Optimal adaptive nonconforming FEM for the Stokes problem

Carstensen Carstensen · Daniel Peterseim ·
Hella Rabus

Abstract This paper presents an optimal nonconforming adaptive finite element algorithm and proves its quasi-optimal complexity for the Stokes equations with respect to natural approximation classes. The proof does not explicitly involve the pressure variable and follows from a novel discrete Helmholtz decomposition of deviatoric functions.

Mathematics Subject Classification (2000) Primary 65N12 · 65N15 · 65N30 · 65N50 · 65Y20

1 Introduction

The convergence and the optimality of conforming adaptive finite element methods (FEM) for Poisson-type problems have recently been established and we refer to the

---

Carsten Carstensen 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 R31-2008-000-10049-0. Daniel Peterseim was supported by the DFG research center Matheon 'Mathematics in the key technologies'. Hella Rabus was supported by the DFG research group 797 'Analysis and Computation of Microstructure in Finite Plasticity'.

C. Carstensen · D. Peterseim (✉) · H. Rabus
Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany
e-mail: peterseim@mathematik.hu-berlin.de

C. Carstensen
e-mail: cc@mathematik.hu-berlin.de

H. Rabus
e-mail: rabus@mathematik.hu-berlin.de

C. Carstensen
Department of Computational Science and Engineering, Yonsei University, Seoul 120-749, Korea

landmarks [2,12,18,23,26]. The extension to nonconforming methods for the same class of problems has been established thereafter in [7,10,20,25] based on the concept of quasi-orthogonality.

Although convergence and optimality of adaptive (nonconforming) finite element methods are well understood in the elliptic setting, the literature regarding convergence and analysis of adaptive methods for the Stokes problem is still rare. One reason might be the lack of a concept of (quasi-)orthogonality which is a key tool in the existing analysis of adaptive nonconforming methods for the Poisson problem [10,20]. Early work [6] for the Stokes problem even suggested an Uzawa algorithm with Poisson solves to circumvent this difficulty.

This paper concerns the optimality of the adaptive mesh-refinement in the nonconforming Crouzeix–Raviart finite element method (NCFEM) for the Stokes equations [14] based on the a posteriori error estimator of [17]. The first convergence and optimality result for an adaptive NCFEM for the Stokes problem was included in the technical report [21] (see [22] for a published version) while similar convergence and optimality analysis appeared recently in [5]. However, there is a gap in the complexity analysis of [5] (the estimate in line 23 on page 983 in the last step of the proof of Lemma 5.2 involves some constant $C = C(H/h)$ which depends on the ratio of the two mesh-sizes and so *cannot* be used in the proof of Theorem 5.4 where $h \ll H$ indicates an arbitrary refinement of H over many levels). In contrast to [5,21,22], the present work bases on a novel discrete Helmholtz decomposition of piecewise constant deviatoric matrices. Helmholtz decompositions have been used already in [24,25] to analyze adaptive nonconforming methods for the Poisson problem and in [16] for linear elasticity. Moreover, our analysis comes without the error in the pressure variable which makes it very brief and neat compared with [5,21,22].

The adaptive NCFEM is based on sequences of shape-regular triangulations $\mathcal{T}_\ell$, discrete spaces $V_\ell := \mathrm{CR}_0^1(\mathcal{T}_\ell) \times \mathrm{CR}_0^1(\mathcal{T}_\ell)$ and $Q_\ell := P_0(\mathcal{T}_\ell) \cap L_0^2(\Omega)$, and the discrete bilinear forms

$$a_{NC(\ell)}(u_\ell, v_\ell) = \int_\Omega \mathrm{D}_\ell u_\ell : \mathrm{D}_\ell v_\ell \, \mathrm{d}x \quad \text{and}$$

$$b_{NC(\ell)}(u_\ell, q_\ell) = \int_\Omega q_\ell \mathrm{div}_\ell u_\ell \, \mathrm{d}x$$

for $u_\ell, v_\ell \in V_\ell$ and $q_\ell \in Q_\ell$. Given some right-hand side $F \in V_\ell^*$, the discrete solution $(u_\ell, p_\ell) \in V_\ell \times Q_\ell$ satisfies, for all $(v_\ell, q_\ell) \in V_\ell \times Q_\ell$, that

$$a_{NC(\ell)}\,(u_\ell, v_\ell) + b_{NC(\ell)}(v_\ell, p_\ell) + b_{NC(\ell)}(u_\ell, q_\ell) = F(v_\ell).$$

The corresponding adaptive algorithm is based on a bulk criterion for the contribution

$$\eta_\ell^2(T) := |T|\,\|f\|\ {L^2(T)}^2 + |T|^{1/2} \sum_{E\in\mathcal{E}(T)} \|[\partial u_\ell/\partial s]_E\|_{L^2(E)}^2$$

for a triangle T with area $|T|$ and edges $E \in \mathcal{E}(T)$; $[\partial u_\ell/\partial s]_E$ denotes the jump of the tangential components of the piecewise constant gradient $\mathrm{D}_\ell u_\ell$ along any edge $E \in \mathcal{E}_\ell$.

The proposed algorithm Acrfem of Subsect. 2.2 is quasi-optimally convergent with respect to some natural approximation class $\mathcal{A}_s$ and its semi-norm $|\cdot|_{\mathcal{A}_s}$ for some $s > 0$ in the following sense. Given the exact velocity u and exact pressure p, the generated sequences of triangulations $(\mathcal{T}_\ell)_\ell$ and discrete solutions $(u_\ell, p_\ell)_\ell$ satisfy on any level $\ell \in \mathbb{N}_0$, that the number of triangles $|\mathcal{T}_\ell|$ of $\mathcal{T}_\ell$ is bounded like

$$\begin{aligned} |\mathcal{T}_\ell| - |\mathcal{T}_0| \le\ & C_{\mathrm{opt}} |(u, p, f)|_{\mathcal{A}_s}{}^{-1/s} \\ & \times \left(\|\mathrm{D}u - D_\ell u_\ell\|^2_{L^2(\Omega)} + \|p - p_\ell\|^2_{L^2(\Omega)} + \operatorname{osc}^2(f, \mathcal{T}_\ell)\right)^{-1/(2s)}. \end{aligned} \tag{1.1}$$

The convergence rate s is optimal in the sense that $|(u, p, f)|_{\mathcal{A}_s}$ is the infimum of all upper bounds of

$$N^s \inf_{|\mathcal{T}|-|\mathcal{T}_0|\le N} \left(\|Du - D_\mathcal{T} u_\mathcal{T}\|^2_{L^2(\Omega)} + \|p - p_\mathcal{T}\|^2_{L^2(\Omega)} + \operatorname{osc}^2(f, \mathcal{T})\right)^{1/2}$$

over all $N \in \mathbb{N}$ where $\mathcal{T}$ is an arbitrary admissible triangulation refined from $\mathcal{T}_0$ (cf. Remark 2.1 for an explanation) with less than or equal to $N + |\mathcal{T}_0|$ triangles and with associated discrete solution $(u_\mathcal{T}, p_\mathcal{T})$. The computed triangulation $\mathcal{T}_\ell$ is optimal up to the factor $C_{\mathrm{opt}} \lesssim 1$ and hence called quasi-optimal.

This paper is organized in the following way. The weak formulation, the algorithm, the definition of the approximation class $\mathcal{A}_s$ and the main theorem of this paper are stated in Sect. 2. The a posteriori error estimator is analyzed in Sect. 3 with the discrete Helmholtz decomposition and discrete reliability. Section 4 is devoted to the proof of the contraction property and its fundamentals such as estimator reduction and quasi-orthogonality. Section 5 concludes the proof of the main theorem on robust optimal convergence rate.

Throughout this paper, standard notation on Lebesgue and Sobolev spaces and their norms is employed; $⨍$ denotes the integral mean and $L_0^2(\Omega) := \{v \in L^2(\Omega) |\ ⨍_\Omega v = 0\}$. The formula $A \lesssim B$ represents $A \le CB$ for some mesh-independent, positive generic constant C; $A \approx B$ abbreviates $A \lesssim B \lesssim A$. By convention, all generic constants do not depend on the mesh-size h_ℓ but they may depend on the fixed coarse triangulation $\mathcal{T}_0$ and its interior angles. The 2×2 unit matrix is denoted by I_2 and the Euclid product of matrices by colon, e.g., $A : B = \sum_{j,k=1}^2 A_{jk} B_{jk}$ for $A, B \in \mathbb{R}^{2\times 2}$; $\operatorname{tr}(A) := A : I_2$ names the trace of A and $\operatorname{dev}(A) := A - \frac{1}{2}\operatorname{tr}(A) I_2$ the deviatoric part of A. The measure $|\cdot|$ is context-sensitive and refers to the number of elements of some finite set or the length of an edge or the area of some domain.

2 Model Stokes problem

2.1 Weak formulation and discretization

The two-dimensional motion of a viscous incompressible fluid in a polygonal simply connected Lipschitz domain $\Omega \subset \mathbb{R}^2$ can be modeled by a velocity field $u : \Omega \to \mathbb{R}^2$ and a pressure distribution $p : \Omega \to \mathbb{R}$ which satisfy the Stokes equations under the standard no-slip boundary condition

$$\left.\begin{aligned} -\Delta u + \nabla p &= f \\ \operatorname{div} u &= 0 \end{aligned}\right\} \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \partial\Omega, \tag{2.1}$$

where $f \in L^2(\Omega; \mathbb{R}^2)$ is a given force density.

Given bilinear and linear forms

$$a(u, v) := \int_\Omega \mathrm{D}u : \mathrm{D}v \,\mathrm{d}x, \quad b(u, q) := \int_\Omega q \operatorname{div} u \,\mathrm{d}x, \quad F(v) := \int_\Omega v \cdot f \,\mathrm{d}x$$

for $u, v \in V := H_0^1(\Omega; \mathbb{R}^2)$, $q \in Q := L_0^2(\Omega)$, the weak formulation of (2.1) seeks a pair $(u, p) \in V \times Q$ that satisfies the mixed variational problem

$$\begin{aligned} a(u, v) + b(v, p) &= F(v) && \text{for all } v \in V; \\ b(u, q) &= 0 && \text{for all } q \in Q. \end{aligned} \tag{2.2}$$

Let $\mathcal{T}_\ell$ be some regular triangulation of Ω into closed triangles $T \in \mathcal{T}_\ell$ with piecewise constant mesh-size h_ℓ. The set $\mathcal{E}_\ell$ contains all edges of $\mathcal{T}_\ell$, $\mathcal{E}_\ell(\Omega)$ all interior edges and $\mathcal{E}_\ell(\partial\Omega)$ all edges on the boundary; the set of edges of a triangle T is denoted with $\mathcal{E}(T)$. Moreover, let $\mathcal{N}_\ell$ be the set of all nodes in $\mathcal{T}_\ell$ and $\mathcal{E}_\ell(z)$ the set of edges that share the node $z \in \mathcal{N}_\ell$. For interior edges, $[\cdot]_E := \cdot|_{T_+} - \cdot|_{T_-}$ denotes the jump across the edge $E = T_+ \cap T_-$ shared by the two elements $T_\pm \in \mathcal{T}_\ell$, and $\omega_E := \operatorname{int}(T_+ \cup T_-)$. If $E \in \mathcal{E}_\ell(\partial\Omega)$ the jump $[\cdot]_E := \cdot|_{T_+}$ is the restriction to the one element $T_+ \in \mathcal{T}_\ell(E)$ and $\omega_E := \operatorname{int}(T_+)$. In addition, for any edge $E \in \mathcal{E}_\ell(\Omega)$, $\operatorname{mid}(E)$ names its midpoint and $\nu_E = \nu_{T_+}$ is the unit normal vector exterior to T_+ along E and τ_E is the unit tangential vector along $E|_{T_+}$.

Throughout the paper, the discrete spaces read

$$\begin{aligned} P_0(\mathcal{T}_\ell) &:= \left\{ v_\ell \in L^2(\Omega) \,\middle|\, v_\ell|_T \text{ is constant for all } T \in \mathcal{T}_\ell \right\}, \\ P_1(\mathcal{T}_\ell) &:= \left\{ v_\ell \in L^2(\Omega) \,\middle|\, v_\ell|_T \text{ is affine for all } T \in \mathcal{T}_\ell \right\}, \\ \mathrm{CR}^1(\mathcal{T}_\ell) &:= \left\{ v_\ell \in P_1(\mathcal{T}_\ell) \,\middle|\, \begin{array}{l} v_\ell \text{ is continuous in } \operatorname{mid}(E) \\ \text{for all } E \in \mathcal{E}_\ell(\Omega) \end{array} \right\}, \\ \mathrm{CR}_0^1(\mathcal{T}_\ell) &:= \left\{ v_\ell \in \mathrm{CR}^1(\mathcal{T}_\ell) \,|\, v_\ell(\operatorname{mid}(E)) = 0 \text{ for all } E \in \mathcal{E}_\ell(\partial\Omega) \right\}, \\ V_\ell &:= V(\mathcal{T}_\ell) := \mathrm{CR}_0^1(\mathcal{T}_\ell) \times \mathrm{CR}_0^1(\mathcal{T}_\ell), \\ Q_\ell &:= Q(\mathcal{T}_\ell) := P_0(\mathcal{T}_\ell) \cap L_0^2(\Omega). \end{aligned}$$

Let D_ℓ and div_ℓ denote the piecewise action of the gradient and the divergence with respect to the triangulation $\mathcal{T}_\ell$. Let

$$a_{NC(\ell)}(u_\ell, v_\ell) := \int_\Omega \mathrm{D}_\ell u_\ell : \mathrm{D}_\ell v_\ell \, \mathrm{d}x \quad \text{for all } u_\ell, v_\ell \in V_\ell$$

define the discrete energy scalar product on V_ℓ and let

$$b_{NC(\ell)}(v_\ell, q_\ell) := \int_\Omega q_\ell \mathrm{div}_\ell v_\ell \, \mathrm{d}x \quad \text{for all } v_\ell \in V_\ell,\ q_\ell \in Q_\ell$$

define the discrete counterpart of the bounded bilinear form b.

The discrete Friedrichs inequality [9, (10.6.14)] shows that $a_{NC(\ell)}$ is positive definite, and hence, defines a norm $|||\cdot|||_{NC(\ell)} := \|\mathrm{D}_\ell \cdot\|_{L^2(\Omega)}$ on V_ℓ. Moreover, the inf-sup stability of b yields discrete inf-sup stability of $b_{NC(\ell)}$ [14]. Thus, there exists a unique discrete solution $(u_\ell, p_\ell) \in V_\ell \times Q_\ell$ with

$$a_{NC(\ell)}(u_\ell, v_\ell) + b_{NC(\ell)}(v_\ell, p_\ell) = F(v_\ell) \quad \text{for all } v_\ell \in V_\ell; \tag{2.3a}$$
$$b_{NC(\ell)}(u_\ell, q_\ell) = 0 \quad \text{for all } q_\ell \in Q_\ell. \tag{2.3b}$$

With the present choice of $Q_\ell = P_0(\mathcal{T}_\ell) \cap L^2_0(\Omega)$, the discrete conservation of volume (2.3a) implies $\mathrm{div}_\ell u_\ell = 0$. Set

$$Z_\ell := Z(\mathcal{T}_\ell) := \left\{ v_\ell \in V_\ell \;\middle|\; \mathrm{div}_\ell v_\ell = 0 \right\}$$

as the subspace of discrete divergence free velocities in V_ℓ. Then, the solution $u_\ell \in Z_\ell$ of the discrete system (2.3) uniquely solves

$$a_{NC(\ell)}(u_\ell, z_\ell) = F(z_\ell) \quad \text{for all } z_\ell \in Z_\ell.$$

2.2 ACRFEM

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

Input: Initial coarse triangulation $\mathcal{T}_0$, $0 < \theta < \theta_0 \leq 1$.
Loop: For $\ell = 0, 1, \ldots$

Update ℓ

SOLVE problem (2.3) with respect to $\mathcal{T}_\ell$, discrete velocity u_ℓ and discrete pressure p_ℓ.

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

$$\eta_\ell^2(T) := |T| \, \|f\|_{L^2(T)}^2 + |T|^{1/2} \sum_{E \in \mathcal{E}(T)} \|[\partial u_\ell / \partial s]_E\|_{L^2(E)}^2 .$$

MARK a minimal subset $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ of triangles with

$$\theta \eta_\ell^2 \leq \eta_\ell^2 (\mathcal{M}_\ell) := \sum_{T \in \mathcal{M}_\ell} \eta_\ell^2(T). \tag{2.4}$$

REFINE $\mathcal{M}_\ell$ in $\mathcal{T}_\ell$ with Newest-Vertex-Bisection (NVB) of Figure 2.1 and generate a regular triangulation $\mathcal{T}_{\ell+1}$.

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

Remark 2.1 The result of REFINE is the smallest shape-regular refinement $\mathcal{T}_{\ell+1}$ of $\mathcal{T}_\ell$ without hanging nodes using NVB, where at least the refinement edges of the marked elements $E(\mathcal{M}_\ell)$ are refined, cf. [1]. Up to rotations, all admissible refinements of a triangle $T \in \mathcal{T}_\ell$ are depicted in Fig. 1 and depend on the set of its edges $\mathcal{E}(T)$ that have to be refined. The refinement edge $E(T)$ of each triangle is accented in Fig. 1. In case that all edges $\mathcal{E}(T)$ have to be refined either *bisec3*(T) or *bisec5*(T) can be applied.

2.3 Approximation class and main result

Here and throughout the paper, $f \in L^2(\Omega; \mathbb{R}^2)$, and the oscillations of f with respect to some subset $\mathcal{F} \subseteq \mathcal{T}_\ell$ read

$$\mathrm{osc}_\ell^2 := \mathrm{osc}^2(f, \mathcal{T}_\ell) \text{ with } \mathrm{osc}^2(f, \mathcal{F}) := \sum_{T \in \mathcal{F}} \mathrm{osc}^2(f, T)$$

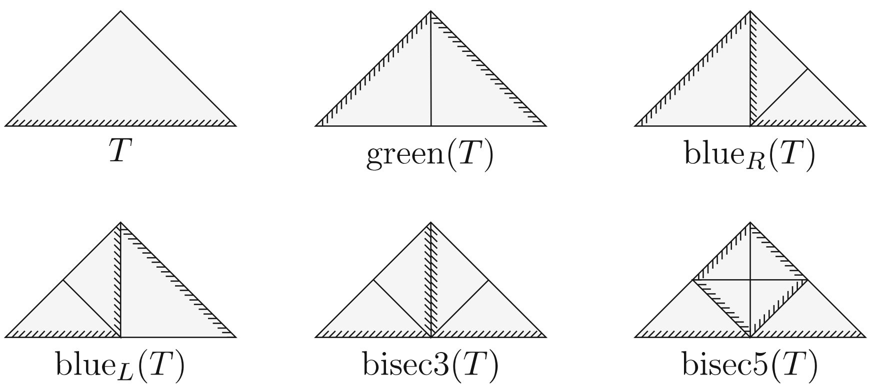

Fig. 1 Possible refinements of a triangle T in one level using NVB

and, for any subset $\omega \subseteq \Omega$,

$$\operatorname{osc}(f, \omega) := |\omega|^{1/2} \, \|f - f_\omega\|_{L^2(\omega)} \text{ with } f_\omega := \fint_\omega f \, \mathrm{d}x := |\omega|^{-1} \int_\omega f \, \mathrm{d}x \, .$$

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

$$\mathcal{A}_s := \left\{ (u, p, f) \in H_0^1(\Omega; \mathbb{R}^2) \times L_0^2(\Omega) \times L^2(\Omega; \mathbb{R}^2) \,\middle|\, |(u, p, f)|_{\mathcal{A}_s} < \infty \right\}$$

with $|(u, p, f)|_{\mathcal{A}_s}$ defined by

$$\sup_{N \in \mathbb{N}} \left(N^s \inf_{|\mathcal{T}| - |\mathcal{T}_0| \le N} \left(|||u - u_{\mathcal{T}}|||_{NC(\mathcal{T})}^2 + \|p - p_{\mathcal{T}}\|_{L^2(\Omega)}^2 + \operatorname{osc}^2(f, \mathcal{T}) \right)^{1/2} \right).$$

In the infimum, $\mathcal{T}$ runs through all admissible triangulations that are refined from $\mathcal{T}_0$ by NVB (cf. Fig. 1) and that satisfy $|\mathcal{T}| - |\mathcal{T}_0| \le N$.

Remark 2.2 For the Poisson problem, [13] shows that in the definition of the approximation class above the error of the Crouzeix–Raviart approximation might be replaced by the best approximation error (see also [19]). By similar techniques it can be shown that for any solution (u, p) of (2.2) with right-hand side $f \in L^2(\Omega; \mathbb{R}^2)$

$$\begin{aligned} |(u, p, f)|_{\mathcal{A}_s} \approx \sup_{N \in \mathbb{N}} \Bigg(N^s & \inf_{|\mathcal{T}| - |\mathcal{T}_0| \le N} \; \inf_{(v, q) \in V(\mathcal{T}) \times Q(\mathcal{T})} \Big(|||u - v|||_{NC(\mathcal{T})}^2 \\ & + \|p - q\|_{L^2(\Omega)}^2 + \operatorname{osc}^2(f, \mathcal{T}) \Big)^{1/2} \Bigg). \end{aligned}$$

Hence, the approximation class $\mathcal{A}_s$ might be replaced by the standard one [12].

The main theorem of this paper states optimal convergence rates of algorithm ACRFEM. Let c_{eff}, C_{rel}, and C_{qo} denote the constants from Theorem 3.1 and Lemma 4.3 below, and let $(\mathcal{T}_\ell)_\ell$ be the sequence of triangulations generated by ACRFEM with discrete velocities $(u_\ell)_\ell$ and pressures $(p_\ell)_\ell$ from (2.3).

Theorem 2.1 (Optimal convergence) *Let (u, p) be the exact solution of (2.2) with right-hand side f. If $(u, p, f) \in \mathcal{A}_s$ then, for any bulk parameter $0 < \theta < \theta_0 := \min\{1, c_{\mathrm{eff}}/(C_{\mathrm{drel}} + C_{\mathrm{qo}} + 1)\}$, algorithm ACRFEM generates sequences of triangulations $(\mathcal{T}_\ell)_\ell$ and discrete solutions $(u_\ell, p_\ell)_\ell$ of optimal rate of convergence in the sense that*

$$|\mathcal{T}_\ell| - |\mathcal{T}_0| \lesssim \left(|||u - u_\ell|||_{NC(\ell)}^2 + \|p - p_\ell\|_{L^2(\Omega)}^2 + osc^2(f, \mathcal{T}_\ell) \right)^{-1/(2s)} .$$

The proof of Theorem 2.1 follows in Sect. 5 based on the preparations in Sects. 3 and 4.

3 A posteriori error analysis

This section recalls some robust a posteriori error analysis of the Stokes problem. The following theorem states efficiency, reliability, and discrete reliability for the estimator η_ℓ from Algorithm ACRFEM.

Theorem 3.1 (Efficiency, reliability, discrete reliability) *Let (u, p) be the exact solution of (2.2) with right-hand side $f \in L^2(\Omega; \mathbb{R}^2)$, and let (u_ℓ, p_ℓ) be the discrete solution of (2.3). There exist positive constants c_{eff}, C_{rel}, C_{drel} depending on $\mathcal{T}_0$ but independent of the mesh-size h_ℓ such that*

$$c_{\mathrm{eff}}\eta_\ell^2 \le \|Du - D_\ell u_\ell\|_{L^2(\Omega)}^2 + \|p - p_\ell\|_{L^2(\Omega)}^2 + osc_\ell^2 \le C_{\mathrm{rel}}\eta_\ell^2.$$

Furthermore, discrete reliability holds in the sense that

$$|\!|\!| u_{\ell+k} - u_\ell |\!|\!|_{NC(\ell+k)} + \|p_{\ell+k} - p_\ell\|_{L^2(\Omega)} \le C_{\mathrm{drel}}^{1/2}\eta_\ell(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+k}).$$

The proofs of efficiency and reliability in Theorem 3.1 are given in [17]. The proof of discrete reliability follows from an orthogonal decomposition and a discrete Poincaré inequality.

The discrete Helmholtz decomposition requires the following notation. Let $\mathbb{R}^{2\times 2}_{dev}$ denote the trace-free 2×2 matrices and Z_{CR} the discrete divergence free Crouzeix–Raviart functions (with homogeneous Dirichlet boundary condition enforced pointwise in the midpoints of boundary edges) with respect to some regular triangulation $\mathcal{T}$. Define

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

with $\operatorname{curl}\beta := \dfrac{\partial\beta_2}{\partial x_1} - \dfrac{\partial\beta_1}{\partial x_2}$ and $\operatorname{Curl}\beta := \begin{pmatrix} -\frac{\partial\beta_1}{\partial x_2} & \frac{\partial\beta_1}{\partial x_1} \\ -\frac{\partial\beta_2}{\partial x_2} & \frac{\partial\beta_2}{\partial x_1} \end{pmatrix}$ for a vector field $\beta = (\beta_1, \beta_2) \in X$.

Theorem 3.2 (Discrete Helmholtz decomposition of piecewise constant deviatoric matrices) *The decomposition*

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

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

The proof of Theorem 3.2 requires the tr-dev-div Lemma.

Lemma 3.3 (tr-dev-div Lemma) *Any $\tau \in L^2(\Omega; \mathbb{R}^{2\times 2})$ with $\int_\Omega tr(\tau)\,\mathrm{d}x = 0$ satisfies*

$$\|\tau\|_{L^2(\Omega)}^2 \lesssim \|dev\tau\|_{L^2(\Omega)}^2 + \|div\tau\|_{H^{-1}(\Omega)}^2 .$$

Proof Proposition 3.1 in Sect. IV.3 of [3] contains this result for a symmetric τ, but the proof applies verbatim to the situation of this lemma. $\square$

Proof of Theorem 3.2 Since, for any $z_{CR} \in Z_{\mathrm{CR}}$ and any $\beta_C \in X$,

$$\int_\Omega \mathrm{D}_{NC} z_{CR} : \operatorname{dev} \operatorname{Curl} \beta_C \, \mathrm{d}x = \int_\Omega \mathrm{D}_{NC} z_{CR} : \operatorname{Curl} \beta_C \, \mathrm{d}x = 0,$$

the decomposition is orthogonal. Moreover, the inclusion

$$\mathrm{D}_{NC} Z_{\mathrm{CR}} \oplus \operatorname{dev} \operatorname{Curl} X \subset P_0(\mathcal{T}; \mathbb{R}^{2\times 2}_{dev})$$

is obvious. Hence, it remains to prove that the dimensions of the two spaces coincide. Since $\dim \left(P_0(\mathcal{T}; \mathbb{R}^{2\times 2}_{dev})\right) = 3|\mathcal{T}|$, we need to show that $\dim(\operatorname{dev} \operatorname{Curl} X \oplus \dim Z_{\mathrm{CR}}) = 3|\mathcal{T}|$.

The operator $\operatorname{dev} \operatorname{Curl} : X \to P_0(\mathcal{T}_\ell; \mathbb{R}^{2\times 2}_{dev})$ is linear and injective. To prove injectivity, let $v_C \in X$ with $\operatorname{dev} \operatorname{Curl} v_C = 0$. Since $\int_\Omega \operatorname{tr}(\operatorname{Curl} v_C)\, \mathrm{d}x = \int_\Omega \operatorname{curl} v_C \, \mathrm{d}x = 0$ by definition of X, and since $\operatorname{div}\operatorname{Curl} v_C = 0$, the trace-dev-div Lemma 3.3 implies that $\operatorname{Curl} v_C = 0$. Since the integral mean of v_C is zero, one concludes $v_C = 0$.

The injectivity of dev Curl implies

$$\dim(\operatorname{dev} \operatorname{Curl} X) = \dim X = 2|\mathcal{N}| - 3.$$

Since Ω is simply connected, Z_{CR} is spanned by the $|\mathcal{N}(\Omega)| + |\mathcal{E}(\Omega)|$ basis functions given in [8, Chapter III, §7]. Euler's formula proves

$$\dim(\operatorname{dev} \operatorname{Curl} X \oplus \dim Z_{\mathrm{CR}}) = \dim(\operatorname{dev} \operatorname{Curl} X) + \dim(Z_{\mathrm{CR}}) = 3|\mathcal{T}|.$$

$\square$

Lemma 3.4 (Discrete Poincaré inequality) *Let* $\alpha_{\ell+k} \in \mathrm{CR}^1_0(\mathcal{T}_{\ell+k})$ *and* $\alpha_\ell \in \mathrm{CR}^1_0(\mathcal{T}_\ell)$ *with equal integral means* $\fint_E \alpha_{\ell+k}\, \mathrm{d}s = \fint_E \alpha_\ell \, \mathrm{d}s$ *along any edge* $E \in \mathcal{E}_\ell$. *Then, for any* $T \in \mathcal{T}_\ell$, *the following discrete Poincaré inequality holds*

$$\|\alpha_{\ell+k} - \alpha_\ell\|_{L^2(T)} \lesssim |T|^{1/2} \, \|D_{\ell+k} \alpha_{\ell+k}\|_{L^2(T)} \, .$$

Proof The proof can be found in [25, Lemma 4.1] and is based on a result of [9]. $\square$

Proof of discrete reliability in Theorem 3.1 We will solely prove that

$$|||u_{\ell+k} - u_\ell|||_{NC(\ell+k)} \le C_{\mathrm{drel}}^{1/2} \eta_\ell(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+k}).$$

Since the upper bound of the pressure difference $\|p_{\ell+k} - p_\ell\|_{L^2(\Omega)}$ is not needed in the remaining analysis of this paper, its proof is omitted; it can be found in [21, Lemma 8.1].

The discrete Helmholtz decomposition from Theorem 3.2 leads to $\alpha^{CR}_{\ell+k} \in Z_{\ell+k}$ and $\beta^{C}_{\ell+k} \in C(\Omega;\mathbb{R}^2) \cap P_1(\mathcal{T}_{\ell+k};\mathbb{R}^2)$ with

$$\int_\Omega \beta^C_{\ell+k}\,\mathrm{d}x = 0, \quad \int_\Omega \operatorname{curl}\beta^C_{\ell+k}\,\mathrm{d}x = 0, \quad \text{and}$$
$$\mathrm{D}_{\ell+k}u_{\ell+k} - \mathrm{D}_\ell u_\ell = \mathrm{D}_{\ell+k}\alpha^{CR}_{\ell+k} + \operatorname{dev}\operatorname{Curl}\beta^C_{\ell+k}.$$

This implies

$$|||u_{\ell+k} - u_\ell|||^2_{NC(\ell+k)} = \left|\left|\left|\alpha^{CR}_{\ell+k}\right|\right|\right|^2_{NC(\ell+k)} + \left\|\operatorname{dev}\operatorname{Curl}\beta^C_{\ell+k}\right\|^2_{L^2(\Omega)}. \tag{3.1}$$

The nonconforming interpolation $\alpha^{CR}_\ell =: \mathcal{I}^{NC}_\ell \alpha^{CR}_{\ell+k} \in V_\ell$ is defined uniquely by

$$\fint_E \alpha^{CR}_\ell\,\mathrm{d}s = \fint_E \alpha^{CR}_{\ell+k}\,\mathrm{d}s \quad \text{for all } E \in \mathcal{E}_\ell. \tag{3.2}$$

In fact, since $\alpha^{CR}_{\ell+k} \in Z_{\ell+k}$, we have $\alpha^{CR}_\ell \in Z_\ell$. The identity (3.2) holds on either side of each $E \in \mathcal{E}_\ell$ and so

$$\int_E [(\alpha^{CR}_{\ell+k} - \alpha^{CR}_\ell)\mathrm{D}_\ell u_\ell]_E \cdot \nu_E\,\mathrm{d}s = 0 \qquad \text{for all } E \in \mathcal{E}_\ell.$$

Moreover, $\alpha^{CR}_{\ell+k} = \alpha^{CR}_\ell$ on $T \in \mathcal{T}_\ell \cap \mathcal{T}_{\ell+k}$. This leads to

$$\begin{aligned}
\left|\left|\left|\alpha^{CR}_{\ell+k}\right|\right|\right|^2_{NC(\ell+k)} &= \int_\Omega (\mathrm{D}_{\ell+k}u_{\ell+k} - \mathrm{D}_\ell u_\ell) : \mathrm{D}_{\ell+k}\alpha^{CR}_{\ell+k}\,\mathrm{d}x \\
&= \int_\Omega f \cdot \alpha^{CR}_{\ell+k}\,\mathrm{d}x - \sum_{E\in\mathcal{E}_\ell}\int_E [\alpha^{CR}_{\ell+k}\mathrm{D}_\ell u_\ell]_E \cdot \nu_E\,\mathrm{d}s \\
&= \sum_{T\in\mathcal{T}_\ell\setminus\mathcal{T}_{\ell+k}} \int_T f \cdot \left(\alpha^{CR}_{\ell+k} - \alpha^{CR}_\ell\right)\mathrm{d}x \\
&\quad - \sum_{E\in\mathcal{E}_\ell}\int_E \left[\left(\alpha^{CR}_{\ell+k} - \alpha^{CR}_\ell\right)\mathrm{D}_\ell u_\ell\right]_E \cdot \nu_E\,\mathrm{d}s \\
&\le \sum_{T\in\mathcal{T}_\ell\setminus\mathcal{T}_{\ell+k}} \left(\|f\|_{L^2(T)}\left\|\alpha^{CR}_{\ell+k} - \alpha^{CR}_\ell\right\|_{L^2(T)}\right).
\end{aligned}$$

The combination of the aforementioned estimates and the discrete Poincaré inequality of Lemma 3.4 results in

$$\left|\left|\left|\alpha^{CR}_{\ell+k}\right|\right|\right|_{NC(\ell+k)} \lesssim \|h_\ell f\|_{L^2(\mathcal{T}_\ell\setminus\mathcal{T}_{\ell+k})}. \tag{3.3}$$

The analysis of the second term on the right hand side of (3.1) requires the Scott-Zhang [28] interpolation $\beta_\ell^C := \mathcal{I}_\ell \beta_{\ell+k}^C$ on $\mathcal{T}_\ell$. For its definition, one chooses an edge $E \in \mathcal{E}_\ell(z) \cap (\mathcal{E}_\ell \setminus \mathcal{E}_{\ell+k})$ for any $z \in \mathcal{N}_\ell$ whenever possible. If $\mathcal{E}_\ell(z) \cap (\mathcal{E}_\ell \setminus \mathcal{E}_{\ell+k}) = \emptyset$, this choice is arbitrary. Then β_ℓ^C satisfies

$$\left\| \beta_{\ell+k}^C - \beta_\ell^C \right\|_{L^2(E)} = 0 \qquad \text{for all } \; E \in \mathcal{E}_{\ell+k} \cap \mathcal{E}_\ell .$$

A standard trace inequality on ω_E in $\mathcal{T}_{\ell+k}$ verifies

$$\left\| \beta_{\ell+k}^C - \beta_\ell^C \right\|_{L^2(E)} \lesssim |E|^{1/2} \left\| \beta_{\ell+k}^C \right\|_{H^1(\omega_E)} \qquad \text{for all } \; E \in \mathcal{E}_{\ell+k} \setminus \mathcal{E}_\ell .$$

Since $\int_\Omega (\mathrm{D}_{\ell+k} u_{\ell+k} - \mathrm{D}_\ell u_\ell) : \mathrm{Curl} \beta_\ell^C \, \mathrm{d}x = 0$, this leads to

$$\begin{aligned}
\left\| \mathrm{dev\ Curl} \beta_{\ell+k}^C \right\|_{L^2(\Omega)}^2 &= \int_\Omega (\mathrm{D}_{\ell+k} u_{\ell+k} - \mathrm{D}_\ell u_\ell) : \mathrm{dev\ Curl} \beta_{\ell+k}^C \, \mathrm{d}x \\
&= \int_\Omega (\mathrm{D}_{\ell+k} u_{\ell+k} - \mathrm{D}_\ell u_\ell) : \mathrm{Curl} \left(\beta_{\ell+k}^C - \beta_\ell^C \right) \mathrm{d}x \\
&= - \sum_{T \in \mathcal{T}_{\ell+k}} \int_T \mathrm{curl}\, (\mathrm{D}_{\ell+k} u_{\ell+k} - \mathrm{D}_\ell u_\ell) \cdot \left(\beta_{\ell+k}^C - \beta_\ell^C \right) \mathrm{d}x \\
&\quad + \sum_{E \in \mathcal{E}_{\ell+k}} \int_E [\partial u_\ell / \partial s]_E \cdot \left(\beta_{\ell+k}^C - \beta_\ell^C \right) \mathrm{d}s \\
&\leq \sum_{E \in \mathcal{E}_{\ell+k} \setminus \mathcal{E}_\ell} \| [\partial u_\ell / \partial s]_E \|_{L^2(E)} \left\| \beta_{\ell+k}^C - \beta_\ell^C \right\|_{L^2(E)} \\
&\lesssim \sum_{E \in \mathcal{E}_{\ell+k} \setminus \mathcal{E}_\ell} |E|^{1/2} \| [\partial u_\ell / \partial s]_E \|_{L^2(E)} \left\| \beta_{\ell+k}^C \right\|_{H^1(\omega_E)} \\
&\lesssim \eta_\ell (\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+k}) \left\| \mathrm{D} \beta_{\ell+k}^C \right\|_{L^2(\Omega)} .
\end{aligned}$$

Since $\left\| \mathrm{D} \beta_{\ell+k}^C \right\|_{L^2(\omega_E)} \lesssim \left\| \mathrm{dev\ Curl} \beta_{\ell+k}^C \right\|_{L^2(\Omega)}$ this proves

$$\left\| \mathrm{dev\ Curl} \beta_{\ell+k}^C \right\|_{L^2(\Omega)} \lesssim \eta_\ell (\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+k}). \tag{3.4}$$

The combination of (3.1) and (3.3) - (3.4) concludes the proof. □

4 Contraction property

The proof of optimality involves the contraction property for some linear combination ξ_ℓ of the estimated error η_ℓ^2, the volume term $\|h_\ell f\|_{L^2(\Omega)}$, and the error in the broken energy norm $|||u - u_\ell|||_{NC(\ell)}$. A similar linear combination including the pressure error was used earlier in [21, Theorem 4.4].

Theorem 4.1 (Contraction property) *Given some bulk parameter* $0 < \theta < 1$ *in (2.4), and any* C_{rel}, Λ, ρ *and* C_{qo} *from Theorem 3.1, and Lemmas 4.2–4.3, there exist positive* α, β, *and* $0 < \varrho < 1$ *such that in* ACRFEM *on each level* $\ell \in \mathbb{N}_0$,

$$\xi_\ell^2 := \eta_\ell^2 + \alpha \, \|h_\ell f\|^2_{L^2(\Omega)} + \beta \, |||u - u_\ell|||^2_{NC(\ell)}$$

satisfies

$$\xi^2_{\ell+1} \le \varrho \, \xi^2_\ell .$$

The proof of the contraction property is based on the subsequent lemmas on the estimator reduction and quasi-orthogonality.

Lemma 4.2 (Estimator reduction) *For any* $0 < \delta < \theta/(\sqrt{2} - \theta)$ *with bulk parameter* $0 < \theta < 1$ *in (2.4) there exists some* $\Lambda > 0$ *such that* η_ℓ *reduces on each level* $\ell \in \mathbb{N}_0$ *of* ACRFEM *with* $\rho := (1 + \delta)(1 - \theta/\sqrt{2}) < 1$ *in the sense of*

$$\eta^2_{\ell+1} \le \rho \eta^2_\ell + \Lambda \, |||u_{\ell+1} - u_\ell|||^2_{NC(\ell+1)} . \tag{4.1}$$

Proof The proof is verbatim the same as that of Lemma 4.2 in [16] and, hence, not repeated here. □

The following lemma states quasi-orthogonality for the Stokes problem and Crouzeix–Raviart FEM as in [16] for the pure displacement problem in elasticity. The result will be essential for the proof of quasi-optimality below. Quasi-orthogonality for the Poisson problem has been introduced in [10,11] and sharpened for mixed FEM [4]. The sharpened form has been employed for nonconforming methods in [20] and later in [7,24,25] for the Poisson problem and in [16] for linear elasticity.

For convenient reading the volume term on the elements of a subset $\mathcal{F}_\ell \subseteq \mathcal{T}_\ell$ is abbreviated by $\|h_\ell f\|^2_{\mathcal{F}_\ell} := \sum_{T \in \mathcal{F}_\ell \cap \mathcal{T}_\ell} |T| \, \|f\|^2_{L^2(T)}$.

Lemma 4.3 (Quasi-orthogonality) *There exists some positive constant* C_{qo}, *which depends on* $\mathcal{T}_0$ *only, such that for admissible refinements* $\mathcal{T}_\ell$ *of* $\mathcal{T}_0$ *and* $\mathcal{T}_{\ell+k}$ *of* $\mathcal{T}_\ell$ *and the respective discrete velocities* u_ℓ *and* $u_{\ell+k}$ *fulfil quasi-orthogonality in the sense of*

$$\left| a_{NC(\ell+k)} \left(u - u_{\ell+k}, u_{\ell+k} - u_\ell \right) \right| \le C_{\mathrm{qo}}^{1/2} \, |||u - u_{\ell+k}|||_{NC(\ell+k)} \, \|h_\ell f\|_{\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+k}} .$$

Proof Let $\mathcal{I}_\ell^{NC}$ be the nonconforming interpolation operator as in (3.2) and $\mathcal{I}_{\ell+k}^{NC}$ denote the standard nonconforming interpolation operator. Since $⨍_E \mathcal{I}^{NC}_{\ell+k} u \, \mathrm{d}s = ⨍_E u \, \mathrm{d}s$ for all $E \in \mathcal{E}_{\ell+k}$ an integration by parts argument shows

$$\mathrm{D}_{\ell+k} \mathcal{I}^{NC}_{\ell+k} u \Big|_T = ⨍_T \mathrm{D}u \, \mathrm{d}x \quad \text{for } T \in \mathcal{T}_{\ell+k} .$$

The interpolation operators lead to

$$\begin{aligned} a_{NC(\ell+k)}\left(u-u_{\ell+k}, u_{\ell+k}-u_\ell\right) &= \int_\Omega \left(\mathrm{D}u - \mathrm{D}_{\ell+k}u_{\ell+k}\right) : \mathrm{D}_{\ell+k}\left(u_{\ell+k}-u_\ell\right)\mathrm{d}x \\ &= \int_\Omega \mathrm{D}_{\ell+k}\left(\mathcal{I}^{NC}_{\ell+k}u - u_{\ell+k}\right) : \mathrm{D}_{\ell+k}u_{\ell+k}\,\mathrm{d}x \\ &\quad - \int_\Omega \mathrm{D}_\ell\left(\mathcal{I}^{NC}_\ell\left(u-u_{\ell+k}\right)\right) : \mathrm{D}_\ell u_\ell\,\mathrm{d}x\,. \end{aligned}$$

The integral mean property of $\mathcal{I}^{NC}_\ell$ proves

$$\mathcal{I}^{NC}_\ell u = \mathcal{I}^{NC}_\ell \mathcal{I}^{NC}_{\ell+k} u \ \text{ and } \ \mathrm{div}_\ell \mathcal{I}^{NC}_\ell u = \mathrm{div}_{\ell+k}\mathcal{I}^{NC}_{\ell+k}u = 0.$$

Hence, with $v_{\ell+k} := \mathcal{I}^{NC}_{\ell+k}u - u_{\ell+k} \in V_{\ell+k}$,

$$\begin{aligned} &a_{NC(\ell+k)}\left(u-u_{\ell+k}, u_{\ell+k}-u_\ell\right) \\ &\quad = a_{NC(\ell+k)}\left(u_{\ell+k}, \mathcal{I}^{NC}_{\ell+k}u - u_{\ell+k}\right) - a_{NC(\ell)}\left(u_\ell, \mathcal{I}^{NC}_\ell\left(\mathcal{I}^{NC}_{\ell+k}u - u_{\ell+k}\right)\right) \\ &\quad = F\left(v_{\ell+k} - \mathcal{I}^{NC}_{\ell+k}v_{\ell+k}\right). \end{aligned}$$

Since $\mathcal{I}^{NC}_{\ell+k}$ sustains the integral mean on any $E \in \mathcal{E}_\ell$, Lemma 3.4 proves

$$\left\| v_{\ell+k} - \mathcal{I}^{NC}_{\ell+k}v_{\ell+k}\right\|_{L^2(T)} \lesssim |T|^{1/2}\,\|\mathrm{D}_{\ell+k}v_{\ell+k}\|_{L^2(T)} \quad \text{for all } T \in \mathcal{T}_\ell.$$

This concludes the proof,

$$\begin{aligned} &\left|a_{NC(\ell+k)}\left(u-u_{\ell+k}, u_{\ell+k}-u_\ell\right)\right| \\ &\quad \lesssim \sum_{T\in\mathcal{T}_\ell\setminus\mathcal{T}_{\ell+k}} \|f\|_{L^2(T)}\, h_T \left\|\mathrm{D}_{\ell+k}\left(\mathcal{I}^{NC}_{\ell+k}u - u_{\ell+k}\right)\right\|_{L^2(T)} \\ &\quad \lesssim \|h_\ell f\|_{\mathcal{T}_\ell\setminus\mathcal{T}_{\ell+k}}\, ⦀u-u_{\ell+k}⦀_{NC(\ell+k)}\,. \end{aligned}$$

□

Proof of Theorem 4.1 The following proof shows convergence of AcrFEM and η_ℓ. Given $0 < \delta < \theta(\sqrt{2}-\theta)$ and $\rho = (1+\delta)(1-\theta/\sqrt{2}) < 1$ from the estimator reduction of Lemma 4.2, one can choose positive γ_1, γ_2 with

$$0 < \gamma_2 < \min\left\{\Lambda, \frac{1-\rho}{C_{\mathrm{rel}}}\right\} \ \text{ and } \ 1 < \frac{\Lambda}{\gamma_2} < \gamma_1.$$

With C_{qo} from Lemma 4.3, set

$$2\Lambda\gamma_1 C_{\mathrm{qo}} < \alpha \text{ and } \quad \beta := \Lambda(1-\gamma_1^{-1}).$$

Here and throughout let $\varepsilon_\ell^2 := |||u - u_\ell|||_{NC(\ell)}^2$ denote the discrete energy error with respect to $\mathcal{T}_\ell$. The estimator reduction of Lemma 4.2, the quasi-orthogonality of Lemma 4.3, and Young's inequality show

$$\begin{aligned}
\eta_{\ell+1}^2 &\le \rho\eta_\ell^2 + \Lambda\, |||u_{\ell+1} - u_\ell|||_{NC(\ell+1)}^2 \\
&\le \rho\eta_\ell^2 + \Lambda a_{NC(\ell+1)}\,(u_{\ell+1} - u_\ell, u_{\ell+1} - u_\ell) \\
&\le \rho\eta_\ell^2 + \Lambda a_{NC(\ell+1)}\,((u_\ell - u) - (u_{\ell+1} - u), (u_\ell - u) + (u_{\ell+1} - u)) \\
&\quad +2\Lambda a_{NC(\ell+1)}\,(u_\ell - u_{\ell+1}, u - u_{\ell+1}) \\
&\le \rho\eta_\ell^2 + \Lambda\left(\varepsilon_\ell^2 - \varepsilon_{\ell+1}^2 + \gamma_1 C_{\mathrm{qo}}\,\|h_\ell f\|_{\mathcal{T}_\ell\setminus\mathcal{T}_{\ell+1}}^2 + \frac{1}{\gamma_1}\varepsilon_{\ell+1}^2\right).
\end{aligned}$$

Hence,

$$\eta_{\ell+1}^2 + \beta\varepsilon_{\ell+1}^2 \le \rho\eta_\ell^2 + \Lambda\varepsilon_\ell^2 + \Lambda\gamma_1 C_{\mathrm{qo}}\,\|h_\ell f\|_{\mathcal{T}_\ell\setminus\mathcal{T}_{\ell+1}}^2 .$$

Let

$$\eta_\ell^2(\mathcal{E}_\ell) := \sum_{T\in\mathcal{T}_\ell} |T|^{1/2} \sum_{E\in\mathcal{E}(T)} \|[\partial u_\ell/\partial s]_E\|_{L^2(E)}^2$$

and recall $\xi_\ell^2 := \eta_\ell^2 + \alpha\,\|h_\ell f\|_{L^2(\Omega)}^2 + \beta\varepsilon_\ell^2$. Since the volume term satisfies

$$\|h_{\ell+1} f\|_{L^2(\Omega)}^2 \le \|h_\ell f\|_{L^2(\Omega)}^2 - \frac{1}{2}\,\|h_\ell f\|_{\mathcal{T}_\ell\setminus\mathcal{T}_{\ell+1}}, \tag{4.2}$$

reliability $\gamma_2\varepsilon_\ell^2 \le \gamma_2 C_{\mathrm{rel}}\eta_\ell^2$ proves

$$\begin{aligned}
\xi_{\ell+1}^2 &\le (\rho + \gamma_2 C_{\mathrm{rel}})\,\eta_\ell^2(\mathcal{E}_\ell) + (\Lambda - \gamma_2)\,\varepsilon_\ell^2 \\
&\quad + \left(\Lambda\gamma_1 C_{\mathrm{qo}} - \alpha/2\right)\|h_\ell f\|_{\mathcal{T}_\ell\setminus\mathcal{T}_{\ell+1}}^2 + (\alpha + \rho + \gamma_2 C_{\mathrm{rel}})\,\|h_\ell f\|_{L^2(\Omega)}^2 .
\end{aligned}$$

This reads $\xi_{\ell+1}^2 \le \varrho\xi_\ell^2$ with

$$\varrho := \max\left\{\rho + \gamma_2 C_{\mathrm{rel}}, \frac{\Lambda - \gamma_2}{\beta}, 1 - \frac{1 - \rho - \gamma_2 C_{\mathrm{rel}}}{\alpha + 1}\right\} < 1. \qquad \square$$

5 Proof of optimal convergence

This section is devoted to the proof of Theorem 2.1 and is based on the contraction property (Theorem 4.1), the discrete reliability (Theorem 3.1), and the quasi-orthogonality (Lemma 4.3) from the previous sections.

For $s > 0$, consider the modified approximation class

$$\tilde{\mathcal{A}}_s := \left\{(u, p, f) \in H_0^1(\Omega; \mathbb{R}^2) \times L_0^2(\Omega) \times L^2(\Omega; \mathbb{R}^2) \big| |(u, p, f)|_{\tilde{\mathcal{A}}_s} < \infty\right\}$$

with

$$|(u, p, f)|_{\tilde{\mathcal{A}}_s} := \sup_{N \in \mathbb{N}} \left(N^s \inf_{|\mathcal{T}| - |\mathcal{T}_0| \leq N} \left(\||u - u_{\mathcal{T}}\||_{NC(\mathcal{T})}^2 + \|h_{\mathcal{T}} f\|_{L^2(\Omega)^2}\right)^{1/2}\right).$$

Proposition 5.1 *Let (u, p) be the exact solution of (2.2) with right-hand side $f \in L^2(\Omega; \mathbb{R}^2)$, let $\mathcal{T}$ be some admissible triangulation that is refined from $\mathcal{T}_0$ by NVB, and let $(u_{\mathcal{T}}, p_{\mathcal{T}})$ be the corresponding discrete solution of (2.3). Then*

$$\||u - u_{\mathcal{T}}\||_{NC(\mathcal{T})}^2 + \|p - p_{\mathcal{T}}\|_{L^2(\Omega)}^2 + osc^2(f, \mathcal{T}) \approx \||u - u_{\mathcal{T}}\||_{NC(\mathcal{T})}^2 + \|h_{\mathcal{T}} f\|_{L^2(\Omega)}^2$$

holds with hidden constants that depend on $\mathcal{T}_0$ but not on the mesh-size $h_{\mathcal{T}}$.

Proof [17, Remark 3.2] shows

$$\|p - p_{\mathcal{T}}\|_{L^2(\Omega)} \lesssim \||u - u_{\mathcal{T}}\||_{NC(\mathcal{T})} + \|h_{\mathcal{T}} f\|_{L^2(\Omega)},$$

which proves one inequality. The efficiency of $\|h_{\mathcal{T}} f\|_{L^2(\Omega)}$ up to oscillations (see Theorem 3.1) proves the other inequality. □

Proposition 5.1 yields $|(u, p, f)|_{\mathcal{A}_s} \approx |(u, p, f)|_{\tilde{\mathcal{A}}_s}$. Hence, it suffices to prove quasi-optimality with regard to $\tilde{\mathcal{A}}_s$.

Given positive c_{eff}, C_{qo}, C_{drel}, α, β, and ϱ from Theorems 3.1 and 4.1 and Lemma 4.3, C_{up} arising from Proposition 5.1 in (5.4) below, and $0 < \theta < \theta_0 \leq 1$ with $\theta_0 := \min\left\{1, c_{\text{eff}}/(C_{\text{drel}} + C_{\text{qo}} + 1)\right\}$, choose some τ with

$$0 < \tau^2 < \tau_0^2 := \frac{c_{\text{eff}} - \theta(C_{\text{drel}} + C_{\text{qo}} + 1)}{c_{\text{eff}} + C_{\text{up}}(1 - \theta)} < 1$$

and set

$$\epsilon^2 := \frac{\tau^2}{\max\{4, 16C_{\text{qo}}\}} \left(\||u - u_\ell\||_{NC(\ell)}^2 + \eta_\ell^2\right) \approx \xi_\ell^2. \tag{5.1}$$

Given this ϵ, the definition of $\tilde{\mathcal{A}}_s$ above implies the existence of an admissible regular triangulation $\mathcal{T}_\epsilon$ refined from $\mathcal{T}_0$ with

$$\||u - u_\epsilon\||_{NC(\epsilon)}^2 + \|h_\epsilon f\|_{L^2(\Omega)}^2 \leq \epsilon^2 \text{ and } |\mathcal{T}_\epsilon| - |\mathcal{T}_0| \lesssim \epsilon^{-1/s}. \tag{5.2}$$

In the last estimate, the factor $|(u, p, f)|_{\tilde{\mathcal{A}}_s}$ is hidden in the generic constant behind the symbol $\lesssim$. This factor is neglected in the sequel for brevity but enters at the end as displayed in (1.1).

The number of elements of the overlay $\mathcal{T}_{\ell+\epsilon} := \mathcal{T}_\epsilon \oplus \mathcal{T}_\ell$ (the coarsest triangulation which refined $\mathcal{T}_\epsilon$ as well as $\mathcal{T}_\ell$) satisfies [15,12]

$$|\mathcal{T}_{\ell+\epsilon}| - |\mathcal{T}_\ell| = |\mathcal{T}_\epsilon \oplus \mathcal{T}_\ell| - |\mathcal{T}_\ell| \leq |\mathcal{T}_\epsilon| - |\mathcal{T}_0|.$$

Let $\mathcal{F} := \mathcal{T}_\ell \setminus \mathcal{T}_{\ell+\epsilon}$ denote the set of triangles in $\mathcal{T}_\ell$ refined in $\mathcal{T}_{\ell+\epsilon}$. The choice of $\mathcal{T}_\epsilon$ with (5.1)–(5.2) implies

$$|\mathcal{F}| \leq |\mathcal{T}_{\ell+\epsilon}| - |\mathcal{T}_\ell| \leq |\mathcal{T}_\epsilon| - |\mathcal{T}_0| \lesssim \epsilon^{-1/s} \approx \xi_\ell^{-1/s}. \tag{5.3}$$

Let $\varepsilon_\ell^2 := \|u - u_\ell\|_{NC(\ell)}^2$, $\varepsilon_\epsilon^2 := \|u - u_\epsilon\|_{NC(\epsilon)}^2$, and $\varepsilon_{\ell+\epsilon}^2 := \|u - u_{\ell+\epsilon}\|_{NC(\ell+\epsilon)}^2$ denote the discrete energy error with respect to $\mathcal{T}_\ell$, $\mathcal{T}_\epsilon$, and $\mathcal{T}_{\ell+\epsilon}$. Quasi-orthogonality from Lemma 4.3, the discrete reliability of Theorem 3.1 and Young's inequality show

$$\begin{aligned}
\varepsilon_\ell^2 &\leq \|u_{\ell+\epsilon} - u_\ell\|_{NC(\ell+\epsilon)}^2 + \varepsilon_{\ell+\epsilon}^2 + 2C_{\mathrm{qo}}^{1/2} \varepsilon_{\ell+\epsilon} \|h_\ell f\|_{\mathcal{F}}^2 \\
&\leq C_{\mathrm{drel}} \eta_\ell^2(\mathcal{F}) + 2\varepsilon_{\ell+\epsilon}^2 + C_{\mathrm{qo}} \|h_\ell f\|_{\mathcal{F}}^2 \\
&\leq \left(C_{\mathrm{drel}} + C_{\mathrm{qo}}\right) \eta_\ell^2(\mathcal{F}) + 2\varepsilon_{\ell+\epsilon}^2.
\end{aligned}$$

Let $\mathcal{F}_\epsilon := \mathcal{T}_\epsilon \setminus \mathcal{T}_{\ell+\epsilon}$. Quasi-orthogonality from Lemma 4.3 and Young's inequality show

$$\varepsilon_{\ell+\epsilon}^2 \leq 2\varepsilon_\epsilon^2 + 4C_{\mathrm{qo}} \|h_\epsilon f\|_{\mathcal{F}_\epsilon}^2.$$

Since the volume term satisfies $\|h_{\ell+\epsilon} f\|_{L^2(\Omega)}^2 \leq \|h_\epsilon f\|_{L^2(\Omega)}^2 - \frac{1}{2} \|h_\epsilon f\|_{\mathcal{F}_\epsilon}^2$,

$$\begin{aligned}
\varepsilon_{\ell+\epsilon}^2 &\leq \max\{2, 8C_{\mathrm{qo}}\}\left(\varepsilon_\epsilon^2 + \|h_\epsilon f\|_{L^2(\Omega)}^2 - \|h_{\ell+\epsilon} f\|_{L^2(\Omega)}^2\right) \\
&\leq \frac{\tau^2}{2}(\varepsilon_\ell^2 + \eta_\ell^2) - \max\{2, 8C_{\mathrm{qo}}\} \|h_{\ell+\epsilon} f\|_{L^2(\Omega)}^2.
\end{aligned}$$

The combination of this upper bound for $\varepsilon_{\ell+\epsilon}^2$ with the aforementioned estimate leads to

$$(1-\tau^2)\varepsilon_\ell^2 \leq \left(C_{\mathrm{drel}} + C_{\mathrm{qo}}\right) \eta_\ell^2(\mathcal{F}) + \tau^2 \eta_\ell^2 - \max\{4, 16C_{\mathrm{qo}}\} \|h_{\ell+\epsilon} f\|_{L^2(\Omega)}^2.$$

Due to Proposition 5.1 there exists some positive generic constant C_{up} such that

$$\begin{aligned}
&\|\mathrm{D}u - \mathrm{D}_\ell u_\ell\|_{L^2(\Omega)}^2 + \|p - p_\ell\|_{L^2(\Omega)}^2 + \mathrm{osc}_\ell^2 \\
&\quad \leq C_{\mathrm{up}} \left(\|\mathrm{D}u - \mathrm{D}_\ell u_\ell\|_{L^2(\Omega)}^2 + \|h_\ell f\|_{L^2(\Omega)}^2\right).
\end{aligned} \tag{5.4}$$

This inequality plus efficiency of η_ℓ^2 of Theorem 3.1 leads to

$$
\begin{aligned}
c_{\mathrm{eff}}(1-\tau^2)\eta_\ell^2 \le C_{\mathrm{up}}\left(C_{\mathrm{drel}}+C_{\mathrm{qo}}\right)\eta_\ell^2(\mathcal{F}) &+ C_{\mathrm{up}}\tau^2\eta_\ell^2 \\
+C_{\mathrm{up}}\left(1-\tau^2\right)&\|h_\ell f\|^2_{L^2(\Omega)} \\
-C_{\mathrm{up}}\max\{4,16C_{\mathrm{qo}}\}&\|h_{\ell+\epsilon} f\|^2_{L^2(\Omega)}.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\left(c_{\mathrm{eff}}(1-\tau^2)-C_{\mathrm{up}}\tau^2\right)\eta_\ell^2 \le{}& C_{\mathrm{up}}\left(C_{\mathrm{drel}}+C_{\mathrm{qo}}\right)\eta_\ell^2(\mathcal{F}) \\
&+C_{\mathrm{up}}\left(1-\tau^2\right)\|h_\ell f\|^2_{\mathcal{F}} \\
&+C_{\mathrm{up}}(1-\tau^2-\max\{4,16C_{\mathrm{qo}}\})\|h_\ell f\|^2_{\mathcal{T}_\ell\cap\mathcal{T}_{\ell+1}}.
\end{aligned}
$$

Since the factor in front of $\|h_\ell f\|^2_{\mathcal{T}_\ell\cap\mathcal{T}_{\ell+1}}$ is negative, this verifies

$$
(c_{\mathrm{eff}}(1-\tau^2)-C_{\mathrm{up}}\tau^2)\eta_\ell^2 \le C_{\mathrm{up}}(C_{\mathrm{drel}}+C_{\mathrm{qo}}+1-\tau^2)\eta_\ell^2(\mathcal{F}).
$$

The choice of $\tau^2 < \tau_0^2$ implies that $\mathcal{F}$ fulfils the bulk criterion for $\theta \le \theta_0$. Therefore, $|\mathcal{M}_\ell| \le |\mathcal{F}|$ on any level ℓ and with (5.3),

$$
|\mathcal{M}_\ell| \lesssim \xi_\ell^{-1/s}.
$$

Moreover, the overhead in the marking procedure is bounded in the sense of [2,27]

$$
|\mathcal{T}_\ell| - |\mathcal{T}_0| \lesssim \sum_{j=0}^{\ell-1} |\mathcal{M}_j|. \tag{5.5}
$$

Since $\alpha, \beta > 0$ are chosen according to contraction property of Theorem 4.1, the assertion follows by (5.3)-(5.5). Indeed,

$$
|\mathcal{T}_\ell| - |\mathcal{T}_0| \lesssim \sum_{j=0}^{\ell-1}\xi_j^{-1/s} \lesssim \xi_\ell^{-1/s}\sum_{j=1}^{\ell-1}\varrho^{-j/(2s)} \lesssim \xi_\ell^{-1/s}.
$$

References

1. Brenner, S.C., Carstensen, C.: Encyclopedia of computational mechanics, chapter 4. In: Finite Element Methods. Wiley, New York (2004)
2. Binev, P., Dahmen, W., De Vore, R.: Adaptive finite element methods with convergence rates. Numer. Math. **97**, 219–268 (2004)
3. Brezzi, F., Fortin, M.: Mixed and Hybrid Finite Element Methods. Springer, New-York (1991)
4. Becker, R., Mao, S.: An optimally convergent adaptive mixed finite element method. Numer. Math. **111**, 35–54 (2008)

5. Becker, R., Mao, S.: Adaptive nonconforming finite elements for the Stokes equations. SIAM J. Numer. Anal. **49**(3), 970–991 (2011)
6. Bänsch, E., Morin, P., Nochetto, R.H.: An adaptive Uzawa FEM for the Stokes problem: convergence without the inf-sup condition. SIAM J. Numer. Anal. **40**(4), 1207–1229 (2002)
7. Becker, R., Mao, S., Shi, Z.: A convergent nonconforming adaptive finite element method with quasi-optimal complexity. SIAM J. Numer. Anal. **47**(6), 4639–4659 (2010)
8. Braess, D.: Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics. Cambridge University Press, Cambridge (2001)
9. Brenner, S.C., Scott, L.R.: The Mathematical theory of finite element methods, vol. 15, 3rd edn. In: Texts in Applied Mathematics. Springer, New York (2008)
10. Carstensen, C., Hoppe, R.H.W.: Convergence analysis of an adaptive nonconforming finite element method. Numer. Math. **103**(2), 251–266 (2006)
11. Carstensen, C., Hoppe, R.H.W.: Error reduction and convergence for an adaptive mixed finite element method. Math. Comp. **75**(255):1033–1042 (electronic) (2006)
12. Cascon, J.M., Kreuzer, Ch., Nochetto, R.H., Siebert, K.G.: Quasi-optimal convergence rate for an adaptive finite element method. SIAM J. Numer. Anal. **46**(5), 2524–2550 (2008)
13. Carstensen, C., Peterseim, D., Schedensack, M: Comparison results of three first-order finite element methods for the Poisson model problem. Preprint 831, DFG Research Center Matheon Berlin (2011)
14. Crouzeix, M., Raviart, P.A.: Conforming and nonconforming finite element methods for solving the stationary Stokes equations. I. Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge **7**(R-3), 33–75 (1973)
15. Carstensen, C., Rabus, H.: An optimal adaptive mixed finite element method. Math. Comp. **80**(274), 649–667 (2011)
16. Carstensen, C., Rabus, H.: The adaptive nonconforming fem for the pure displacement problem in linear elasticity is optimal and robust. SIAM J. Numer. Anal., accepted (2012)
17. Dari, E., Durán, R., Padra, C.: Error estimators for nonconforming finite element approximations of the Stokes problem. Math. Comp. **64**(211), 1017–1033 (1995)
18. Dörfler, W.: A convergent adaptive algorithm for Poisson's equation. SIAM J. Numer. Anal. **33**(3), 1106–1124 (1996)
19. Gudi, T.: A new error analysis for discontinuous finite element methods for linear elliptic problems. Math. Comp. **79**(272), 2169–2189 (2010)
20. Hu, J., Shi, Z., Xu, J.: Convergence and optimality of the adaptive morley element method. Numerische Mathematik 1–22 (2012). doi:10.1007/s00211-012-0445-0
21. Hu, J., Xu, J.: 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 (2007). Available at www.math.pku.edu.cn:8000/var/preprint/7297.pdf
22. Hu, J., Xu, J.: Convergence and optimality of the adaptive nonconforming linear element method for the Stokes problem. J. Sci. Comput. 1–24 (2012). doi:10.1007/s10915-012-9625-4
23. Morin, P., Nochetto, R. H., Siebert, K. G.: Convergence of adaptive finite element methods. SIAM Rev. **44**(4):631–658 (electronic) (2003) 2002. Revised reprint of "Data oscillation and convergence of adaptive FEM". SIAM J. Numer. Anal. **38**(2):466–488 (2000) (electronic); MR1770058 (2001g:65157)
24. Mao, S., Zhao, X., Shi, Z.: Convergence of a standard adaptive nonconforming finite element method with optimal complexity. Appl. Numer. Math. **60**, 673–688 (2010)
25. Rabus, H.: A natural adaptive nonconforming FEM of quasi-optimal complexity. Comput. Methods Appl. Math. **10**(3), 315–325 (2010)
26. Stevenson, R.: Optimality of a standard adaptive finite element method. Found. Comput. Math. **7**(2), 245–269 (2007)
27. Stevenson, R.: The completion of locally refined simplicial partitions created by bisection. Math. Comput. **77**(261), 227–241 (2008)
28. Scott, L.R., Zhang, S.: Finite element interpolation of nonsmooth functions satisfying boundary conditions. Math. Comput. **54**(190), 483–493 (1990)