step of the proof of Lemma 5.2 is wrong for refinements over many levels) and

*Received by the editors October 19, 2011; accepted for publication (in revised form) March 7, 2013; published electronically June 13, 2013.
http://www.siam.org/journals/sinum/51-3/85234.html

†Institut für Mathematik, Humboldt-Universität zu Berlin, D-10099 Berlin, Germany, and Department of CSE, Yonsei University, Seoul, Korea (cc@math.hu-berlin.de). This author's research was supported by the WCU program through KOSEF (R31-2008-000-10049-0).

‡Institut für Mathematik, Humboldt-Universität zu Berlin, D-10099 Berlin, Germany (gallistl@math.hu-berlin.de, schedens@math.hu-berlin.de).

the series of work based on [25] utilizes an inappropriate Strang–Fix-like procedural error and eventually proves equivalence to an approximation class [26].

The proof of the quasi orthogonality in this paper employs a representation formula in which the solution of the pseudostress method is obtained from some postprocessing of the Crouzeix–Raviart nonconforming finite element method [15, 28]. The discrete reliability proof employs the discrete Helmholtz decomposition of piecewise constant deviatoric matrices introduced in [16].

The adaptive algorithm Apsfem is introduced in section 3 and is shown to be quasi-optimally convergent with respect to the approximation class

$$\mathcal{A}_s := \big\{(\sigma, f, g) \in H(\operatorname{div}, \Omega; \mathbb{R}^{2\times 2})/\mathbb{R} \times L^2(\Omega; \mathbb{R}^2) \\ \times \big(H^1(\Omega; \mathbb{R}^2) \cap H^1(\mathcal{E}(\partial\Omega); \mathbb{R}^2)\big) \,\big|\, |(\sigma, f, g)|_{\mathcal{A}_s} < \infty\big\}$$

with

$$|(\sigma, f, g)|_{\mathcal{A}_s} := \sup_{N\in\mathbb{N}} N^s \inf_{\mathcal{T}\in\mathbb{T}(N)} \left( \|\sigma - \sigma_\mathcal{T}\|^2_{L^2(\Omega)} + \operatorname{osc}^2(f, \mathcal{T}) + \operatorname{osc}^2\left(\frac{\partial g}{\partial s}, \mathcal{E}(\partial\Omega)\right)\right)^{1/2}.$$

In the infimum, $\mathcal{T}$ runs through all admissible triangulations $\mathbb{T}(N)$ that are refined from $\mathcal{T}_0$ by NVB (cf. Figure 3.1) with a number $|\mathcal{T}|$ of triangles bounded as $|\mathcal{T}|-|\mathcal{T}_0| \le N$ and the solution $\sigma_\mathcal{T}$ of (2.2) with respect to $\mathcal{T}$. (Further details and notation, in particular on $g$, are given in sections 2 and 3.) Given the exact stress $\sigma := Du - p\, I_{2\times 2}$ and some bulk parameter $\theta$ sufficiently small, Apsfem generates sequences of triangulations $(\mathcal{T}_\ell)_\ell$ and discrete solutions $(u_\ell, \sigma_\ell)_\ell$ of optimal convergence rate in the sense that

$$(1.2)\qquad (|\mathcal{T}_\ell| - |\mathcal{T}_0|)^s \Big( \|\sigma - \sigma_\ell\|^2_{L^2(\Omega)} + \operatorname{osc}^2(f, \mathcal{T}_\ell) + \operatorname{osc}^2(\partial g/\partial s, \mathcal{E}_\ell(\partial\Omega))\Big)^{1/2} \\ \le C\, |(\sigma, f, g)|_{\mathcal{A}_s}.$$

This estimate states optimality for $C = 1$ (then $\mathcal{T}_\ell$ is optimal amongst all possible triangulations) while the main result shows that $C$ is bounded in terms of the initial triangulation $\mathcal{T}_0$, and so $\mathcal{T}_\ell$ performs optimal in (1.2) up to a positive generic constant $C$ which does not depend on the mesh-size (denoted by $C \approx 1$ in what follows) and is called a quasi-optimal triangulation. Therefore, the convergence rates are optimal while the convergence is said to be quasi optimal in terms of the approximation class $\mathcal{A}_s$.

The remaining parts of this paper are organized as follows. Section 2 introduces the basic notation as well as the pseudostress method and some equivalence to a nonconforming Crouzeix–Raviart discretization for one triangulation $\mathcal{T}$. It also recalls the a posteriori error estimates of [14] for the pseudostress method. Section 3 introduces the adaptive pseudostress method Apsfem, specifies more details on the approximation class $\mathcal{A}_s$ and some equivalent characterization $\mathcal{A}'_s$, and states the aforementioned optimality result. Section 4 shows convergence of Apsfem and contraction of a convex combination of estimator, error, and data oscillations. Section 5 establishes the discrete reliability and concludes the optimality proof. Computational experiments conclude the paper in section 6.

Throughout this paper, standard notation on Lebesgue and Sobolev spaces and their norms is employed with $(\cdot,\cdot)_\Omega$ the $L^2$ inner product, $H(\operatorname{div}, \Omega) := \{v \in L^2(\Omega) \,|\, \operatorname{div} v \in L^2(\Omega)\}$, while $\langle\cdot,\cdot\rangle := \langle\cdot,\cdot\rangle_{H^{1/2}(\partial\Omega)\times H^{-1/2}(\partial\Omega)}$ denotes the duality pairing of $H^{1/2}(\partial\Omega)$ with $H^{-1/2}(\partial\Omega)$ on the boundary $\partial\Omega$.

The formula $A \lesssim B$ represents an inequality $A \leq CB$ for some mesh-independent, positive generic constant $C$; $A \approx B$ abbreviates $A \lesssim B \lesssim A$. By convention, all generic constants $C \approx 1$ do not depend on the mesh-size but may depend on the fixed coarse triangulation $\mathcal{T}_0$ and its interior angles.

The measure $|\cdot|$ is context-sensitive and refers to the number of elements of some finite set (e.g., the number $|\mathcal{T}|$ of triangles in a triangulation $\mathcal{T}$) or the length $|E|$ of an edge $E$ or the area $|T|$ of some domain $T$ and not just the modulus of a real number or the Euclidean length of a vector.

**2. Preliminaries.** Let $\Omega$ be a simply connected bounded Lipschitz domain with polygonal boundary $\partial\Omega$ and outer unit normal $\nu$, and let $\mathcal{T}$ be some shape-regular triangulation of $\Omega$ into closed triangles $T \in \mathcal{T}$. The set $\mathcal{E}$ contains all edges of $\mathcal{T}$, $\mathcal{E}(\Omega)$ all interior edges, and $\mathcal{E}(\partial\Omega)$ all edges on the boundary; $\mathcal{E}(T)$ is the set of edges of a triangle $T$. For interior edges, $[\cdot]_E := \cdot|_{T_+} - \cdot|_{T_-}$ denotes the jump across the edge $E = T_+ \cap T_-$ shared by the two elements $T_\pm \in \mathcal{T}$, and $\omega_E := \operatorname{int}(T_+ \cup T_-)$ denotes the edge-patch. For $E \in \mathcal{E}(\partial\Omega)$, the jump includes the boundary conditions, namely $[\operatorname{dev}\sigma_{\mathrm{PS}}\tau_E]_E := \operatorname{dev}\sigma_{\mathrm{PS}}|_{T_+}\tau_E - (\partial g/\partial s)$ for the one element $T_+$ with $E \subset T_+$, and $\omega_E := \operatorname{int}(T_+)$. In addition, for any edge $E \in \mathcal{E}$, $\operatorname{mid}(E)$ names its midpoint and $\nu_E = \nu_{T_+}$ is the unit normal vector exterior to $T_+$ along $E$ and $\tau_E$ is the unit tangential vector along $E|_{T_+}$. For any triangle $T \in \mathcal{T}$, $\operatorname{mid}(T)$ denotes the center of inertia, and the piecewise constant function $\operatorname{mid}(\mathcal{T}) \in P_0(\mathcal{T};\mathbb{R}^2)$ is defined through $\operatorname{mid}(\mathcal{T})|_T = \operatorname{mid}(T)$.

Let $D_{\mathrm{NC}}$ and $\operatorname{div}_{\mathrm{NC}}$ denote the piecewise action of the gradient and the divergence with respect to the triangulation $\mathcal{T}$. For a vector field $\beta = (\beta_1, \beta_2)$ the operators Curl and curl read as

$$\operatorname{Curl}\beta := \begin{pmatrix} -\partial\beta_1/\partial x_2 & \partial\beta_1/\partial x_1 \\ -\partial\beta_2/\partial x_2 & \partial\beta_2/\partial x_1 \end{pmatrix} \quad \text{and} \quad \operatorname{curl}\beta := \frac{\partial\beta_2}{\partial x_1} - \frac{\partial\beta_1}{\partial x_2}.$$

For matrices $\sigma \in \mathbb{R}^{2\times 2}$ the divergence and curl are defined rowwise,

$$\operatorname{div}\sigma := \begin{pmatrix} \partial\sigma_{11}/\partial x_1 + \partial\sigma_{12}/\partial x_2 \\ \partial\sigma_{21}/\partial x_1 + \partial\sigma_{22}/\partial x_2 \end{pmatrix} \quad \text{and} \quad \operatorname{curl}\sigma := \begin{pmatrix} \partial\sigma_{12}/\partial x_1 - \partial\sigma_{11}/\partial x_2 \\ \partial\sigma_{22}/\partial x_1 - \partial\sigma_{21}/\partial x_2 \end{pmatrix}.$$

The $2\times 2$ unit matrix is denoted by $I_{2\times 2}$ and the Euclid product of matrices is denoted by a colon, e.g., $A : B = \sum_{j,k=1}^{2} A_{jk}B_{jk}$ for $A$, $B \in \mathbb{R}^{2\times 2}$; $\operatorname{tr}(A) := A : I_{2\times 2}$ is the trace, and $\operatorname{dev}(A) := A - 1/2\operatorname{tr}(A)I_{2\times 2}$ is the deviatoric part of $A \in \mathbb{R}^{2\times 2}$. The dot denotes the product of two one-dimensional lists of the same length while $\otimes$ denotes the rank-one matrix product, e.g., $a \cdot b = a^\top b \in \mathbb{R}$ and $a \otimes b = ab^\top \in \mathbb{R}^{2\times 2}$ for $a, b \in \mathbb{R}^2$. The interior of a set $\omega \subset \mathbb{R}^2$ is denoted by $\operatorname{int}(\omega)$.

Throughout the paper, the discrete spaces read as

$$P_0(\mathcal{T}) := \left\{ v \in L^2(\Omega) \,|\; v|_T \text{ is constant for all } T \in \mathcal{T} \right\},$$
$$P_1(\mathcal{T}) := \left\{ v \in L^2(\Omega) \,|\; v|_T \text{ is affine for all } T \in \mathcal{T} \right\}.$$

Analogous notation applies to vectors and matrices. For $f \in L^2(\Omega;\mathbb{R}^2)$, $\Pi_{\mathcal{T}} f \in P_0(\mathcal{T})$ denotes the $L^2$ best approximation in $P_0(\mathcal{T};\mathbb{R}^2)$. The lowest-order Raviart–Thomas space is defined as

$$\mathrm{RT}_0(T) := \{v \in P_1(T;\mathbb{R}^2) \,|\, \exists\, a, b, c \in \mathbb{R},\ v = (a,b) + c(x_1,x_2)\},$$
$$\mathrm{RT}_0(\mathcal{T}) := \{q \in H(\operatorname{div},\Omega) \,|\, \forall T \in \mathcal{T},\ q|_T \in \mathrm{RT}_0(T)\}.$$

Define

$$H(\mathrm{div},\Omega;\mathbb{R}^{2\times 2})/\mathbb{R} := \left\{\tau \in L^2(\Omega;\mathbb{R}^{2\times 2}) \,\middle|\, \forall j=1,2,\ (\tau_{j1},\tau_{j2}) \in H(\mathrm{div},\Omega) \text{ and } \int_\Omega \mathrm{tr}(\tau)\,dx = 0\right\},$$

$$\mathrm{PS}(\mathcal{T}) := \left\{\tau \in P_1(\mathcal{T};\mathbb{R}^{2\times 2}) \cap H(\mathrm{div},\Omega;\mathbb{R}^{2\times 2})/\mathbb{R} \,\middle|\, \forall j=1,2,\ (\tau_{j1},\tau_{j2}) \in \mathrm{RT}_0(\mathcal{T})\right\}.$$

The weak form of problem (1.1) is formally equivalent and reads as follows: Given $f \in L^2(\Omega;\mathbb{R}^2)$ and $g \in H^1(\Omega;\mathbb{R}^2) \cap H^1(\mathcal{E}(\partial\Omega);\mathbb{R}^2)$ with $\int_{\partial\Omega} g\cdot\nu\,ds = 0$ seek $\sigma \in H(\mathrm{div},\Omega;\mathbb{R}^{2\times 2})/\mathbb{R}$ and $u \in L^2(\Omega;\mathbb{R}^2)$ such that

$$\begin{aligned} (\mathrm{dev}\ \sigma,\tau)_\Omega + (\mathrm{div}\ \tau,u)_\Omega &= \langle g,\tau\,\nu\rangle && (\tau \in H(\mathrm{div},\Omega;\mathbb{R}^{2\times 2})/\mathbb{R}),\\ (\mathrm{div}\ \sigma,v)_\Omega &= -(f,v)_\Omega && (v \in L^2(\Omega;\mathbb{R}^2)). \end{aligned} \tag{2.1}$$

The discrete formulation of (2.1) seeks $\sigma_{\mathrm{PS}} \in \mathrm{PS}(\mathcal{T})$ and $u_{\mathrm{PS}} \in P_0(\mathcal{T}_\ell;\mathbb{R}^2)$ such that

$$\begin{aligned} (\mathrm{dev}\ \sigma_{\mathrm{PS}},\tau_{\mathrm{PS}})_\Omega + (\mathrm{div}\ \tau_{\mathrm{PS}},u_{\mathrm{PS}})_\Omega &= \langle g,\tau_{\mathrm{PS}}\,\nu\rangle && (\tau_{\mathrm{PS}} \in \mathrm{PS}(\mathcal{T})),\\ (\mathrm{div}\ \sigma_{\mathrm{PS}},v_{\mathrm{PS}})_\Omega &= -(f,v_{\mathrm{PS}})_\Omega && (v_{\mathrm{PS}} \in P_0(\mathcal{T};\mathbb{R}^2)). \end{aligned} \tag{2.2}$$

For the inf-sup condition and the quasi-optimal convergence of the discrete pseudostress problem, see [8, 14]. The subsequent notation on nonconforming finite element schemes plays a dominant role in the analysis of this paper,

$$\begin{aligned} \mathrm{CR}^1(\mathcal{T}) &:= \left\{v \in P_1(\mathcal{T}) \,\middle|\, \begin{array}{l} v \text{ is continuous in } \mathrm{mid}(E)\\ \text{for all } E \in \mathcal{E}\end{array}\right\},\\ \mathrm{CR}^1_0(\mathcal{T}) &:= \left\{v \in \mathrm{CR}^1(\mathcal{T}) \,|\, v(\mathrm{mid}(E)) = 0 \ \text{ for all } E \in \mathcal{E}(\partial\Omega)\right\},\\ \mathrm{Z}_{\mathrm{CR}}(\mathcal{T}) &:= \left\{v \in \mathrm{CR}^1_0(\mathcal{T};\mathbb{R}^2) \,|\, \mathrm{div}_{\mathrm{NC}}\, v = 0 \text{ a.e. in } \Omega\right\},\\ Q_{\mathrm{CR}}(\mathcal{T}) &:= \left\{q_{\mathrm{CR}} \in P_0(\mathcal{T}) \,\middle|\, \int_\Omega q_{\mathrm{CR}}dx = 0\right\}. \end{aligned}$$

The Crouzeix–Raviart nonconforming finite element formulation of (1.1) [4, 6, 7, 24] employs some discretization $g_{\mathrm{CR}} \in \mathrm{CR}^1(\mathcal{T};\mathbb{R}^2)$ such that $\int_E g_{\mathrm{CR}}\,ds = \int_E g\,ds$ for all $E \in \mathcal{E}(\partial\Omega)$. Given $f_\mathcal{T} = \Pi_\mathcal{T} f$ seek $\tilde u_{\mathrm{CR}} \in g_{\mathrm{CR}} + \mathrm{CR}^1_0(\mathcal{T};\mathbb{R}^2)$ and $\tilde p_{\mathrm{CR}} \in Q_{\mathrm{CR}}(\mathcal{T})$ such that

$$\begin{aligned} (D_{\mathrm{NC}}\tilde u_{\mathrm{CR}}, D_{\mathrm{NC}}v_{\mathrm{CR}})_\Omega - (\tilde p_{\mathrm{CR}}, \mathrm{div}_{\mathrm{NC}}\, v_{\mathrm{CR}})_\Omega &= (f_\mathcal{T}, v_{\mathrm{CR}})_\Omega && (v_{\mathrm{CR}} \in \mathrm{CR}^1_0(\mathcal{T};\mathbb{R}^2)),\\ (q_{\mathrm{CR}}, \mathrm{div}_{\mathrm{NC}}\, \tilde u_{\mathrm{CR}})_\Omega &= 0 && (q_{\mathrm{CR}} \in Q_{\mathrm{CR}}(\mathcal{T})). \end{aligned} \tag{2.3}$$

The following result of [15] utilizes the notation $\bullet - \mathrm{mid}(\mathcal{T})$ to abbreviate the function $x - \mathrm{mid}(T)$ for $x \in T \in \mathcal{T}$ with midpoint $\mathrm{mid}(T)$.

THEOREM 2.1 (pseudostress representation formula). *Let* $(\tilde u_{\mathrm{CR}}, \tilde p_{\mathrm{CR}}) \in (g_{\mathrm{CR}} + \mathrm{CR}^1_0(\mathcal{T};\mathbb{R}^2)) \times Q_{\mathrm{CR}}(\mathcal{T})$ *be the solution of* (2.3) *for the right-hand side* $f_\mathcal{T} := \Pi_\mathcal{T} f$ *for* $f \in L^2(\Omega;\mathbb{R}^2)$. *Then* $\sigma_{\mathrm{PS}} \in \mathrm{PS}(\mathcal{T})$ *and* $u_{\mathrm{PS}} \in P_0(\mathcal{T};\mathbb{R}^2)$, *defined by*

$$\sigma_{\mathrm{PS}} := D_{\mathrm{NC}}\tilde u_{\mathrm{CR}} - \frac{f_\mathcal{T}}{2} \otimes (\bullet - \mathrm{mid}(\mathcal{T})) - \tilde p_{\mathrm{CR}}\, I_{2\times 2} \tag{2.4}$$

*and*

$$u_{\mathrm{PS}} := \Pi_\mathcal{T}\tilde u_{\mathrm{CR}} + \frac{1}{4}\Pi_\mathcal{T}\left(\mathrm{dev}(f_\mathcal{T} \otimes (\bullet - \mathrm{mid}(\mathcal{T})))\,(\bullet - \mathrm{mid}(\mathcal{T}))\right), \tag{2.5}$$

*solve* (2.2).

*Proof.* The proof is given here for completeness and to stress the consequences of the inhomogeneous boundary conditions. The first claim

$$\sigma_{\mathrm{PS}}(j) := (\sigma_{\mathrm{PS}}(j,1), \sigma_{\mathrm{PS}}(j,2)) \in H(\mathrm{div}, \Omega) \quad \text{ for } j = 1, 2$$

is equivalent to $[\sigma_{\mathrm{PS}}(j)\, \nu_E]_E = 0$ for all $E \in \mathcal{E}(\Omega)$. Given an interior edge $E$, define the edge-oriented nonconforming Crouzeix–Raviart basis function $\psi_E \in \mathrm{CR}^1_0(\mathcal{T})$ through $\psi_E|_E = 1$ and $\psi_E(\mathrm{mid}(F)) = 0$ for all $E \neq F \in \mathcal{E}$. For $e_1 = (1,0)^\top, e_2 = (0,1)^\top$, a piecewise integration by parts shows, for $j = 1, 2$, that

$$\begin{aligned}|E|[\sigma_{\mathrm{PS}}(j)\, \nu_E]_E &= \int_E [\sigma_{\mathrm{PS}}\, \nu_E]_E \cdot e_j \psi_E \, ds \\ &= \int_{\omega_E} \sigma_{\mathrm{PS}} : D_{\mathrm{NC}}(\psi_E\, e_j)\, dx + \int_{\omega_E} (\psi_E e_j) \cdot \mathrm{div}_{\mathrm{NC}}\, \sigma_{\mathrm{PS}}\, dx.\end{aligned}$$

The definitions (2.4)–(2.5) show that this equals

$$\left( D_{\mathrm{NC}} \tilde{u}_{\mathrm{CR}} - \frac{f_{\mathcal{T}}}{2} \otimes (\bullet - \mathrm{mid}(\mathcal{T})) - \tilde{p}_{\mathrm{CR}}\, I_{2\times 2},\ D_{\mathrm{NC}}(\psi_E\, e_j) \right)_\Omega - (f_{\mathcal{T}}, \psi_E e_j)_\Omega .$$

The discrete nonconforming problem (2.3) and the fact that $\int_T (x - \mathrm{mid}(T)) dx = 0$ for any triangle $T$ eventually prove that this vanishes. Hence $|E|[\sigma_{\mathrm{PS}}(j)\, \nu_E]_E = 0$ and so $\sigma_{\mathrm{PS}} \in H(\mathrm{div}, \Omega; \mathbb{R}^{2\times 2})$.

Since $\mathrm{div}_{\mathrm{NC}}\, \tilde{u}_{\mathrm{CR}} = 0$, since $\int_T (\bullet - \mathrm{mid}(T))\, dx = 0$ for all $T \in \mathcal{T}$, and since $\int_\Omega \tilde{p}_{\mathrm{CR}}\, dx = 0$, the definitions prove $\sigma_{\mathrm{PS}} \in \mathrm{PS}(\mathcal{T})$.

In order to show that $(\sigma_{\mathrm{PS}}, u_{\mathrm{PS}})$ solves (2.2), (2.4)–(2.5) imply

$$\begin{aligned}&(\mathrm{dev}\, \sigma_{\mathrm{PS}}, \mathrm{dev}\, \tau_{PS})_\Omega \\ &= (D_{\mathrm{NC}} \tilde{u}_{\mathrm{CR}}, \tau_{\mathrm{PS}})_\Omega - \left( \mathrm{dev} \left( \frac{f_{\mathcal{T}}}{2} \otimes (\bullet - \mathrm{mid}(\mathcal{T})) \right),\ \tau_{\mathrm{PS}} \right)_\Omega \\ &= -\, (\tilde{u}_{\mathrm{CR}}, \mathrm{div}\, \tau_{\mathrm{PS}})_\Omega + \langle \tilde{u}_{\mathrm{CR}}, \tau_{\mathrm{PS}}\, \nu \rangle\ - \left( \mathrm{dev} \left( \frac{f_{\mathcal{T}}}{2} \otimes (\bullet - \mathrm{mid}(\mathcal{T})) \right),\ \tau_{\mathrm{PS}} - \Pi_{\mathcal{T}} \tau_{\mathrm{PS}} \right)_\Omega \\ &= -\, (\Pi_{\mathcal{T}} \tilde{u}_{\mathrm{CR}}, \mathrm{div}\, \tau_{\mathrm{PS}})_\Omega + \langle g, \tau_{\mathrm{PS}}\, \nu \rangle \\ &\qquad - \left(\mathrm{dev}\, (f_{\mathcal{T}} \otimes (\bullet - \mathrm{mid}(\mathcal{T})))\, ,\ (\mathrm{div}\, \tau_{\mathrm{PS}} \otimes (\bullet - \mathrm{mid}(\mathcal{T})))\right)_\Omega / 4 \\ &= -\, (\Pi_{\mathcal{T}} \tilde{u}_{\mathrm{CR}} + \Pi_{\mathcal{T}} (\mathrm{dev}(f_{\mathcal{T}} \otimes (\bullet - \mathrm{mid}(\mathcal{T})))\, (\bullet - \mathrm{mid}(\mathcal{T})))/4, \mathrm{div}\, \tau_{\mathrm{PS}})_\Omega\ + \langle g, \tau_{\mathrm{PS}}\, \nu \rangle .\end{aligned}$$

This is the first equality in (2.2). Since the piecewise divergence equals the distributional divergence for any $H(\mathrm{div}, \Omega)$ function, $\mathrm{div}\, \sigma_{\mathrm{PS}} = -f_{\mathcal{T}}$ proves the second equality in (2.2). ∎

Throughout the paper, the oscillations of the data $f \in L^2(\Omega; \mathbb{R}^2)$ with respect to some subset $\omega \subset \Omega$ read as

$$\mathrm{osc}\,(f, \omega) := |\omega|^{1/2} \| f - f_\omega \|_{L^2(\omega)} \qquad \text{with } f_\omega := |\omega|^{-1} \int_\omega f dx$$

and, for any $\mathcal{F} \subset \mathcal{T}$,

$$\mathrm{osc}^2(f, \mathcal{F}) := \sum_{T \in \mathcal{F}} \mathrm{osc}^2(f, T).$$

For the data $g \in H^1(\Omega;\mathbb{R}^2)$ with $\int_{\partial\Omega} g\cdot\nu\,ds = 0$ such that the piecewise derivative $\partial g/\partial s$ of $g$ along any $E \in \mathcal{E}(\partial\Omega)$ exists in $L^2(E)$, the oscillations of $\partial g/\partial s$ with respect to some edge $E \in \mathcal{E}(\partial\Omega)$ read as

$$\operatorname{osc}^2(\partial g/\partial s, E) := \min_{\gamma_E\in P_1(E;\mathbb{R}^2)} |E|\|\partial g/\partial s - \gamma_E\|^2_{L^2(E)}.$$

The total oscillations read as

$$\operatorname{osc}^2(\partial g/\partial s, \mathcal{E}(\partial\Omega)) := \sum_{E\in\mathcal{E}(\partial\Omega)} \operatorname{osc}^2(\partial g/\partial s, E).$$

The residual-based error estimator from [14] reads, for $T \in \mathcal{T}$, as

$$\eta^2(T) := \operatorname{osc}^2(f,T) + |T|\,\|\operatorname{curl}(\operatorname{dev}\sigma_{\mathrm{PS}})\|^2_{L^2(T)} + |T|^{1/2} \sum_{E\in\mathcal{E}(T)} \|[\operatorname{dev}\sigma_{\mathrm{PS}}]_E\tau_E\|^2_{L^2(E)}$$

(with slightly different but equivalent weights) and

$$\eta^2 := \eta^2(\mathcal{T}) := \sum_{T\in\mathcal{T}} \eta^2(T).$$

It is already shown in [14] that $\eta$ is reliable and efficient up to data oscillations. This work considers the following refined efficiency result for a modified definition of the oscillations $\operatorname{osc}(\partial g/\partial s, \mathcal{E}(\partial\Omega))$. For $g \in H^1(\Omega;\mathbb{R}^2)$ such that $g|_E \in H^1(E;\mathbb{R}^2)$ for all $E \in \mathcal{E}(\partial\Omega)$ (written $g \in H^1(\mathcal{E}(\partial\Omega);\mathbb{R}^2)$), let $\gamma_E \in P_1(E;\mathbb{R}^2)$ be the $L^2$ best approximation of $(\partial g/\partial s)|_E$, and let $\operatorname{osc}(\partial g/\partial s, E) := |E|^{1/2}\|(\partial g/\partial s) - \gamma_E\|_{L^2(E)}$ and $\operatorname{osc}^2(\partial g/\partial s, \mathcal{E}(\partial\Omega)) = \sum_{E\in\mathcal{E}(\partial\Omega)} \operatorname{osc}^2(\partial g/\partial s, E)$.

The velocity variable and its approximation merely play the role of a Lagrange multiplier and appear to be of minor relevance. The a posteriori error analysis is indeed free of the velocity.

THEOREM 2.2 (efficiency and reliability of $\eta$). *The reliability and efficiency of $\eta$ hold in the sense that*

$$\begin{aligned}(1/C_{\mathrm{rel}})\|\operatorname{dev}(\sigma - \sigma_{\mathrm{PS}})\|^2_{L^2(\Omega)} &\le \eta^2\\ &\le C_{\mathrm{eff}}\big(\|\operatorname{dev}(\sigma-\sigma_{\mathrm{PS}})\|^2_{L^2(\Omega)} + \operatorname{osc}^2(f,\mathcal{T}) + \operatorname{osc}^2(\partial g/\partial s, \mathcal{E}(\partial\Omega))\big).\end{aligned}$$

*Proof.* The assertion is essentially contained in [14] with different oscillations of the boundary data. To complete the proof of the presented version, it suffices to verify

$$\||E|^{1/2}[\operatorname{dev}(\sigma_{\mathrm{PS}})\tau_E]_E\|_{L^2(E)} \lesssim \|\operatorname{dev}(\sigma-\sigma_{\mathrm{PS}})\|_{L^2(\omega_E)} + \operatorname{osc}(\partial g/\partial s, E)$$

for $E \in \mathcal{E}(\partial\Omega)$. Let $b_E$ be the quadratic edge bubble function of a boundary edge $E \in \mathcal{E}(\partial\Omega)$ defined as the product of the two affine nodal basis functions associated with the two nodes of $E$. With the triangle inequality and an equivalence of norms argument, the jump terms on the boundary are estimated as

$$\begin{aligned}|E|\,\|[\operatorname{dev}(\sigma_{\mathrm{PS}})]_E\tau_E\|^2_{L^2(E)} &= |E|\,\|(\partial g/\partial s) - \operatorname{dev}(\sigma_{\mathrm{PS}})\tau_E\|^2_{L^2(E)}\\ &\lesssim |E|\,\|b_E^{1/2}((\partial g/\partial s) - \operatorname{dev}(\sigma_{\mathrm{PS}})\tau_E)\|^2_{L^2(E)} + \operatorname{osc}^2(\partial g/\partial s, E)\\ &= |E|\,\|b_E^{1/2}\operatorname{dev}(\sigma-\sigma_{\mathrm{PS}})\tau_E\|^2_{L^2(E)} + \operatorname{osc}^2(\partial g/\partial s, E).\end{aligned}$$

An integration by parts on $T = \overline{\omega}_E$, the Cauchy inequality, and the stability properties $\|b_E\|_{L^\infty(T)} \approx 1 \approx |E|\|\nabla b_E\|_{L^\infty(T)}$ yield

$$\begin{aligned}|E|\ \|b_E^{1/2}\operatorname{dev}(\sigma-\sigma_{\mathrm{PS}})\tau_E\|^2_{L^2(E)} &= |E|\int_{\partial T}(\operatorname{dev}(\sigma-\sigma_{\mathrm{PS}})\tau_E)\cdot(b_E\operatorname{dev}(\sigma-\sigma_{\mathrm{PS}}))\tau\,ds\\ &\lesssim \|\operatorname{dev}(\sigma-\sigma_{\mathrm{PS}})\|^2_{L^2(T)} + |E|^2\|\operatorname{curl}(\operatorname{dev}(\sigma-\sigma_{\mathrm{PS}}))\|^2_{L^2(T)}\\ &\lesssim \|\operatorname{dev}(\sigma-\sigma_{\mathrm{PS}})\|^2_{L^2(T)} + |T|\ \|\operatorname{curl}(\operatorname{dev}(\sigma_{\mathrm{PS}}))\|^2_{L^2(T)}.\end{aligned}$$

The efficiency of $|T|^{1/2}\|\operatorname{curl}(\operatorname{dev}(\sigma_{\mathrm{PS}}))\|_{L^2(T)}$ from [14] concludes the proof. ∎

## 3. Adaptive algorithm and main result.

This section is devoted to the adaptive pseudostress finite element method and its optimality in terms of approximation classes.

### 3.1. Apsfem.

This subsection presents an optimal adaptive algorithm Apsfem with an error estimator based on triangles.

**Input:** Initial coarse triangulation $\mathcal{T}_0$ with refinement edges $RE(\mathcal{T}_0)$, $0 < \theta < \theta_0 \leq 1$.

**Loop:** For $\ell = 0, 1, 2, \dots$

**Solve** problem (2.2) with respect to the regular triangulation $\mathcal{T}_\ell$ into triangles with discrete velocity $u_\ell \in P_0(\mathcal{T}_\ell;\mathbb{R}^2)$ and discrete stress $\sigma_\ell \in \mathrm{PS}(\mathcal{T}_\ell)$.

**Estimate** $\eta_\ell^2 := \sum_{T\in\mathcal{T}_\ell}\eta_\ell^2(T)$ with

$$\eta_\ell^2(T) := \operatorname{osc}^2(f,T) + |T|\|\operatorname{curl}(\operatorname{dev}\sigma_\ell)\|^2_{L^2(T)} \tag{3.1}$$

$$+\,|T|^{1/2}\sum_{E\in\mathcal{E}(T)}\|[\operatorname{dev}(\sigma_\ell)\tau_E]_E\|^2_{L^2(E)}. \tag{3.2}$$

**Mark** a minimal subset $\mathcal{M}_\ell \subset \mathcal{T}_\ell$ of triangles with

$$\theta\eta_\ell^2 \leq \eta_\ell^2(\mathcal{M}_\ell) := \sum_{T\in\mathcal{M}_\ell}\eta_\ell^2(T). \tag{3.3}$$

**Refine** $\mathcal{M}_\ell$ in $\mathcal{T}_\ell$ with newest-vertex-bisection (NVB) of Figure 3.1 and generate a regular triangulation $\mathcal{T}_{\ell+1}$.

**Output:** Sequence of triangulations $(\mathcal{T}_\ell)_\ell$ and discrete solutions $(u_\ell, \sigma_\ell)_\ell$.

*Remark* 3.1. The refinement edge $RE : \mathcal{T}_0 \to \mathcal{E}$, with $RE(T) \in \mathcal{E}(T)$ for any $T \in \mathcal{T}_0$, is fixed for the initial triangulation $\mathcal{T}_0$. The configuration of the refinement edges in triangles which are refined is depicted in Figure 3.1. The result of Refine $\mathcal{T}_{\ell+1}$ is the smallest shape-regular refinement of $\mathcal{T}_\ell$ without hanging nodes using NVB, where at least the refinement edges of the triangles in $\mathcal{M}_\ell$ are refined; cf. [3, 5, 31]. Up to rotations, all admissible refinements of a triangle $T \in \mathcal{T}_\ell$ are depicted in Figure 3.1 and depend on the set of its edges $\mathcal{E}_\ell(T)$ that have to be refined.

*Remark* 3.2. Given an initial triangulation $\mathcal{T}_0$, a triangulation $\mathcal{T}_\ell$ is called an admissible triangulation if there exist regular triangulations $\mathcal{T}_0, \mathcal{T}_1, \dots, \mathcal{T}_\ell$ such that, for $j = 1, \dots, \ell$, each $\mathcal{T}_j$ is generated from $\mathcal{T}_{j-1}$ using only the refinements from Figure 3.1.

*Remark* 3.3. There is no need for the inner node property. In fact, bisec5 can be used but does not need to be.

*Remark* 3.4. Throughout the rest of this paper $(\mathcal{T}_\ell)_\ell$ denotes a sequence of regular triangulations of $\Omega$ and $\mathcal{E}_\ell$ denotes the set of edges of $\mathcal{T}_\ell$. To link the notation from section 2 to that of section 3, the $L^2$ projection onto the piecewise constant functions with respect to the triangulation $\mathcal{T}_\ell$ is denoted by $\Pi_\ell := \Pi_{\mathcal{T}_\ell}$, and $f_\ell := f_{\mathcal{T}_\ell} := \Pi_\ell f$.

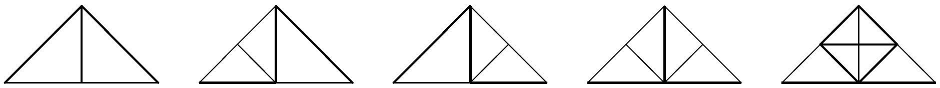


Fig. 3.1. *Possible refinements of a triangle T in one level using NVB. The thick lines indicate the refinement edges of the new triangles.*

The piecewise derivative with respect to $\mathcal{T}_\ell$ is denoted by $D_\ell$. For any $K \in \mathcal{T}_\ell$ define $\mathcal{T}_{\ell+m}(K) := \{T \in \mathcal{T}_{\ell+m} \,|\, T \subset K\}$ and for any $\mathcal{F} \subset \mathcal{T}_\ell$

$$\eta_\ell^2(\mathcal{F}) := \sum_{T\in\mathcal{F}} \eta_\ell^2(T).$$

**3.2. Approximation class.** The definition of quasi-optimal convergence is based on the concept of approximation classes. For $s > 0$, let

$$\mathcal{A}'_s := \big\{(\sigma, f, g) \in H(\mathrm{div}, \Omega; \mathbb{R}^{2\times 2})/\mathbb{R} \times L^2(\Omega; \mathbb{R}^2) \times \big(H^1(\Omega; \mathbb{R}^2) \cap H^1(\mathcal{E}(\partial\Omega); \mathbb{R}^2)\big) \,\big|\, |(\sigma, f, g)|_{\mathcal{A}'_s} < \infty\big\}$$

with $|(\sigma, f, g)|_{\mathcal{A}'_s}$

$$:= \sup_{N\in\mathbb{N}} N^s \inf_{\mathcal{T}\in\mathbb{T}(N)} \left(\|\mathrm{dev}(\sigma - \sigma_{\mathcal{T}})\|^2_{L^2(\Omega)} + \mathrm{osc}^2(f, \mathcal{T}) + \mathrm{osc}^2\left(\frac{\partial g}{\partial s}, \mathcal{E}(\partial\Omega)\right)\right)^{1/2}.$$

In the infimum, $\mathcal{T}$ runs through all admissible triangulations $\mathbb{T}(N)$ that are refined from $\mathcal{T}_0$ by NVB (cf. Figure 3.1) of [3, 32] and satisfy $|\mathcal{T}| - |\mathcal{T}_0| \le N$.

**3.3. Quasi optimality.** The main theorem of this paper states optimal convergence rates of the algorithm Apsfem and will be proven in section 5. Let $C_{\mathrm{eff}}$, $C_{\mathrm{drel}}$, and $C_{\mathrm{qo}}$ denote the constants from Theorems 2.2, 5.1, and 4.2, and let $(\mathcal{T}_\ell)_\ell$ be the sequence of triangulations generated by Apsfem with discrete velocities $(u_\ell)_\ell$ and stresses $(\sigma_\ell)_\ell$ from (2.2).

Theorem 3.5 (quasi-optimal convergence). *For any bulk parameter* $0 < \theta < \theta_0 := 1/(2C_{\mathrm{eff}}(C_{\mathrm{drel}} + C_{\mathrm{qo}} + 2))$ *and* $(\sigma, f, g) \in \mathcal{A}'_s$, Apsfem *generates sequences of triangulations* $(\mathcal{T}_\ell)_\ell$ *and discrete solutions* $(u_\ell, \sigma_\ell)_\ell$ *of optimal rate of convergence in the sense that*

$$(|\mathcal{T}_\ell| - |\mathcal{T}_0|)^s \Big(\|\mathrm{dev}(\sigma - \sigma_\ell)\|^2_{L^2(\Omega)} + \mathrm{osc}^2(f, \mathcal{T}_\ell) + \mathrm{osc}^2(\partial g/\partial s, \mathcal{E}_\ell(\partial\Omega))\Big)^{1/2} \lesssim |(\sigma, f, g)|_{\mathcal{A}'_s}.$$

Some remarks on the error terms and the approximation classes are in order before the proof follows in sections 4–5.

**3.4. Equivalence of approximation classes.** The tr-dev-div lemma [7, Proposition 3.1 in section IV.3] states for any function, like $\sigma - \sigma_\ell$, in $H(\mathrm{div}, \Omega; \mathbb{R}^{2\times 2})/\mathbb{R}$ that

$$\|\mathrm{tr}(\sigma - \sigma_\ell)\|_{L^2(\Omega)} \lesssim \|\mathrm{dev}(\sigma - \sigma_\ell)\|_{L^2(\Omega)} + \|\mathrm{div}(\sigma - \sigma_\ell)\|_{H^{-1}(\Omega)}.$$

(The proof of this stronger version is an obvious modification of the proof of [7].) Since $\mathrm{div}(\sigma - \sigma_\ell) = f_\ell - f$ is perpendicular to $P_0(\mathcal{T}; \mathbb{R}^2)$, some piecewise Poincaré inequality leads to

$$\|\mathrm{div}(\sigma - \sigma_\ell)\|_{H^{-1}(\Omega)} \leq \mathrm{osc}(f, \mathcal{T}).$$

This and the orthogonal splitting of matrices into isochoric and deviatoric parts prove that

$$\|\sigma - \sigma_\ell\|_{L^2(\Omega)} \leq \|\mathrm{dev}(\sigma - \sigma_\ell)\|_{L^2(\Omega)} + \mathrm{osc}(f, \mathcal{T}_\ell).$$

Since the converse is obvious, $\|\mathrm{dev}(\sigma - \sigma_\ell)\|_{L^2(\Omega)} \leq \|\sigma - \sigma_\ell\|_{L^2(\Omega)}$, it follows that

$$\begin{aligned}\|\sigma - \sigma_\ell\|^2_{L^2(\Omega)} &+ \mathrm{osc}^2(f, \mathcal{T}_\ell) + \mathrm{osc}^2(g, \mathcal{E}(\partial\Omega)) \\ &\approx \|\mathrm{dev}(\sigma - \sigma_\ell)\|^2_{L^2(\Omega)} + \mathrm{osc}^2(f, \mathcal{T}_\ell) + \mathrm{osc}^2(g, \mathcal{E}(\partial\Omega)).\end{aligned}$$

In other words, the approximation class $\mathcal{A}'_s$ from subsection 3.2 and the approximation class $\mathcal{A}_s$ from the introduction are identical,

$$\mathcal{A}_s = \mathcal{A}'_s \quad \text{(with equivalent norms)}.$$

Therefore, Theorem 3.5 implies the quasi-optimality result (1.2) of the introduction (and is even equivalent).

## 4. Contraction property.

This section is devoted to the proof of the contraction property of some convex combination of estimator, error, and data oscillations. The first step is the error estimator reduction property which follows as in [18].

THEOREM 4.1 (estimator reduction property). *There exist constants* $0 < \rho_1 < 1$ *and* $\Lambda > 0$ *such that for an admissible refinement* $\mathcal{T}_{\ell+1}$ *of* $\mathcal{T}_\ell$ *generated by* APSFEM *with bulk parameter* $0 < \theta \leq 1$ *and discrete solutions* $\sigma_\ell \in \mathrm{PS}(\mathcal{T}_\ell)$ *and* $\sigma_{\ell+1} \in \mathrm{PS}(\mathcal{T}_{\ell+1})$ *it holds that*

$$\eta^2_{\ell+1} \leq \rho_1\, \eta^2_\ell + \Lambda\, \|\mathrm{dev}(\sigma_{\ell+1} - \sigma_\ell)\|^2_{L^2(\Omega)}. \tag{4.1}$$

*Proof.* The main arguments of the proof will be an inverse estimate [6, p. 112]

$$\|\mathrm{curl}\,\mathrm{dev}(\sigma_{\ell+1} - \sigma_\ell)\|_{L^2(T)} \lesssim |T|^{-1/2}\, \|\mathrm{dev}(\sigma_{\ell+1} - \sigma_\ell)\|_{L^2(T)}$$

and a trace inequality [6, p. 282] (for an edge $E$ of a triangle $T$)

$$\begin{aligned}\|[\mathrm{dev}(\sigma_{\ell+1} - \sigma_\ell)]_E\, \tau_E\|_{L^2(E)} &\lesssim |T|^{-1/4}\, \|\mathrm{dev}(\sigma_{\ell+1} - \sigma_\ell)\|_{L^2(T)} \\ &\quad + |T|^{1/4}\, \|D\,\mathrm{dev}(\sigma_{\ell+1} - \sigma_\ell)\|_{L^2(T)} \\ &\lesssim |T|^{-1/4}\, \|\mathrm{dev}(\sigma_{\ell+1} - \sigma_\ell)\|_{L^2(T)}.\end{aligned}$$

The triangle inequality shows, for $T \in \mathcal{T}_{\ell+1} \cap \mathcal{T}_\ell$ and $0 < \delta < \infty$, that

$$\begin{aligned}\eta^2_{\ell+1}(T) &= |T|\, \|f - f_{\ell+1}\|_{L^2(T)} \\ &\quad + |T|\, \|\mathrm{curl}\,\mathrm{dev}\,\sigma_{\ell+1}\|^2_{L^2(T)} + |T|^{1/2} \sum_{E \in \mathcal{E}_{\ell+1}(T)} \|[\mathrm{dev}\ \sigma_{\ell+1}]_E\, \tau_E\|^2_{L^2(E)} \\ &\leq |T|\, \|f - f_\ell\|_{L^2(T)} + (1+\delta)\, |T|\, \|\mathrm{curl}\,\mathrm{dev}\,\sigma_\ell\|^2_{L^2(T)} \\ &\quad + (1+\delta)\, |T|^{1/2} \sum_{E \in \mathcal{E}_\ell(T)} \|[\mathrm{dev}\ \sigma_\ell]_E\, \tau_E\|^2_{L^2(E)} \\ &\quad + (1 + 1/\delta)\, |T|\, \|\mathrm{curl}\,\mathrm{dev}(\sigma_{\ell+1} - \sigma_\ell)\|^2_{L^2(T)} \\ &\quad + (1 + 1/\delta)\, |T|^{1/2} \sum_{E \in \mathcal{E}_\ell(T)} \|[\mathrm{dev}(\sigma_{\ell+1} - \sigma_\ell)]_E\, \tau_E\|^2_{L^2(E)}.\end{aligned}$$

The aforementioned inverse and trace inequalities lead to some generic constant $C \approx 1$ with

$$\eta_{\ell+1}^2(T) \le (1+\delta)\,\eta_\ell^2(T) + C(1+(1/\delta))\|\mathrm{dev}(\sigma_{\ell+1}-\sigma_\ell)\|_{L^2(T)}^2.$$

For $K \in \mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}$ and with $|T| \le (1/2)\,|K|$ for $T \subset K$ the above arguments show

$$\begin{aligned}\eta_{\ell+1}^2(\mathcal{T}_{\ell+1}(K)) \le (1+\delta) \sum_{T\subset K} \Bigg( &\frac{1}{2}\,|K|\,\|f-f_T\|_{L^2(T)} + \frac{1}{2}\,|K|\,\|\mathrm{curl}\,\mathrm{dev}\,\sigma_\ell\|_{L^2(T)}^2 \\ &+ \frac{1}{\sqrt{2}}\,|K|^{1/2} \sum_{E\in\mathcal{E}_{\ell+1}(T)} \|[\mathrm{dev}\ \sigma_\ell]_E\,\tau_E\|_{L^2(E)}^2 \Bigg) \\ &+ \left(1+\frac{1}{\delta}\right) \sum_{T\subset K} \Bigg( |T|\,\|\mathrm{curl}\,\mathrm{dev}(\sigma_{\ell+1}-\sigma_\ell)\|_{L^2(T)}^2 \\ &+ |T|^{1/2} \sum_{E\in\mathcal{E}_{\ell+1}(T)} \|[\mathrm{dev}(\sigma_{\ell+1}-\sigma_\ell)]_E\,\tau_E\|_{L^2(E)}^2 \Bigg).\end{aligned}$$

Since $[\mathrm{dev}\ \sigma_\ell]_E = 0$ for $E \in \mathcal{E}_{\ell+1}(\mathrm{int}(K))$, $K \in \mathcal{T}_\ell$,

$$\eta_{\ell+1}^2(\mathcal{T}_{\ell+1}(K)) \le (1+\delta)\,\frac{1}{\sqrt{2}}\eta_\ell^2(K) + C(1+1/\delta)\|\mathrm{dev}(\sigma_{\ell+1}-\sigma_\ell)\|_{L^2(K)}^2.$$

The sum over all $T \in \mathcal{T}_{\ell+1}$ yields

$$\begin{aligned}\eta_{\ell+1}^2 &= \eta_{\ell+1}^2(\mathcal{T}_{\ell+1}\cap\mathcal{T}_\ell) + \eta_{\ell+1}^2(\mathcal{T}_{\ell+1}\setminus\mathcal{T}_\ell) \\ &\le (1+\delta)\,\left(\eta_\ell^2(\mathcal{T}_{\ell+1}\cap\mathcal{T}_\ell) + \frac{1}{\sqrt{2}}\eta_\ell^2(\mathcal{T}_\ell\setminus\mathcal{T}_{\ell+1})\right) \\ &\quad + (1+1/\delta)\,C\,\|\mathrm{dev}(\sigma_{\ell+1}-\sigma_\ell)\|_{L^2(\Omega)}^2.\end{aligned}$$

The bulk criterion $\theta\,\eta_\ell^2 \le \eta_\ell^2(\mathcal{T}_\ell\setminus\mathcal{T}_{\ell+1})$ leads to

$$\begin{aligned}\eta_\ell^2(\mathcal{T}_{\ell+1}\cap\mathcal{T}_\ell) + \frac{1}{\sqrt{2}}\,\eta_\ell^2(\mathcal{T}_\ell\setminus\mathcal{T}_{\ell+1}) &= \eta_\ell^2 - \left(1-\frac{1}{\sqrt{2}}\right)\,\eta_\ell^2(\mathcal{T}_\ell\setminus\mathcal{T}_{\ell+1}) \\ &\le \left(1-\theta\left(1-\frac{1}{\sqrt{2}}\right)\right)\,\eta_\ell^2.\end{aligned}$$

For $\delta$ sufficiently small, $\rho_1 := (1+\delta)\,(1-\theta\,(1-1/\sqrt{2}))$ and $\Lambda = (1+1/\delta)\,C$ satisfy (4.1). ∎

THEOREM 4.2 (quasi orthogonality). *There exists a positive constant $C_{\mathrm{qo}} \approx 1$ which solely depends on $\mathcal{T}_0$ such that, for any refinement $\mathcal{T}_{\ell+m}$ of $\mathcal{T}_\ell$, the exact solution $\sigma$ and the discrete solutions $\sigma_{\ell+m}$ and $\sigma_\ell$ (with respect to $\mathcal{T}_{\ell+m}$ and $\mathcal{T}_\ell$) satisfy*

$$\begin{aligned}&|\,(\mathrm{dev}(\sigma-\sigma_{\ell+m}), \mathrm{dev}(\sigma_{\ell+m}-\sigma_\ell))_\Omega| \\ &\qquad \le C_{\mathrm{qo}}^{1/2}\,\|\mathrm{dev}(\sigma-\sigma_{\ell+m})\|_{L^2(\Omega)}\ \mathrm{osc}(f,\mathcal{T}_\ell\setminus\mathcal{T}_{\ell+m}).\end{aligned} \tag{4.2}$$

*Proof.* Let $\sigma_{\ell+m}^*$ be the solution of the intermediate problem on $\mathcal{T}_{\ell+m}$ where the right-hand side $f$ in (2.2) is replaced by the piecewise constant projection $f_\ell := \Pi_\ell f$. Since $\sigma_{\ell+m}^* - \sigma_\ell \in \mathrm{PS}(\mathcal{T}_{\ell+m})$ and

$$\mathrm{div}(\sigma_{\ell+m}^* - \sigma_\ell) = 0 \quad \text{a.e. in } \Omega,$$

the problems (2.1)–(2.2) yield

$$\big(\operatorname{dev}(\sigma-\sigma_{\ell+m}),\operatorname{dev}(\sigma^*_{\ell+m}-\sigma_\ell)\big)_\Omega=0.$$

This orthogonality implies

$$\begin{aligned}(\operatorname{dev}(\sigma-\sigma_{\ell+m}),\operatorname{dev}(\sigma_{\ell+m}-\sigma_\ell))_\Omega&=\big(\operatorname{dev}(\sigma-\sigma_{\ell+m}),\operatorname{dev}(\sigma_{\ell+m}-\sigma^*_{\ell+m})\big)_\Omega\\&\le\|\operatorname{dev}(\sigma-\sigma_{\ell+m})\|_{L^2(\Omega)}\,\|\operatorname{dev}(\sigma_{\ell+m}-\sigma^*_{\ell+m})\|_{L^2(\Omega)}.\end{aligned}$$

Let $(\tilde u_{\mathrm{CR},\ell+m},\tilde p_{\mathrm{CR},\ell+m})$ and $(\tilde u^*_{\mathrm{CR},\ell+m},\tilde p^*_{\mathrm{CR},\ell+m})$ be the Crouzeix–Raviart solutions of problem (2.3) with right-hand sides $f_{\ell+m}$ and $f_\ell$. By Theorem 2.1, $\sigma_{\ell+m}$ and $\sigma^*_{\ell+m}$ can be represented as

$$\begin{aligned}\sigma_{\ell+m}&=D_{\ell+m}\tilde u_{\mathrm{CR},\ell+m}+1/2\,f_{\ell+m}\otimes(\bullet-\operatorname{mid}(\mathcal T_{\ell+m}))-\tilde p_{\mathrm{CR},\ell+m}\,I_{2\times2},\\\sigma^*_{\ell+m}&=D_{\ell+m}\tilde u^*_{\mathrm{CR},\ell+m}+1/2\,f_\ell\otimes(\bullet-\operatorname{mid}(\mathcal T_{\ell+m}))-\tilde p^*_{\mathrm{CR},\ell+m}\,I_{2\times2}.\end{aligned}$$

Therefore, the triangle inequality reveals

$$\begin{aligned}\|\operatorname{dev}(\sigma_{\ell+m}-\sigma^*_{\ell+m})\|_{L^2(\Omega)}&\le\|D_{\ell+m}(\tilde u_{\mathrm{CR},\ell+m}-\tilde u^*_{\mathrm{CR},\ell+m})\|_{L^2(\Omega)}\\&\quad+1/2\,\|\operatorname{dev}((f_{\ell+m}-f_\ell)\otimes(\bullet-\operatorname{mid}(\mathcal T_{\ell+m})))\|_{L^2(\Omega)}.\end{aligned}\tag{4.3}$$

Since $\tilde u_{\mathrm{CR},\ell+m}\in g_{\mathrm{CR}}+\mathrm{CR}^1_0(\mathcal T_{\ell+m};\mathbb R^2)$ and $\tilde u^*_{\mathrm{CR},\ell+m}\in g_{\mathrm{CR}}+\mathrm{CR}^1_0(\mathcal T_{\ell+m};\mathbb R^2)$ are the Crouzeix–Raviart solutions and $\int_K(f_{\ell+m}-f_\ell)\,dx=0$ for all $K\in\mathcal T_\ell$, one obtains

$$\begin{aligned}\|D_{\ell+m}(\tilde u_{\mathrm{CR},\ell+m}-\tilde u^*_{\mathrm{CR},\ell+m})\|^2_{L^2(\Omega)}&=\big(f_{\ell+m}-f_\ell,\tilde u_{\mathrm{CR},\ell+m}-\tilde u^*_{\mathrm{CR},\ell+m}\big)_\Omega\\&\lesssim\operatorname{osc}(f_{\ell+m}-f_\ell,\mathcal T_\ell\setminus\mathcal T_{\ell+m})\,\|D_{\ell+m}(\tilde u_{\mathrm{CR},\ell+m}-\tilde u^*_{\mathrm{CR},\ell+m})\|_{L^2(\Omega)}.\end{aligned}\tag{4.4}$$

Since $|\bullet-\operatorname{mid}(\mathcal T_{\ell+m})|\lesssim|T|^{1/2}$ and $\int_K(f_{\ell+m}-f_\ell)\,dx=0$ for all $K\in\mathcal T_\ell$, it holds that

$$\begin{aligned}&\|\operatorname{dev}((f_{\ell+m}-f_\ell)\otimes(\bullet-\operatorname{mid}(\mathcal T_{\ell+m})))\|_{L^2(\Omega)}\\&\quad\le\|f_{\ell+m}-f_\ell\|_{L^2(\Omega)}\,\|\bullet-\operatorname{mid}(\mathcal T_{\ell+m})\|_{L^2(\Omega)}\\&\quad\lesssim\operatorname{osc}(f_{\ell+m}-f_\ell,\mathcal T_\ell\setminus\mathcal T_{\ell+m}).\end{aligned}\tag{4.5}$$

The combination of (4.3)–(4.5) shows

$$\|\operatorname{dev}(\sigma_{\ell+m}-\sigma^*_{\ell+m})\|_{L^2(\Omega)}\lesssim\operatorname{osc}(f_{\ell+m}-f_\ell,\mathcal T_\ell\setminus\mathcal T_{\ell+m})\le\operatorname{osc}(f,\mathcal T_\ell\setminus\mathcal T_{\ell+m}).\tag{4.6}$$ ◻

THEOREM 4.3 (contraction property). *There exist positive constants $\beta$, $\gamma$, and $0<\rho_2<1$ such that, for any $\ell\in\mathbb N_0$, the solution $\sigma_\ell$ and error estimator $\eta_\ell$ with respect to the triangulation $\mathcal T_\ell$ of* APSFEM,

$$\xi^2_\ell:=\eta^2_\ell+\beta\,\|\operatorname{dev}(\sigma-\sigma_\ell)\|^2_{L^2(\Omega)}+\gamma\operatorname{osc}^2(f,\mathcal T_\ell),$$

*satisfies*

$$\xi^2_{\ell+1}\le\rho_2\xi^2_\ell.\tag{4.7}$$

*Proof.* The estimator reduction property (4.1) and the quasi orthogonality (4.2) yield

$$\begin{aligned}\eta^2_{\ell+1}\le\rho_1\eta^2_\ell+\Lambda\big(&\|\operatorname{dev}(\sigma-\sigma_\ell)\|^2_{L^2(\Omega)}-\|\operatorname{dev}(\sigma-\sigma_{\ell+1})\|^2_{L^2(\Omega)}\\&+2C_{\mathrm{qo}}^{1/2}\|\operatorname{dev}(\sigma-\sigma_{\ell+1})\|_{L^2(\Omega)}\operatorname{osc}(f,\mathcal T_\ell\backslash\mathcal T_{\ell+1})\big).\end{aligned}$$

For any $0 < \lambda < 1$ it holds that

$$\begin{aligned} &2C_{\mathrm{qo}}^{1/2}\|\operatorname{dev}(\sigma-\sigma_{\ell+1})\|_{L^2(\Omega)}\operatorname{osc}(f,\mathcal{T}_\ell\backslash\mathcal{T}_{\ell+1}) \\ &\qquad \le \lambda\|\operatorname{dev}(\sigma-\sigma_{\ell+1})\|_{L^2(\Omega)}^2 + \frac{4C_{\mathrm{qo}}}{\lambda}\operatorname{osc}^2(f,\mathcal{T}_\ell\backslash\mathcal{T}_{\ell+1}). \end{aligned}$$

The combination of the previous estimates for $\beta := \Lambda(1-\lambda)$ leads to

$$\begin{aligned} &\beta\|\operatorname{dev}(\sigma-\sigma_{\ell+1})\|_{L^2(\Omega)}^2 + \eta_{\ell+1}^2 \\ &\qquad \le \rho_1\eta_\ell^2 + \frac{\beta}{1-\lambda}\|\operatorname{dev}(\sigma-\sigma_\ell)\|_{L^2(\Omega)}^2 + \frac{4\Lambda C_{\mathrm{qo}}}{\lambda}\operatorname{osc}^2(f,\mathcal{T}_\ell\backslash\mathcal{T}_{\ell+1}). \end{aligned}$$

Bisection implies $2\operatorname{osc}^2(f,\mathcal{T}_{\ell+1}) \le \operatorname{osc}^2(f,\mathcal{T}_\ell)$ and, hence,

$$\operatorname{osc}^2(f,\mathcal{T}_\ell\backslash\mathcal{T}_{\ell+1}) \le 2\operatorname{osc}^2(f,\mathcal{T}_\ell) - 2\operatorname{osc}^2(f,\mathcal{T}_{\ell+1}).$$

This implies for $\gamma := 8\Lambda C_{\mathrm{qo}}/\lambda$ and $\varepsilon = 2\lambda$ with $C_{\mathrm{rel}}$ from Theorem 2.2 that

$$\begin{aligned} &\eta_{\ell+1}^2 + \beta\|\operatorname{dev}(\sigma-\sigma_{\ell+1})\|_{L^2(\Omega)}^2 + \gamma\operatorname{osc}^2(f,\mathcal{T}_{\ell+1}) \\ &\le \rho_1\eta_\ell^2 + \frac{\beta}{1-\lambda}\|\operatorname{dev}(\sigma-\sigma_\ell)\|_{L^2(\Omega)}^2 + \gamma\operatorname{osc}^2(f,\mathcal{T}_\ell) \\ &\le \rho_1\eta_\ell^2 + \frac{\beta}{1-\lambda}\Big((1-\varepsilon)\|\operatorname{dev}(\sigma-\sigma_\ell)\|_{L^2(\Omega)}^2 + \varepsilon C_{\mathrm{rel}}\eta_\ell^2\Big) \\ &\qquad + (\gamma-\varepsilon)\operatorname{osc}^2(f,\mathcal{T}_\ell) + \varepsilon\eta_\ell^2 \\ &\le \max\left\{\rho_1 + \varepsilon\left(1+\frac{C_{\mathrm{rel}}\beta}{1-\lambda}\right), \frac{1-\varepsilon}{1-\lambda}, 1-\frac{\varepsilon}{\gamma}\right\} \\ &\qquad \big(\eta_\ell^2 + \beta\|\operatorname{dev}(\sigma-\sigma_\ell)\|_{L^2(\Omega)}^2 + \gamma\operatorname{osc}^2(f,\mathcal{T}_\ell)\big). \end{aligned}$$

For sufficiently small $\lambda$ this leads to

$$\rho_2 := \max\left\{\rho_1 + \varepsilon\left(1+\frac{C_{\mathrm{rel}}\beta}{1-\lambda}\right), \frac{1-\varepsilon}{1-\lambda}, 1-\frac{\varepsilon}{\gamma}\right\} < 1. \qquad \square$$

## 5. Proof of optimality.

The key argument in the proof of Theorem 3.5 is the discrete reliability.

THEOREM 5.1 (discrete reliability). *There exists a constant $C_{\mathrm{drel}} \approx 1$ which depends solely on $\mathcal{T}_0$ such that any refinement $\mathcal{T}_{\ell+m}$ of $\mathcal{T}_\ell$ with respective solutions $\sigma_{\ell+m}$ and $\sigma_\ell$ satisfies*

$$\|\operatorname{dev}(\sigma_{\ell+m}-\sigma_\ell)\|_{L^2(\Omega)}^2 \le C_{\mathrm{drel}}\eta_\ell^2(\mathcal{T}_\ell\backslash\mathcal{T}_{\ell+m}).$$

One key argument in the proof of Theorem 5.1 is some novel Helmholtz decomposition of piecewise constant deviatoric matrices which is proven in [16]. Let $\mathbb{R}^{2\times 2}_{dev} := \{A \in \mathbb{R}^{2\times 2} \,|\, A = \operatorname{dev} A\}$ denote the trace-free $2\times 2$ matrices, and let

$$X(\mathcal{T}) := \left\{v_C \in C(\Omega;\mathbb{R}^2)\cap P_1(\mathcal{T};\mathbb{R}^2)\,\Big|\, \int_\Omega v_C\,dx = 0 \text{ and } \int_\Omega \operatorname{curl} v_C\,dx = 0\right\}.$$

THEOREM 5.2 (discrete Helmholtz decomposition [16, Theorem 3.2]). *The direct decomposition*

$$P_0(\mathcal{T};\mathbb{R}^{2\times 2}_{dev}) = D_{\mathrm{NC}}\,\mathrm{Z}_{\mathrm{CR}}(\mathcal{T}) \oplus \operatorname{dev}\operatorname{Curl} X(\mathcal{T})$$

*is orthogonal in $L^2(\Omega;\mathbb{R}^{2\times 2}_{dev})$.* $\square$

*Proof of Theorem* 5.1. Let $\sigma^*_{\ell+m}$ denote the intermediate solution on the mesh $\mathcal{T}_{\ell+m}$ with right-hand side $f_\ell$ as in the proof of Theorem 4.2 and recall (4.6), namely

$$\|\operatorname{dev}(\sigma_{\ell+m} - \sigma^*_{\ell+m})\|^2_{L^2(\Omega)} \lesssim \operatorname{osc}^2(f, \mathcal{T}_\ell \backslash \mathcal{T}_{\ell+m}).$$

By the triangle inequality,

$$\|\operatorname{dev}(\sigma_{\ell+m} - \sigma_\ell)\|_{L^2(\Omega)} \leq \|\operatorname{dev}(\sigma_{\ell+m} - \sigma^*_{\ell+m})\|_{L^2(\Omega)} + \|\operatorname{dev}(\sigma^*_{\ell+m} - \sigma_\ell)\|_{L^2(\Omega)},$$

it remains to analyze the term $\|\operatorname{dev}(\sigma^*_{\ell+m} - \sigma_\ell)\|_{L^2(\Omega)}$. Since the difference $\sigma^*_{\ell+m} - \sigma_\ell$ is divergence-free and hence piecewise constant, Theorem 5.2 guarantees the existence of $z_{\ell+m} \in \mathrm{Z}_{\mathrm{CR}}(\mathcal{T}_{\ell+m})$ and $\beta_{\ell+m} \in X(\mathcal{T}_{\ell+m})$ such that

$$\operatorname{dev}(\sigma^*_{\ell+m} - \sigma_\ell) = D_{\ell+m} z_{\ell+m} + \operatorname{dev} \operatorname{Curl} \beta_{\ell+m}.$$

The orthogonality of the decomposition (followed by a piecewise integration by parts) reveals

$$\begin{aligned} \|D_{\ell+m} z_{\ell+m}\|^2_{L^2(\Omega)} &= \big(\operatorname{dev}(\sigma^*_{\ell+m} - \sigma_\ell), D_{\ell+m} z_{\ell+m}\big)_\Omega \\ &= -\big(\operatorname{div}_{\mathrm{NC}}(\sigma^*_{\ell+m} - \sigma_\ell), z_{\ell+m}\big)_\Omega = 0. \end{aligned}$$

This implies

$$\operatorname{dev}(\sigma^*_{\ell+m} - \sigma_\ell) = \operatorname{dev} \operatorname{Curl} \beta_{\ell+m}. \tag{5.1}$$

Any $\beta_\ell \in P_1(\mathcal{T}_\ell; \mathbb{R}^2) \cap C(\bar{\Omega}; \mathbb{R}^2)$ satisfies $\operatorname{Curl} \beta_\ell \in \mathrm{RT}_0(\mathcal{T}_\ell; \mathbb{R}^{2\times 2})$. Note that the discrete equation (2.2) is satisfied for all test functions in $\mathrm{RT}_0(\mathcal{T}_{\ell+m}; \mathbb{R}^{2\times 2})$. The discrete equation (2.2) for the level $\ell + m$ and $\ell$ with respective solutions $\sigma^*_{\ell+m}$ and $\sigma_\ell$ results in

$$\big(\operatorname{Curl} \beta_\ell, \operatorname{dev}(\sigma^*_{\ell+m} - \sigma_\ell)\big)_\Omega = 0.$$

The same argument on the level $\ell + m$ for the test function $\operatorname{Curl}(\beta_{\ell+m} - \beta_\ell)$ leads to

$$\big(\operatorname{dev} \sigma^*_{\ell+m}, \operatorname{Curl}(\beta_{\ell+m} - \beta_\ell)\big)_\Omega = \langle g, \operatorname{Curl}(\beta_{\ell+m} - \beta_\ell)\nu \rangle.$$

The combination of the previous identities reads as

$$\|\operatorname{dev}(\sigma^*_{\ell+m} - \sigma_\ell)\|^2_{L^2(\Omega)} = \langle g, \operatorname{Curl}(\beta_{\ell+m} - \beta_\ell)\nu \rangle - (\operatorname{dev} \sigma_\ell, \operatorname{Curl}(\beta_{\ell+m} - \beta_\ell))_\Omega.$$

Define $\beta_\ell$ as the Scott–Zhang quasi interpolant [30] of $\beta_{\ell+m}$ such that $\beta_{\ell+m} = \beta_\ell$ on all $E \in \mathcal{E}_{\ell+m} \cap \mathcal{E}_\ell$. The piecewise integration by parts and the stability and approximation property of the Scott–Zhang quasi-interpolation operator imply that

$$\begin{aligned} &\|\operatorname{dev}(\sigma^*_{\ell+m} - \sigma_\ell)\|^2_{L^2(\Omega)} = -\langle g, D(\beta_{\ell+m} - \beta_\ell)\tau \rangle - (\operatorname{dev} \sigma_\ell, \operatorname{Curl}(\beta_{\ell+m} - \beta_\ell))_\Omega \\ &= -\sum_{E \in \mathcal{E}_\ell \backslash \mathcal{E}_{\ell+m}} \int_E [\operatorname{dev} \sigma_\ell]_E \, \tau_E \cdot (\beta_{\ell+m} - \beta_\ell)\, ds + (\operatorname{curl} \operatorname{dev} \sigma_\ell, \beta_{\ell+m} - \beta_\ell)_\Omega \\ &\lesssim \eta_\ell(\mathcal{T}_\ell \backslash \mathcal{T}_{\ell+m}) \|D\beta_{\ell+m}\|_{L^2(\Omega)}. \end{aligned} \tag{5.2}$$

The second Korn inequality [6, p. 316] plus some algebra leads to

$$\|D\beta_{\ell+m}\|_{L^2(\Omega)} \lesssim \|\operatorname{dev} \operatorname{Curl} \beta_{\ell+m}\|_{L^2(\Omega)}.$$

The combination with (5.1)–(5.2) concludes the proof. ∎

The remaining part of this section adopts the strategy from [18, 31] to the present situation of section 3 with the output of the sequence of pseudostress approximations $(\sigma_\ell)_\ell$ and triangulations $(\mathcal{T}_\ell)_\ell$.

In the *first step* of the proof, set

$$\xi_\ell := \eta_\ell^2 + \beta\,\|\mathrm{dev}(\sigma-\sigma_\ell)\|^2_{L^2(\Omega)} + \gamma\,\mathrm{osc}^2(f,\mathcal{T}_\ell)$$

as in Theorem 4.3. Without loss of generality, the pathological case $\xi_0 = 0$ can be excluded. Choose $0 < \tau \le |(\sigma,f,g)|^2_{\mathcal{A}'_s}/\xi_0^2$, and set $\varepsilon^2(\ell) := \tau\xi_\ell^2$. Choose minimal $N(\ell) \in \mathbb{N}$ with the property

$$|(\sigma,f,g)|_{\mathcal{A}'_s} \le \varepsilon(\ell)\,N(\ell)^s. \tag{5.3}$$

Claim A. *Then it holds that*

$$N(\ell) \le 2\;|(\sigma,f,g)|^{1/s}_{\mathcal{A}'_s}\,\varepsilon(\ell)^{-1/s} \quad \textit{for } \ell \in \mathbb{N}_0.$$

*Proof of Claim* A. For $N(\ell) = 1$, (5.3) implies by the contraction property (4.7) that

$$|(\sigma,f,g)|^2_{\mathcal{A}'_s} \le \varepsilon(\ell)^2 = \tau\xi_\ell^2 \le \tau\xi_0^2.$$

This implies equality $|(\sigma,f,g)|^2_{\mathcal{A}'_s} = \varepsilon(\ell)^2$. For $N(\ell) \ge 2$ the minimality of $N(\ell)$ in (5.3) yields

$$\varepsilon(\ell)\big(N(\ell)-1\big)^s < |(\sigma,f,g)|_{\mathcal{A}'_s}.$$

Therefore,

$$N(\ell) \le 2\big(N(\ell)-1\big) \le 2\,|(\sigma,f,g)|^{1/s}_{\mathcal{A}'_s}\,\varepsilon(\ell)^{-1/s}. \qquad \square$$

The definition of $|(\sigma,f,g)|_{\mathcal{A}'_s}$ as a supremum over $N$ shows for $N = N(\ell)$ that there exists some optimal triangulation $\tilde{\mathcal{T}}_\ell \in \mathbb{T}(N)$ (which is possibly not related to $\mathcal{T}_\ell$) of cardinality $|\tilde{\mathcal{T}}_\ell| \le |\mathcal{T}_0| + N(\ell)$ with approximate stress $\tilde{\sigma_\ell} \in \mathrm{PS}(\tilde{\mathcal{T}}_\ell)$ and

$$\begin{aligned}\|\mathrm{dev}(\sigma-\tilde{\sigma_\ell})\|^2_{L^2(\Omega)} + \mathrm{osc}^2(f,\tilde{\mathcal{T}}_\ell) + \mathrm{osc}^2(\partial g/\partial s,\tilde{\mathcal{E}}_\ell(\partial\Omega))&\\ \le N(\ell)^{-2s}\,|(\sigma,f,g)|^2_{\mathcal{A}'_s} \le \varepsilon(\ell)^2.&\end{aligned} \tag{5.4}$$

The overlay $\hat{\mathcal{T}}_\ell := \mathcal{T}_\ell \otimes \tilde{\mathcal{T}}_\ell$ is defined as the smallest common refinement of $\mathcal{T}_\ell$ and $\tilde{\mathcal{T}}_\ell$. It is known [18, 32] that

$$|\hat{\mathcal{T}}_\ell| - |\mathcal{T}_\ell| \le |\tilde{\mathcal{T}}_\ell| - |\mathcal{T}_0| \le N(\ell).$$

The number of triangles in $\mathcal{T}_\ell \setminus \hat{\mathcal{T}}_\ell$ can be estimated as

$$|\mathcal{T}_\ell \setminus \hat{\mathcal{T}}_\ell| \le \sum_{K\in\mathcal{T}_\ell\setminus\hat{\mathcal{T}}_\ell} \big(|\hat{\mathcal{T}}_\ell(K)| - 1\big) = |\hat{\mathcal{T}}_\ell \setminus \mathcal{T}_\ell| - |\mathcal{T}_\ell \setminus \hat{\mathcal{T}}_\ell| = |\hat{\mathcal{T}}_\ell| - |\mathcal{T}_\ell|.$$

Thus

$$|\mathcal{T}_\ell \setminus \hat{\mathcal{T}}_\ell| \le N(\ell) \le 2\,|(\sigma,f,g)|^{1/s}_{\mathcal{A}'_s}\,\varepsilon(\ell)^{-1/s}. \tag{5.5}$$

CLAIM B. *In the second step the following estimate will be established. There exists $C_1 \approx 1$ such that the stress approximation $\hat\sigma_\ell \in \mathrm{PS}(\hat{\mathcal{T}}_\ell)$ with respect to $\hat{\mathcal{T}}_\ell$ satisfies*

$$\|\mathrm{dev}(\sigma-\hat\sigma_\ell)\|^2_{L^2(\Omega)} + \mathrm{osc}^2(f,\hat{\mathcal{T}}_\ell) + \mathrm{osc}^2(\partial g/\partial s, \hat{\mathcal{E}}_\ell(\partial\Omega)) \le C_1\varepsilon^2(\ell). \tag{5.6}$$

*Proof of Claim* B. The quasi orthogonality shows

$$\begin{aligned}\|\mathrm{dev}(\sigma-\hat\sigma_\ell)\|^2_{L^2(\Omega)} &= \|\mathrm{dev}(\sigma-\tilde\sigma_\ell)\|^2_{L^2(\Omega)} - \|\mathrm{dev}(\tilde\sigma_\ell-\hat\sigma_\ell)\|^2_{L^2(\Omega)}\\ &\quad + 2\int_\Omega \mathrm{dev}(\sigma-\hat\sigma_\ell) : \mathrm{dev}(\tilde\sigma_\ell-\hat\sigma_\ell)dx\\ &\le \|\mathrm{dev}(\sigma-\tilde\sigma_\ell)\|^2_{L^2(\Omega)} - \|\mathrm{dev}(\tilde\sigma_\ell-\hat\sigma_\ell)\|^2_{L^2(\Omega)}\\ &\quad + 2C_{\mathrm{qo}}\,\mathrm{osc}^2(f,\tilde{\mathcal{T}}_\ell\setminus\hat{\mathcal{T}}_\ell) + \frac12\|\mathrm{dev}(\sigma-\hat\sigma_\ell)\|^2_{L^2(\Omega)}.\end{aligned}$$

Hence

$$\begin{aligned}\frac12\|\mathrm{dev}(\sigma-\hat\sigma_\ell)\|^2_{L^2(\Omega)}&+\|\mathrm{dev}(\tilde\sigma_\ell-\hat\sigma_\ell)\|^2_{L^2(\Omega)}\\ &\le \|\mathrm{dev}(\sigma-\tilde\sigma_\ell)\|^2_{L^2(\Omega)} + 2C_{\mathrm{qo}}\,\mathrm{osc}^2(f,\tilde{\mathcal{T}}_\ell\setminus\hat{\mathcal{T}}_\ell).\end{aligned}$$

Equation (5.4) and the choice $C_1 := \max\{2, 4\,C_{\mathrm{qo}}+1\}$ conclude the proof. ∎

CLAIM C. *It holds that*

$$\eta_\ell \lesssim \eta_\ell(\mathcal{T}_\ell\setminus\hat{\mathcal{T}}_\ell). \tag{5.7}$$

*Proof of Claim* C. Theorem 2.2 shows

$$\frac{\eta_\ell^2}{C_{\mathrm{eff}}} \le \|\mathrm{dev}(\sigma-\sigma_\ell)\|^2_{L^2(\Omega)} + \mathrm{osc}^2(f,\mathcal{T}_\ell) + \mathrm{osc}^2(\partial g/\partial s, \mathcal{E}_\ell(\partial\Omega)). \tag{5.8}$$

The quasi orthogonality leads to

$$\begin{aligned}\|\mathrm{dev}(\sigma-\sigma_\ell)\|^2_{L^2(\Omega)} &= \|\mathrm{dev}(\sigma-\hat\sigma_\ell)\|^2_{L^2(\Omega)} + \|\mathrm{dev}(\hat\sigma_\ell-\sigma_\ell)\|^2_{L^2(\Omega)}\\ &\quad + 2\int_\Omega \mathrm{dev}(\sigma-\hat\sigma_\ell) : \mathrm{dev}(\hat\sigma_\ell-\sigma_\ell)dx\\ &\le 2\|\mathrm{dev}(\sigma-\hat\sigma_\ell)\|^2_{L^2(\Omega)} + \|\mathrm{dev}(\hat\sigma_\ell-\sigma_\ell)\|^2_{L^2(\Omega)}\\ &\quad + C_{\mathrm{qo}}\,\mathrm{osc}^2(f,\mathcal{T}_\ell\setminus\hat{\mathcal{T}}_\ell).\end{aligned}$$

This and discrete reliability from Theorem 5.1 with constant $C_{\mathrm{drel}}$ lead to

$$\|\mathrm{dev}(\sigma-\sigma_\ell)\|^2_{L^2(\Omega)} \le 2\|\mathrm{dev}(\sigma-\hat\sigma_\ell)\|^2_{L^2(\Omega)} + (C_{\mathrm{drel}}+C_{\mathrm{qo}})\eta_\ell^2(\mathcal{T}_\ell\setminus\hat{\mathcal{T}}_\ell). \tag{5.9}$$

The oscillations can be controlled by

$$\mathrm{osc}^2(f,\mathcal{T}_\ell) \le \eta_\ell^2(\mathcal{T}_\ell\setminus\hat{\mathcal{T}}_\ell) + \mathrm{osc}^2(f,\mathcal{T}_\ell\cap\hat{\mathcal{T}}_\ell) \le \eta_\ell^2(\mathcal{T}_\ell\setminus\hat{\mathcal{T}}_\ell) + \mathrm{osc}^2(f,\hat{\mathcal{T}}_\ell).$$

Since $\mathrm{osc}(\partial g/\partial s, E) \le |E|^{1/2}\|(\partial g/\partial s) - \mathrm{dev}\,\sigma_\ell\|_{L^2(E)}$, it follows that

$$\mathrm{osc}^2(\partial g/\partial s, \mathcal{E}_\ell(\partial\Omega)) \le \eta_\ell(\mathcal{T}_\ell\setminus\hat{\mathcal{T}}_\ell) + \mathrm{osc}^2(\partial g/\partial s, \hat{\mathcal{E}}_\ell(\partial\Omega)).$$

The combination of (5.6) and (5.8)–(5.9) leads to

$$\frac{\eta_\ell^2}{C_{\rm eff}} \le (C_{\rm drel} + C_{\rm qo} + 2)\eta_\ell^2(\mathcal{T}_\ell \setminus \hat{\mathcal{T}}_\ell) + 2C_1\varepsilon(\ell)^2$$
$$\le (C_{\rm drel} + C_{\rm qo} + 2)\eta_\ell^2(\mathcal{T}_\ell \setminus \hat{\mathcal{T}}_\ell) + 2\tau C_1 C_{\rm eq}\eta_\ell^2$$

with equivalence constant $C_{\rm eq}$ from $\eta_\ell^2 \le \xi_\ell^2 \le C_{\rm eq}\eta_\ell^2$. The choice of $0 < \tau < 1/(4C_{\rm eff}C_1C_{\rm eq})$ leads to

$$\eta_\ell^2 \le 2C_{\rm eff}(C_{\rm drel} + C_{\rm qo} + 2)\,\eta_\ell^2(\mathcal{T}_\ell \setminus \hat{\mathcal{T}}_\ell). \qquad \square$$

CLAIM D. *Let $C_2 \approx 1$ be such that $\eta_\ell^2 \le C_2\eta_\ell^2(\mathcal{T}_\ell \setminus \hat{\mathcal{T}}_\ell)$ for all $\ell \in \mathbb{N}_0$. Then $0 < \theta \le \theta_0 := 1/C_2$ implies*

$$(|\mathcal{T}_\ell| - |\mathcal{T}_0|)^s \Big(\|{\rm dev}(\sigma - \sigma_\ell)\|^2_{L^2(\Omega)} + {\rm osc}^2(f, \mathcal{T}_\ell) + {\rm osc}^2(\partial g/\partial s, \mathcal{E}_\ell(\partial\Omega))\Big)^{1/2}$$
$$\lesssim |(\sigma, f, g)|_{\mathcal{A}'_s}.$$

*Proof of Claim* D. MARK selects $\mathcal{M}_\ell \subset \mathcal{T}_\ell$ with minimal cardinality $|\mathcal{M}_\ell|$ such that $\theta\eta_\ell^2 \le \eta_\ell^2(\mathcal{M}_\ell)$. Since

$$\theta\eta_\ell^2 \le \theta_0\eta_\ell^2 = \eta_\ell^2/C_2 \le \eta_\ell^2(\mathcal{T}_\ell \setminus \hat{\mathcal{T}}_\ell),$$

$\mathcal{T}_\ell \setminus \hat{\mathcal{T}}_\ell$ also satisfies the bulk criterion and the minimality of $\mathcal{M}_\ell$ proves

$$|\mathcal{M}_\ell| \le |\mathcal{T}_\ell \setminus \hat{\mathcal{T}}_\ell| \le 2\,|(\sigma, f, g)|^{1/s}_{\mathcal{A}'_s}\,\varepsilon(\ell)^{-1/s} = 2\,|(\sigma, f, g)|^{1/s}_{\mathcal{A}'_s}\,\tau^{-1/(2s)}\xi_\ell^{-1/s}$$

with $\tau \approx 1$ and for all $\ell \in \mathbb{N}_0$. The theorem [3, Theorem 2.4] (see also [32, Theorem 6.1]) leads to a constant $C_{\rm BDV} \approx 1$ with

$$|\mathcal{T}_\ell| - |\mathcal{T}_0| \le C_{\rm BDV}\sum_{k=0}^{\ell-1}|\mathcal{M}_k| \le 2C_{\rm BDV}\,|(\sigma, f, g)|^{1/s}_{\mathcal{A}'_s}\,\tau^{-1/(2s)}\sum_{k=0}^{\ell-1}\xi_k^{-1/s}.$$

The contraction property (Theorem 4.3) reads as $\xi_{k+1}^2 \le \rho_2\xi_k^2$ for all $k \in \mathbb{N}_0$. Mathematical induction proves

$$\xi_\ell^2 \le \rho_2^{\ell-k}\xi_k^2 \quad \text{for } 0 \le k \le \ell.$$

Since $0 < \rho_2 < 1$ it follows that

$$\sum_{k=0}^{\ell-1}\xi_k^{-1/s} \le \xi_\ell^{-1/s}\sum_{k=0}^{\ell-1}\rho_2^{(\ell-k)/(2s)} \le \xi_\ell^{-1/s}\frac{\rho_2^{1/(2s)}}{1-\rho_2^{1/(2s)}}.$$

Altogether,

$$|\mathcal{T}_\ell| - |\mathcal{T}_0| \le 2C_{\rm BDV}\,|(\sigma, f, g)|^{1/s}_{\mathcal{A}'_s}\,\tau^{1/(2s)}\xi_\ell^{-1/s}\frac{\rho_2^{1/(2s)}}{1-\rho_2^{1/(2s)}}. \qquad \square$$

## 6. Numerical experiments.

Four benchmark examples provide numerical evidence for optimality even for large parameters $\theta$.

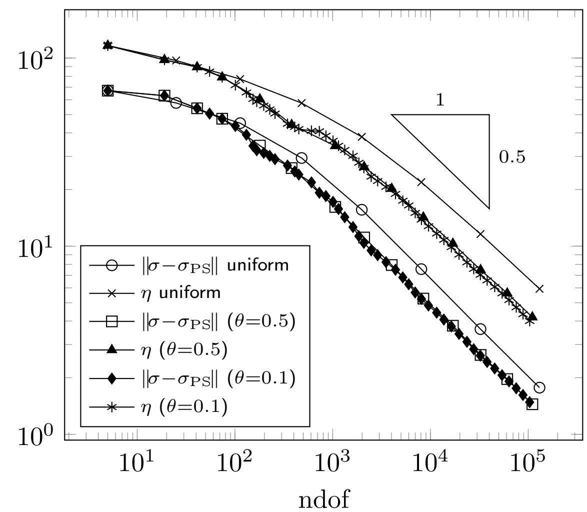


FIG. 6.1. *Convergence history for the colliding flow example.*

**6.1. Colliding flow.** On the square domain $\Omega = (-1,1) \times (-1,1)$, the exact velocity is given by $u(x,y) = (20xy^4 - 4x^5, 20x^4y - 4y^5)$ with pressure $p(x,y) = 120x^2y^2 - 20x^4 - 20y^4 - 32/6$. For this smooth example, both uniform and adaptive mesh-refinement yield optimal convergence rates (see Figure 6.1).

**6.2. L-shaped domain.** On the L-shaped domain $\Omega = (-1,1) \times (-1,1) \setminus ([0,1] \times [-1,0])$, the exact solution reads as

$$u(r,\vartheta) = \begin{pmatrix} r^\alpha((1+\alpha)\sin(\vartheta)w(\vartheta) + \cos(\vartheta)w_\vartheta(\vartheta)) \\ r^\alpha(-(1+\alpha)\cos(\vartheta)w(\vartheta) + \sin(\vartheta)w_\vartheta(\vartheta)) \end{pmatrix}$$

in polar coordinates with $\alpha = 0.54448373$, and

$$\begin{aligned} w(\vartheta) =& (\sin((1+\alpha)\vartheta)\cos(\alpha\omega))/(1+\alpha) - \cos((1+\alpha)\vartheta) \\ &- (\sin((1-\alpha)\vartheta)\cos(\alpha\omega))/(1-\alpha) + \cos((1-\alpha)\vartheta), \end{aligned}$$

and $f = 0$. Figure 6.2 shows the suboptimal convergence rate for uniform mesh-refinement and optimal convergence for adaptive mesh-refinement.

**6.3. Slit domain.** On the slit domain $\Omega = (-1,1)^2 \setminus ([0,1) \times \{0\})$, the exact velocity reads in polar coordinates as

$$u(r,\vartheta) = \frac{3\sqrt{r}}{2}\big(\cos(\vartheta/2) - \cos(3\vartheta/2), 3\sin(\vartheta/2) - \sin(3\vartheta/2)\big)$$

with pressure $p(r,\vartheta) = -6r^{-1/2}\cos(\vartheta/2)$. The suboptimal convergence rate from uniform mesh-refinement is improved towards the optimal one by the adaptive algorithm (see Figure 6.3).

**6.4. Backward-facing step.** This benchmark example considers the domain $\Omega = ((-2,8) \times (-1,1)) \setminus ([-2,0] \times [-1,0])$ from Figure 6.4. Let $f = 0$ and $g(x,y) = (0,0)$ for $-2 < x < 8$, $g(x,y) = (-y(y-1)/10, 0)$ for $x = -2$, $g(x,y) = (-(y+1)(y-1)/80, 0)$ for $x = 8$. Figure 6.5 shows the convergence history with optimal convergence rates.

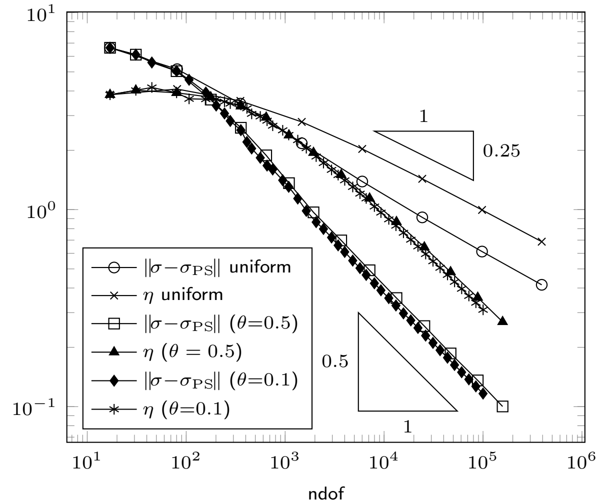


FIG. 6.2. *Convergence history for the L-shaped domain.*

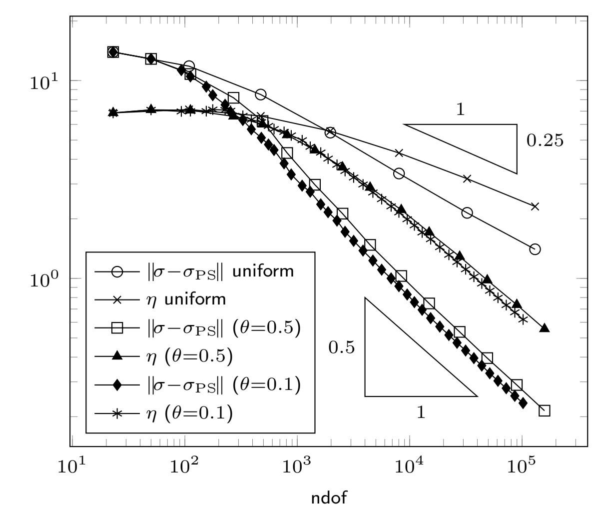


FIG. 6.3. *Convergence history for the slit domain.*

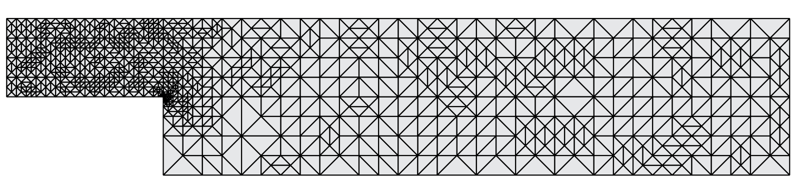

FIG. 6.4. *Adaptive mesh for the backward-facing step ($\theta = 0.1$).*

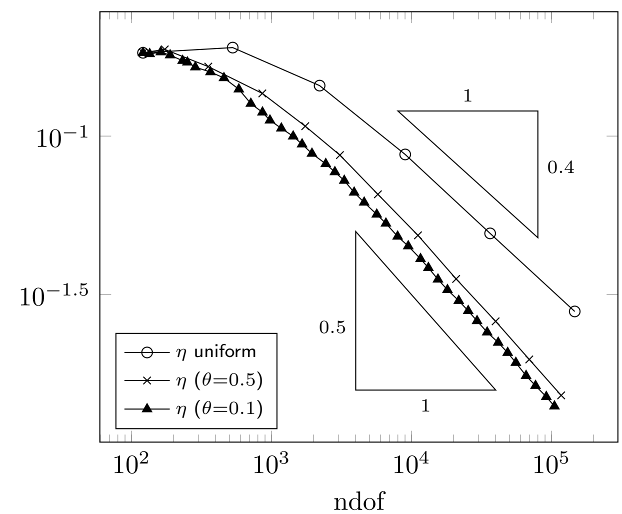


FIG. 6.5. *Convergence history for the backward-facing step.*

**Acknowledgments.** The second and third authors acknowledge the kind hospitality of the Department of Computational Science and Engineering of Yonsei University.

## REFERENCES

[1] R. BECKER AND S. MAO, *Quasi-optimality of adaptive nonconforming finite element methods for the Stokes equations*, SIAM J. Numer. Anal., 49 (2011), pp. 970–991.
[2] R. BECKER, S. MAO, AND Z. SHI, *A convergent nonconforming adaptive finite element method with quasi-optimal complexity*, SIAM J. Numer. Anal., 47 (2010), pp. 4639–4659.
[3] P. BINEV, W. DAHMEN, AND R. DEVORE, *Adaptive finite element methods with convergence rates*, Numer. Math., 97 (2004), pp. 219–268.
[4] D. BRAESS, *Finite Elements. Theory, Fast Solvers, and Applications in Solid Mechanics*, Cambridge University Press, Cambridge, UK, 2001.
[5] S. C. BRENNER AND C. CARSTENSEN, *Encyclopedia of Computational Mechanics*, John Wiley and Sons, Chichester, UK, 2004.
[6] S. C. BRENNER AND L. R. SCOTT, *The Mathematical Theory of Finite Element Methods*, 3rd ed., Texts Appl. Math. 15, Springer, New York, 2008.
[7] F. BREZZI AND M. FORTIN, *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, New York, 1991.
[8] 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.
[9] Z. CAI, C. WANG, AND S. ZHANG, *Mixed finite element methods for incompressible flow: Stationary Navier–Stokes equations*, SIAM J. Numer. Anal., 48 (2010), pp. 79–94.
[10] Z. CAI AND Y. WANG, *A multigrid method for the pseudostress formulation of Stokes problems*, SIAM J. Sci. Comput., 29 (2007), pp. 2078–2095.
[11] Z. CAI AND Y. WANG, *Pseudostress-velocity formulation for incompressible Navier-Stokes equations*, Internat. J. Numer. Methods Fluids, 63 (2010), pp. 341–356.
[12] C. CARSTENSEN AND R. H. W. HOPPE, *Convergence analysis of an adaptive nonconforming finite element method*, Numer. Math., 103 (2006), pp. 251–266.
[13] C. CARSTENSEN AND R. H. W. HOPPE, *Error reduction and convergence for an adaptive mixed finite element method*, Math. Comp., 75 (2006), pp. 1033–1042.
[14] C. CARSTENSEN, D. KIM, AND E.-J. PARK, *A priori and a posteriori pseudostress-velocity mixed finite element error analysis for the Stokes problem*, SIAM J. Numer. Anal., 49 (2011), pp. 2501–2523.

[15] C. Carstensen and E.-J. Park, *Equivalence between pseudostress and nonconforming schemes*, untitled work in progress.
[16] C. Carstensen, D. Peterseim, and H. Rabus, *Optimal adaptive nonconforming FEM for the Stokes problem*, Numer. Math., 123 (2013), pp. 291–308.
[17] C. Carstensen and H. Rabus, *An optimal adaptive mixed finite element method*, Math. Comp., 80 (2011), pp. 649–667.
[18] J. M. Cascon, C. Kreuzer, R. H. Nochetto, and K. G. Siebert, *Quasi-optimal convergence rate for an adaptive finite element method*, SIAM J. Numer. Anal., 46 (2008), pp. 2524–2550.
[19] L. Chen, M. Holst, and J. Xu, *Convergence and optimality of adaptive mixed finite element methods*, Math. Comp., 78 (2009), pp. 35–53.
[20] W. Dörfler, *A convergent adaptive algorithm for Poisson's equation*, SIAM J. Numer. Anal., 33 (1996), pp. 1106–1124.
[21] L. E. Figueroa, G. N. Gatica, and A. Márquez, *Augmented mixed finite element methods for the stationary Stokes equations*, SIAM J. Sci. Comput., 31 (2009), pp. 1082–1119.
[22] G. N. Gatica, A. Márquez, and M. A. Sánchez, *Analysis of a velocity-pressure-pseudostress formulation for the stationary Stokes equations*, Comput. Methods Appl. Mech. Engrg., 199 (2010), pp. 1064–1079.
[23] G. N. Gatica, A. Márquez, and M. A. Sánchez, *A priori and a posteriori error analyses of a velocity-pseudostress formulation for a class of quasi-Newtonian Stokes flows*, Comput. Methods Appl. Mech. Engrg., 200 (2011), pp. 1619–1636.
[24] V. Girault and P.-A. Raviart, *Finite Element Methods for Navier-Stokes Equations*, Springer Ser. Comput. Math. 5, Springer-Verlag, Berlin, 1986.
[25] J. Hu and J. Xu, *Convergence of Adaptive Conforming and Nonconforming Finite Element Methods for the Perturbed Stokes Equation*, Research Report, School of Mathematical Sciences and Institute of Mathematics, Peking University, Beijing, China, 2007; available online from www.math.pku.edu.cn:8000/var/preprint/7297.pdf.
[26] J. Hu and J. Xu, *Convergence and optimality of the adaptive nonconforming linear element method for the Stokes problem*, J. Sci. Comput., 55 (2013), pp. 125–148.
[27] C. Johnson and V. Thomée, *Error estimates for some mixed finite element methods for parabolic type problems*, RAIRO Anal. Numér., 15 (1981), pp. 41–78.
[28] L. D. Marini, *An inexpensive method for the evaluation of the solution of the lowest order Raviart–Thomas mixed method*, SIAM J. Numer. Anal., 22 (1985), pp. 493–496.
[29] H. Rabus, *A natural adaptive nonconforming FEM of quasi-optimal complexity*, Comput. Methods Appl. Math., 10 (2010), pp. 315–325.
[30] L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493.
[31] R. Stevenson, *Optimality of a standard adaptive finite element method*, Found. Comput. Math., 7 (2007), pp. 245–269.
[32] R. Stevenson, *The completion of locally refined simplicial partitions created by bisection*, Math. Comp., 77 (2008), pp. 227–241.