Computers and Mathematics with Applications 74 (2017) 1923–1939

Contents lists available at ScienceDirect

Computers and Mathematics with Applications

journal homepage: www.elsevier.com/locate/camwa

h-adaptive least-squares finite element methods for the 2D Stokes equations of any order with optimal convergence rates☆

P. Bringmann, C. Carstensen*

Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany

ARTICLE INFO

Article history:
Available online 14 March 2017

Keywords:
Least-squares finite element method
Stokes equations
Adaptive finite element method
Higher-order discretization
Separate marking
Quasi-optimal convergence

ABSTRACT

The adaptive least-squares finite element method (LS-FEM) for the Stokes equations has recently been based on alternative error estimators in Bringmann and Carstensen (2016) for the lowest-order case. Since the first-order div LS-FEM measures the flux errors in $H(\mathrm{div})$, the data resolution error measures the L^2 norm of the right-hand side f minus the piecewise polynomial approximation $\Pi_k f$ without a mesh-size factor. This enforces separate marking with an overall abstract theory (Carstensen and Rabus, 2016). This paper contributes (a) a discussion of the scaling of the LS-FEM, (b) optimal rates of the adaptive *h*-version of any order *k*, and (c) comparing numerical results in 2D with all details on inhomogeneous Dirichlet data covered by the analysis.

1. Introduction

The least-squares finite element methods (LS-FEMs) for the Stokes equations form a competitive scheme in CFD with the minimization of L^2 residuals of the divergence of the pseudostress tensor $\boldsymbol{\sigma}$ in $H(\mathrm{div})$-conforming Raviart–Thomas spaces of order $k-1$ and of the constitutive relation for the gradient $\mathbf{D}u$ of the velocity field u in H^1-conforming piecewise polynomials of degree at most k minus the deviatoric part $\mathbf{dev}\boldsymbol{\sigma}$ of the pseudostress,

$$LS(0;\boldsymbol{\sigma},u) := \|\mathrm{div}\,\boldsymbol{\sigma}\|^2_{L^2(\Omega)} + \|\mathbf{dev}\boldsymbol{\sigma} - \mathbf{D}u\|^2_{L^2(\Omega)}.$$

The lowest-order case ($k = 1$) is equivalent to the nonconforming Crouzeix–Raviart scheme [1, Thm. 2.1], but higher-order nonconforming schemes are restricted (e.g. to 2D and odd degrees). An adaptive version of a scheme with automatic recovery of the optimal rate (for corner singularities) appears mandatory for any polynomial degree. It should be remarked, that conforming finite element methods (e.g. MINI or Taylor–Hood) are empirically observed to improve the convergence rates through adaptive mesh-refinement but – at the best knowledge of the authors – are *not* known to lead to optimal convergence rates. In other words, the optimal convergence rates for the adaptive LS-FEM of any order *k* of this paper makes the overall methodology even more competitive.

More information on the history of least-squares finite element schemes may be found in [2] and on the mathematical foundation of adaptive algorithms in [3]. Quasi-optimality of an adaptive first-order system LS-FEM has been invented for the 2D Poisson model problem in [4] and exploited for the Stokes equations in [5] for the lowest-order case $k = 1$ only.

☆ The authors gratefully acknowledge support by the Deutsche Forschungsgemeinschaft in the Priority Program 1748 'Reliable simulation techniques in solid mechanics. Development of non-standard discretization methods, mechanical and mathematical analysis' under the project 'Foundation and application of generalized mixed FEM towards nonlinear problems in solid mechanics' (CA 151/22-1).

* Corresponding author.
E-mail address: cc@math.hu-berlin.de (C. Carstensen).

http://dx.doi.org/10.1016/j.camwa.2017.02.019
0898-1221/

Numerical experiments show optimal behaviour of an adaptive algorithm with least-squares formulations driven by the local contributions of the least-squares functional, e.g., in [6] for the Poisson model problem. However, this approach does not fit into the known mathematical techniques to guarantee optimal convergence rates [3,7]. The affirmative result in [6] requires the bulk parameter θ to be close to one while the known optimality [3,7] follows exclusively for θ sufficiently small. An alternative a posteriori error estimator is therefore derived in [5] within the framework of the axioms of adaptivity and separate marking [7].

Given some right-hand side $f \in L^2(\Omega;\mathbb{R}^2)$ and piecewise differentiable Dirichlet boundary data $g \in H^1(\mathcal{E}(\partial\Omega);\mathbb{R}^2) \cap C(\partial\Omega;\mathbb{R}^2)$ with $\int_{\partial\Omega} g\cdot\nu\,ds = 0$ on a bounded simply-connected Lipschitz domain $\Omega \subseteq \mathbb{R}^2$ with polygonal boundary $\partial\Omega$, the Stokes equations seek a velocity field $u \in \mathcal{A} := \{v \in H^1(\Omega;\mathbb{R}^2) : v = g \text{ on } \partial\Omega\}$ and a pressure distribution $p \in L^2_0(\Omega)$ (i.e. $p \in L^2(\Omega)$ and $\int_\Omega p\,dx = 0$) with

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

The LS-FEM considers the equivalent first-order system [8]

$$f + \operatorname{div}\boldsymbol{\sigma} = 0 \quad \text{and} \quad \mathbf{dev}\boldsymbol{\sigma} - \mathbf{D}u = 0 \quad \text{in } \Omega \tag{1}$$

with the deviatoric part $\mathbf{dev}\boldsymbol{\sigma} := \boldsymbol{\sigma} - \operatorname{tr}(\boldsymbol{\sigma})/2\, I_{2\times 2}$ and seeks a discrete minimizer of the least-squares functional

$$LS(f;\boldsymbol{\sigma},v) := \|f + \operatorname{div}\boldsymbol{\sigma}\|^2_{L^2(\Omega)} + \|\mathbf{dev}\boldsymbol{\sigma} - \mathbf{D}u\|^2_{L^2(\Omega)} \tag{2}$$

for $\boldsymbol{\sigma} \in \boldsymbol{\Sigma} := \{\boldsymbol{\tau} \in H(\operatorname{div},\Omega;\mathbb{R}^{2\times 2}) : \operatorname{tr}\boldsymbol{\tau} \in L^2_0(\Omega)\}$ and $u \in \mathcal{A}$. The equivalence of the homogeneous functional $LS(0;\boldsymbol{\sigma},v)$ to the natural norm of the underlying function space $\boldsymbol{\Sigma}\times H^1_0(\Omega;\mathbb{R}^2)$ [8, Thm. 4.2] reads, for all $(\boldsymbol{\sigma},u) \in \boldsymbol{\Sigma}\times H^1_0(\Omega;\mathbb{R}^2)$,

$$LS(0;\boldsymbol{\sigma},u) \approx \|\boldsymbol{\sigma}\|^2_{H(\operatorname{div},\Omega)} + |||u|||^2 \tag{3}$$

and leads to efficiency and reliability of the a posteriori error estimator $LS(f;\boldsymbol{\sigma}_{LS},u_{LS})$ for some discrete minimizer $(\boldsymbol{\sigma}_{LS},u_{LS})$. Section 3 contributes an explicit constant for the norm equivalence and a discussion of its dependence on the domain Ω. The analysis from [5] for optimal rates of adaptive mesh-refinement requires exact solve to apply an alternative a posteriori error estimator with the volume contributions

$$|T|\,\|\operatorname{div}(\mathbf{dev}\boldsymbol{\sigma}_{LS} - \mathbf{D}u_{LS})\|_{L^2(T)} + |T|\,\|\operatorname{curl}\mathbf{dev}(\boldsymbol{\sigma}_{LS} - \mathbf{D}u_{LS})\|_{L^2(T)}$$

for any triangle T of area $|T|$ and with the edge contributions

$$|T|^{1/2}\|[\mathbf{dev}\boldsymbol{\sigma}_{LS} - \mathbf{D}u_{LS}]_E\nu_E\|_{L^2(E)} + |T|^{1/2}\|[\mathbf{dev}(\boldsymbol{\sigma}_{LS} - \mathbf{D}u_{LS})]_E\tau_E\|_{L^2(E)}$$

with jumps $[\bullet]_E$ across an interior edge E plus additional terms on the boundary.

The main result of this paper is the quasi-optimality of the novel adaptive algorithm (with the number $|\mathcal{T}_\ell|$ of triangles in the triangulation $\mathcal{T}_\ell$)

$$\sup_{\ell\in\mathbb{N}}\big(|\mathcal{T}_\ell| - |\mathcal{T}_0| + 1\big)^s\big(LS(f;\boldsymbol{\sigma}_\ell,u_\ell) + \operatorname{osc}^2(g',\mathcal{E}_\ell(\partial\Omega))\big)^{1/2} \approx |(u,f)|_{\mathcal{A}_s} \tag{4}$$

with the non-linear approximation class

$$\mathcal{A}_s := \big\{(u,f) \in \mathcal{A}\times L^2(\Omega;\mathbb{R}^2) : |(u,f)|^2_{\mathcal{A}_s} := \sup_{N\in\mathbb{N}} N^{2s}E(u,f,N) < \infty\big\}$$

and the best possible error

$$E(u,f,N) := \min_{\mathcal{T}\in\mathbb{T}(N)}\min_{(\boldsymbol{\tau}_{LS},v_{LS})\in\boldsymbol{\Sigma}_{k-1}(\mathcal{T})\times\mathcal{A}(\mathcal{T})}\big(LS(f;\boldsymbol{\tau}_{LS},v_{LS}) + \operatorname{osc}^2(g',\mathcal{E}(\partial\Omega))\big).$$

The proofs require a discrete Helmholtz decomposition for any polynomial degree on a simply-connected domain in [9], which has also been established based on exterior calculus by [10–12].

The paper is organized as follows. Section 2 presents details on the approximation of non-homogeneous Dirichlet boundary data with two possible interpolation operators and states the employed LS-FEM for the Stokes equations. Section 3 is devoted to the analysis of the different residuals in the least-squares functional. The statement of the alternative a posteriori error estimator and the quasi-optimal adaptive algorithm in Section 4 is followed by the main result of this paper Theorem 4.3 and the proofs of the axioms of adaptivity. Section 5 outlines the proof of the Axiom A3 (discrete reliability). Numerical experiments conclude this paper in Section 6.

This paper employs standard notation of Sobolev and Lebesgue spaces $H^k(\Omega)$, $H(\operatorname{div},\Omega)$, and $L^2(\Omega)$ and the corresponding spaces of vector- or matrix-valued functions $H^k(\Omega;\mathbb{R}^2)$, $L^2(\Omega;\mathbb{R}^2)$, $H^k(\Omega;\mathbb{R}^{2\times 2})$, $H(\operatorname{div},\Omega;\mathbb{R}^{2\times 2})$, and $L^2(\Omega;\mathbb{R}^{2\times 2})$. Let $\langle\bullet,\bullet\rangle_{\partial\Omega}$ denote the duality pairing of $H^{1/2}(\partial\Omega)$ and its dual $H^{-1/2}(\Omega)$, which extends the L^2-scalar product on $\partial\Omega$. Let $|||\bullet||| := |\bullet|_{H^1(\Omega)} = \|\mathbf{D}\bullet\|_{L^2(\Omega)}$ abbreviate the energy norm.

Throughout the paper, $A \lesssim B$ abbreviates the relation $A \leqslant CB$ with a generic constant $0 < C$ which solely depends on the interior angles $\sphericalangle\mathcal{T}$ of the underlying triangulation $\mathcal{T}$ and so solely on the initial triangulation $\mathcal{T}_0$; $A \approx B$ abbreviates $A \lesssim B \lesssim A$.

This paper is restricted to 2D for the ease of this presentation although most of the arguments carry over to 3D as well, cf. [13] for some additional arguments in 3D and [14] for more involved modifications of the treatment of Dirichlet data.

2. Preliminaries

2.1. Triangulations

Let $\mathcal{T}_0$ denote a shape-regular initial triangulation of the simply-connected polygonal Lipschitz domain Ω into closed triangles. Any admissible refinement of it in this paper is understood in the context of newest-vertex bisection (NVB), see e.g. [15–17] for details on this method of mesh-refinement in 2D.

The NVB requires an initial condition on the refinement edges in $\mathcal{T}_0$. With reference to [18] for the suppressed details, this is assumed throughout this paper in the definition of the set

$$\begin{aligned}\mathbb{T} := \{&\mathcal{T}_\ell \text{ regular triangulation of } \Omega \text{ into triangles :}\\ &\exists \ell \in \mathbb{N}_0\ \exists \mathcal{T}_0, \mathcal{T}_1, \dots, \mathcal{T}_\ell \text{ successive one-level refinements in the sense}\\ &\text{that } \mathcal{T}_{j+1} \text{ is a one-level refinement of } \mathcal{T}_j \text{ for } j = 0, 1, \dots, \ell - 1\}.\end{aligned}$$

of admissible triangulations created from $\mathcal{T}_0$ for refinement control and existence of overlays as summarized in [3, Sect. 2.4] with further references. For any natural number $N \in \mathbb{N}$, set

$$\mathbb{T}(N) := \{\mathcal{T} \in \mathbb{T} : \left|\mathcal{T}\right| - \left|\mathcal{T}_0\right| \leqslant N\}.$$

For any $\mathcal{T} \in \mathbb{T}$, $\mathcal{N}$ denotes the set of nodes and $\mathcal{E}$ the set of edges and the corresponding sets $\mathcal{N}(\partial\Omega)$ and $\mathcal{E}(\partial\Omega)$ on the boundary $\partial\Omega$, $\mathcal{N}(\Omega)$ and $\mathcal{E}(\Omega)$ in the interior Ω. For a triangle $T \in \mathcal{T}$, let $\mathcal{N}(T)$ denote the set of its three nodes and $\mathcal{E}(T)$ the set of its three edges.

2.2. Approximation of dirichlet boundary data

Given some triangulation $\mathcal{T}$ of the polygonal domain Ω, let $H^1(\mathcal{E}(\partial\Omega); \mathbb{R}^2)$ consist of all boundary functions $g \in L^2(\partial\Omega; \mathbb{R}^2)$ with square-integrable piecewise arc-length derivative $g' = \partial g/\partial s \in L^2(\partial\Omega; \mathbb{R}^2)$ along the boundary edges $\mathcal{E}(\partial\Omega)$. Let $P_k(\mathcal{E}(\partial\Omega))$ denote the space of piecewise polynomials of degree at most $k \in \mathbb{N}_0$ on the boundary. For $k \in \mathbb{N}$ and any function $g \in H^1(\mathcal{E}(\partial\Omega); \mathbb{R}^2) \cap C(\partial\Omega; \mathbb{R}^2)$, let $I_k g \in S^k(\mathcal{E}(\partial\Omega); \mathbb{R}^2) := P_k(\mathcal{E}(\partial\Omega); \mathbb{R}^2) \cap C(\partial\Omega; \mathbb{R}^2)$ denote the nodal interpolation with respect to $k + 1$ distinct nodes $x_0, \dots, x_k \in E$ for any boundary edge $E = \operatorname{conv}\{x_0, x_k\} \in \mathcal{E}(\partial\Omega)$. If not stated otherwise, let the nodes be distributed equally in every boundary edge $E \in \mathcal{E}(\partial\Omega)$.

Lemma 2.1. *For any $E = \operatorname{conv}\{x_0, x_k\} \in \mathcal{E}(\partial\Omega)$ with arbitrary but distinct nodes $x_0, \dots, x_k \in E$ and $g \in H^1(E; \mathbb{R}^2) \cap C(E; \mathbb{R}^2)$, it holds that*

$$\left\|(I_k g)'\right\|_{L^2(E)} \leqslant C(x_0, \dots, x_k) \left\|g'\right\|_{L^2(E)}$$

with a constant $C(x_0, \dots, x_k)$ that does not depend on the length of E.

Proof. Without loss of generality, let $E = [0, h] = [x_0, x_k] \subset \mathbb{R}$. Hence, for any $x \in \mathbb{R}$,

$$I_k g(x) = \sum_{j=0}^{k} g(x_j) L_j(x) \quad \text{with } L_j(x) := \prod_{\ell=0,\dots,k,\ell\neq j} \frac{x - x_\ell}{x_j - x_\ell}.$$

Since $(I_k g)'$ and g' do not change from an additive shift of g, we may and will suppose without loss of generality that $g(0) = 0$. Thus,

$$\left|g(x_j)\right| \leqslant \int_0^{x_j} \left|g'(y)\right| \,\mathrm{d}y \leqslant \sqrt{h} \left\|g'\right\|_{L^2(E)}.$$

Moreover,

$$\left\|(I_k g)'\right\|_{L^2(E)} \leqslant \sum_{j=0}^{k} \left|g(x_j)\right| \left\|L_j'\right\|_{L^2(E)} \leqslant \sqrt{h} \left\|g'\right\|_{L^2(E)} \sum_{j=0}^{k} \left\|L_j'\right\|_{L^2(E)}.$$

This implies the claim via an inverse estimate for the polynomial L_j, namely

$$\left\|L_j'\right\|_{L^2(E)} \leqslant C_{\mathrm{inv}} h^{-1} \left\|L_j\right\|_{L^2(E)} \leqslant C_{\mathrm{inv}} h^{-1/2} \left\|L_j\right\|_{L^\infty(E)}.$$

For uniformly distributed nodes, the Lebesgue constant $\sum_{j=0}^{k} \|L_j\|_{L^\infty(E)}$ satisfies [19]

$$\sum_{j=0}^{k} \left\|L_j\right\|_{L^\infty(E)} \approx \frac{2^{k+1}}{ek \log k}.$$

In general, the constant $C(x_0, \dots, x_k) := C_{\mathrm{inv}} \sum_{j=0}^{k} \left\|L_j\right\|_{L^\infty(E)}$ depends on k and the partition $x_0, \dots, x_k$ of interpolation nodes, but it is independent of the size h. It degenerates to ∞ as $\min_{j\neq k} |x_j - x_k|/h$ tends to zero or as k tends to infinity. □

Lemma 2.1 asserts a generic stability constant $C_{\mathrm{stab}} > 0$ of the nodal interpolation I_k such that, for all $g \in H^1(\mathcal{E}(\partial\Omega);\mathbb{R}^2)\cap C(\partial\Omega;\mathbb{R}^2)$,

$$\big\|(I_k g)'\big\|_{L^2(\partial\Omega)} \leqslant C_{\mathrm{stab}}\big\|g'\big\|_{L^2(\partial\Omega)}. \tag{5}$$

Let $J_k g \in S^k(\mathcal{E}(\partial\Omega);\mathbb{R}^2)$ denote an alternative interpolation of g with $(J_k g)(z) = g(z)$ for all $z \in \mathcal{N}(\partial\Omega)$ and, if $k \geqslant 2$,

$$\int_E (g - J_k g)\varphi_{k-2}\,\mathrm{d}s = 0 \quad \text{for all } \varphi_{k-2} \in P_{k-2}(E) \text{ and } E \in \mathcal{E}(\partial\Omega). \tag{6}$$

This defines J_k uniquely and for the $L^2(\partial\Omega)$-orthogonal projection $\Pi_{k-1}g'$ of g' onto $P_{k-1}(\mathcal{E}(\partial\Omega);\mathbb{R}^2)$, (6) ensures that

$$(J_k g)' = \Pi_{k-1}g'. \tag{7}$$

With the k-independent stability constant 1 of Π_{k-1}, this implies boundedness of J_k,

$$\big\|(J_k g)'\big\|_{L^2(\partial\Omega)} = \big\|\Pi_{k-1}g'\big\|_{L^2(\partial\Omega)} \leqslant \big\|g'\big\|_{L^2(\partial\Omega)}.$$

Recall that C_{stab} denotes the constant in (5).

Lemma 2.2. *It holds* $\|((1-I_k)g)'\|_{L^2(E)} \leqslant (1+C_{\mathrm{stab}})\|((1-J_k)g)'\|_{L^2(E)}$ *for any* $E \in \mathcal{E}(\partial\Omega)$.

Proof. Since $I_k J_k g = J_k g$, Lemma 2.1 implies

$$\big\|((1-I_k)g)'\big\|_{L^2(E)} \leqslant \big\|((1-J_k)g)'\big\|_{L^2(E)} + \big\|(I_k(1-J_k)g)'\big\|_{L^2(E)} \leqslant (1+C_{\mathrm{stab}})\big\|((1-J_k)g)'\big\|_{L^2(E)}. \quad \square$$

Let $h_{\mathcal{E}} \in P_0(\mathcal{E})$ denote the piecewise constant function with $h_{\mathcal{E}}|_E \equiv |\omega_E|^{1/2}$ for every $E \in \mathcal{E}(\partial\Omega)$ and its adjacent triangle $\bar{\omega}_E \in \mathcal{T}$ to define the Dirichlet data oscillation

$$\operatorname{osc}(g', \mathcal{E}(\partial\Omega)) := \big\|h_{\mathcal{E}}^{1/2}(1-\Pi_{k-1})g'\big\|_{L^2(\partial\Omega)}.$$

The proof of the following lemma employs the ideas from [5, Lem. 2.1]. In addition, it utilizes the stability from Lemma 2.2 to control the discrete approximation of the Dirichlet boundary data. Therefore, the generic constant $C_{\mathrm{ext}} > 0$ in (8) includes C_{stab} from Lemma 2.3 and, thus, depends on the polygonal degree k. Let $C_{\mathrm{SZ}} > 0$ denote the stability constant from [20, Thm. 3.1] of the Scott–Zhang quasi-interpolation which also depends on k.

Lemma 2.3. *Given any boundary data* $g \in H^1(\mathcal{E}(\partial\Omega);\mathbb{R}^2) \cap C(\partial\Omega;\mathbb{R}^2)$, *there exists some extension* $w \in H^1(\Omega;\mathbb{R}^2)$ *with*

$$w\big|_{\partial\Omega} = (1-I_k)g \quad \textit{and} \quad |||w||| \leqslant C_{\mathrm{ext}}\operatorname{osc}(g', \mathcal{E}(\partial\Omega)). \tag{8}$$

If in addition $g \equiv \widehat{g} \in S^k(\widehat{\mathcal{E}}(\partial\Omega);\mathbb{R}^2)$ *for any admissible refinement* $\widehat{\mathcal{T}}$ *of* $\mathcal{T}$, *there exists some discrete extension* $w \in S^k(\widehat{\mathcal{T}};\mathbb{R}^2)$ *with* (8) *and the modified constant* $\widehat{C}_{\mathrm{ext}} := C_{\mathrm{ext}}C_{\mathrm{SZ}}$.

Proof. An analogous proof to the one of [5, Lem. 2.1] yields existence of a generic constant $C > 0$ and $w \in H^1(\Omega;\mathbb{R}^2)$ with

$$w\big|_{\partial\Omega} = (1-I_k)g \quad \text{and} \quad |||w||| \leqslant C\big\|h_{\mathcal{E}}^{1/2}((1-I_k)g)'\big\|_{L^2(\partial\Omega)}.$$

The constant $C > 0$ consists of the universal equivalent constant of different $H^{1/2}$-norms times the generic constant from [21, Thm. 1] solely depending on the ratio $r := \max\{|E_j|/|E_k| : E_j, E_k \in \mathcal{E}(\partial\Omega) \text{ and } E_j \cap E_k \neq \emptyset\}$.

With $C_{\mathrm{ext}} := C(1+C_{\mathrm{stab}})$, Lemma 2.2 and the commutativity (7) of J_k complete the proof of

$$|||w||| \leqslant C_{\mathrm{ext}}\operatorname{osc}(g', \mathcal{E}(\partial\Omega)).$$

The existence of a discrete extension employs the Scott–Zhang quasi-interpolation with stability constant C_{SZ} and follows verbatim as in the proof of [5, Lem. 2.1] with the modified constant $\widehat{C}_{\mathrm{ext}} := C_{\mathrm{ext}}C_{\mathrm{SZ}}$. $\square$

Lemma 2.4. *The constant* $C_{\mathrm{osc}} := \sqrt{2}C_{\mathrm{stab}}^2/(\sqrt{2}-1) > 0$ *and any admissible refinement* $\widehat{\mathcal{T}}$ *of* $\mathcal{T}$ *satisfy*

$$\operatorname{osc}^2((\widehat{I}_k g)', \mathcal{E}(\partial\Omega)) \leqslant C_{\mathrm{osc}}\big(\operatorname{osc}^2(g', \mathcal{E}(\partial\Omega)) - \operatorname{osc}^2(g', \widehat{\mathcal{E}}(\partial\Omega))\big).$$

Proof. Since the left-hand side equals

$$\operatorname{osc}^2((\widehat{I}_k g)', \mathcal{E}(\partial\Omega)) = \sum_{E\in\mathcal{E}(\partial\Omega)\setminus\widehat{\mathcal{E}}(\partial\Omega)} |\omega_E|^{1/2} \|(1-\Pi_{k-1})(\widehat{I}_k g)'\|^2_{L^2(E)}, \tag{9}$$

it remains to consider $E \in \mathcal{E}(\partial\Omega)\setminus\widehat{\mathcal{E}}(\partial\Omega)$ with (even with equality in the first step)

$$\|(1-\Pi_{k-1})(\widehat{I}_k g)'\|_{L^2(E)} = \|(1-\Pi_{k-1})((\widehat{I}_k g)' - (J_k g)')\|_{L^2(E)} \leqslant \|(\widehat{I}_k g)' - (J_k g)'\|_{L^2(E)}.$$

Lemma 2.1 (with $g - J_k g$ replacing g) and (7) imply the upper bound

$$\|(\widehat{I}_k(g - J_k g))'\|_{L^2(E)} \leqslant C_{\mathrm{stab}} \|(1-\Pi_{k-1})g'\|_{L^2(E)}.$$

This leads to

$$|\omega_E|^{1/2} \|(1-\Pi_{k-1})(\widehat{I}_k g)'\|^2_{L^2(E)} \leqslant C^2_{\mathrm{stab}} |\omega_E|^{1/2} \sum_{F\in\widehat{\mathcal{E}}(E)} \|(1-\Pi_{k-1})g'\|^2_{L^2(F)}. \tag{10}$$

Since $E \in \mathcal{E}(\partial\Omega)\setminus\widehat{\mathcal{E}}(\partial\Omega)$, any $F \in \widehat{\mathcal{E}}(E)$ has a subordinated subtriangle $\omega_F \in \widehat{\mathcal{T}}$ of $\omega_E \in \mathcal{T}$ that is refined and so $|\omega_F| \leqslant |\omega_E|/2$. This and $\|(1-\widehat{\Pi}_{k-1})g'\|_{L^2(F)} \leqslant \|(1-\Pi_{k-1})g'\|_{L^2(F)}$ result in

$$\begin{aligned}(1-1/\sqrt{2})|\omega_E|^{1/2} \|(1-\Pi_{k-1})g'\|^2_{L^2(F)} &\leqslant \big(|\omega_E|^{1/2} - |\omega_F|^{1/2}\big) \|(1-\Pi_{k-1})g'\|^2_{L^2(F)} \\ &\leqslant |\omega_E|^{1/2} \|(1-\Pi_{k-1})g'\|^2_{L^2(F)} - |\omega_F|^{1/2} \|(1-\widehat{\Pi}_{k-1})g'\|_{L^2(F)} \\ &= |\omega_E|^{1/2} \|(1-\Pi_{k-1})g'\|^2_{L^2(F)} - \operatorname{osc}^2(g', F).\end{aligned}$$

The substitution of this in (10) shows

$$|\omega_E|^{1/2} \|(1-\Pi_{k-1})(\widehat{I}_k g)'\|^2_{L^2(E)} \leqslant C_{\mathrm{osc}}\big(\operatorname{osc}^2(g', E) - \operatorname{osc}^2(g', \widehat{\mathcal{E}}(E))\big).$$

This and (9) conclude the proof. □

Remark 2.5. The commutativity (7) implies

$$\begin{aligned}\|h_{\mathcal{E}}^{1/2}(1-\Pi_{k-1})(\widehat{J}_k g)'\|^2_{L^2(\partial\Omega)} &= \|h_{\mathcal{E}}^{1/2}(\widehat{\Pi}_{k-1} - \Pi_{k-1})g'\|^2_{L^2(\partial\Omega)} \\ &\leqslant \|h_{\mathcal{E}}^{1/2}(1-\Pi_{k-1})g'\|^2_{L^2(\partial\Omega)} - \|h_{\widehat{\mathcal{E}}}^{1/2}(1-\widehat{\Pi}_{k-1})g'\|^2_{L^2(\partial\Omega)}\end{aligned}$$

(with $C_{\mathrm{osc}} = 1$). Hence, the interpolation $J_k g$ is an alternative choice for the approximation of the Dirichlet boundary data in Section 2.5.

Remark 2.6. The corresponding situation in 3D is more involved and does not allow the analysis of this section because (7) fails. This led in [14] to different algorithms to circumvent this difficulty with a separate Dörfler marking to control the boundary approximation terms.

2.3. Auxiliary problem

Given $f \in L^2(\Omega;\mathbb{R}^2)$ and $g \in H^{1/2}(\mathcal{E}(\partial\Omega);\mathbb{R}^2) \cap C(\partial\Omega;\mathbb{R}^2)$, let $z \in H^1(\Omega;\mathbb{R}^2)$ denote the unique solution to

$$-\Delta z = f \quad \text{in } \Omega \quad \text{and} \quad z|_{\partial\Omega} = g \quad \text{on } \partial\Omega.$$

Since $f \in L^2(\Omega;\mathbb{R}^2)$ and $g \in H^{1/2}(\mathcal{E}(\partial\Omega);\mathbb{R}^2) \cap C(\partial\Omega;\mathbb{R}^2)$, the reduced elliptic regularity [22–24] implies $\boldsymbol{\tau} := \mathbf{D}z \in H^s(\Omega;\mathbb{R}^{2\times 2})$ for some $s > 0$ and

$$\|\boldsymbol{\tau}\|_{H(\mathrm{div},\Omega)} \lesssim \|f\|_{L^2(\Omega)} + \|g\|_{H^{1/2}(\partial\Omega)}.$$

Let $\boldsymbol{\tau}_{\mathrm{RT}} := \mathcal{I}_{\mathrm{F}}\boldsymbol{\tau} \in RT_k(\mathcal{T};\mathbb{R}^{2\times 2})$ denote the Fortin interpolation of $\boldsymbol{\tau}$ [25, eq. (2.5.26)]. Then it holds $-\operatorname{div}\boldsymbol{\tau}_{\mathrm{RT}} = -\Pi_k \operatorname{div}\boldsymbol{\tau} = \Pi_k f$ and

$$\|\boldsymbol{\tau}_{\mathrm{RT}}\|_{H(\mathrm{div},\Omega)} \lesssim \|f\|_{L^2(\Omega)} + \|g\|_{H^{1/2}(\partial\Omega)}.$$

2.4. Divergence-free Raviart–Thomas functions

The 2D rotation operators read, for $v \in H^1(\Omega;\mathbb{R}^2)$,

$$\mathbf{Curl} v := \begin{pmatrix} -\partial v_1/\partial x_2 & \partial v_1/\partial x_1 \\ -\partial v_2/\partial x_2 & \partial v_2/\partial x_1 \end{pmatrix} \quad \text{and} \quad \operatorname{curl} v := \operatorname{tr} \mathbf{Curl} v.$$

Let $C_P(\Omega) > 0$ denote the Poincaré constant such that any $\beta \in H^1(\Omega) \cap L^2_0(\Omega)$ satisfies

$$\|\beta\|_{L^2(\Omega)} \leqslant C_P(\Omega) \|\nabla \beta\|_{L^2(\Omega)}.$$

Proposition 2.7. *Given $\boldsymbol{\rho}_{RT} \in RT_k(\mathcal{T};\mathbb{R}^{2\times 2})$ with $\operatorname{div} \boldsymbol{\rho}_{RT} = 0$, there exists $\beta_C \in S^k(\mathcal{T};\mathbb{R}^2)$ such that $\mathbf{Curl}\beta_C = \boldsymbol{\rho}_{RT}$ and*

$$\|\beta_C\|_{L^2(\Omega)} \leqslant C_P(\Omega) \|\boldsymbol{\rho}_{RT}\|_{H(\operatorname{div},\Omega)}.$$

Proof. Since Ω is simply-connected, a known discrete Helmholtz decomposition [10,9,11] guarantees the existence of $\beta_C \in S^k(\Omega;\mathbb{R}^2)$ with $\mathbf{Curl}\beta_C = \boldsymbol{\rho}_{RT}$. Without loss of generality, it holds that $\int_\Omega \beta_C \,dx = 0$ and the Poincaré inequality proves

$$C_P^{-1} \|\beta_C\|_{L^2(\Omega)} \leqslant |||\beta_C||| = \|\mathbf{Curl}\beta_C\|_{L^2(\Omega)} = \|\boldsymbol{\rho}\|_{H(\operatorname{div},\Omega)}. \quad \square$$

2.5. Least-squares FEM

The space of discrete admissible pseudostress functions on a regular triangulation $\mathcal{T}$ reads

$$\boldsymbol{\Sigma}_{k-1}(\mathcal{T}) := \{\boldsymbol{\tau}_{RT} \in RT_{k-1}(\mathcal{T};\mathbb{R}^2) : \operatorname{tr}\boldsymbol{\tau} \in L^2_0(\Omega)\}$$

and the set of discrete admissible velocity functions reads

$$\mathcal{A}_k(\mathcal{T}) := \{v \in S^k(\mathcal{T};\mathbb{R}^2) : v = I_k g \text{ on } \partial\Omega\}.$$

A conforming discretization seeks $(\boldsymbol{\sigma}_{LS}, u_{LS}) \in \boldsymbol{\Sigma}_{k-1}(\mathcal{T}) \times \mathcal{A}_k(\mathcal{T})$ such that, for all $(\boldsymbol{\tau}_{LS}, v_{LS}) \in \boldsymbol{\Sigma}_{k-1}(\mathcal{T}) \times S^k_0(\mathcal{T};\mathbb{R}^2)$,

$$\int_\Omega \operatorname{div}\boldsymbol{\sigma}_{LS} \cdot \operatorname{div}\boldsymbol{\tau}_{LS}\,dx + \int_\Omega (\mathbf{dev}\boldsymbol{\sigma}_{LS} - \mathbf{D}u_{LS}) : (\mathbf{dev}\boldsymbol{\tau}_{LS} - \mathbf{D}v_{LS})\,dx = -\int_\Omega f \cdot \operatorname{div}\boldsymbol{\tau}_{LS}\,dx. \tag{11}$$

This yields the unique discrete minimizer of the least-squares functional $LS(f;\boldsymbol{\sigma}_{LS},u_{LS})$ defined in (2). Given $\mathcal{T}$ and an admissible refinement $\widehat{\mathcal{T}}$, let $(\widehat{\boldsymbol{\sigma}}_{LS},\widehat{u}_{LS}) \in \boldsymbol{\Sigma}_{k-1}(\widehat{\mathcal{T}}) \times \mathcal{A}_k(\widehat{\mathcal{T}})$ denote the subordinated discrete solution.

Proposition 2.8. *The constant $C_{qo} := C_{ext}C_{osc}C_{SZ}$ and any $\varepsilon > 0$ satisfy*

$$\begin{aligned}&(1-\varepsilon)LS(f;\widehat{\boldsymbol{\sigma}}_{LS},\widehat{u}_{LS}) + LS(0;\widehat{\boldsymbol{\sigma}}_{LS}-\boldsymbol{\sigma}_{LS},\widehat{u}_{LS}-u_{LS})\\ &\quad\leqslant LS(f;\boldsymbol{\sigma}_{LS},u_{LS}) + C_{qo}/\varepsilon\big(\operatorname{osc}^2(g',\mathcal{E}(\partial\Omega)) - \operatorname{osc}^2(g',\widehat{\mathcal{E}}(\partial\Omega))\big).\end{aligned} \tag{12}$$

Proof. The proof employs the arguments from [5, Thm. 4.8] and is merely outlined here for brevity. Throughout this proof, the upper index $\widehat{\ }$ refers to $\widehat{\mathcal{T}}$. Lemma 2.3 guarantees the existence of some $\widehat{w} \in S^k(\widehat{\mathcal{T}};\mathbb{R}^2)$ with $\widehat{w}|_{\partial\Omega} = (\widehat{I}_k - I_k)g$ and

$$|||\widehat{w}|||^2 \leqslant C_{ext}C_{SZ} \operatorname{osc}^2((\widehat{I}_k g)', \mathcal{E}(\partial\Omega)). \tag{13}$$

Elementary algebra proves

$$LS(0;\widehat{\boldsymbol{\sigma}}_{LS}-\boldsymbol{\sigma}_{LS},\widehat{u}_{LS}-u_{LS}) = LS(f;\boldsymbol{\sigma}_{LS},u_{LS}) - LS(f;\widehat{\boldsymbol{\sigma}}_{LS},\widehat{u}_{LS}) - 2\int_\Omega (\mathbf{dev}\widehat{\boldsymbol{\sigma}}_{LS} - \mathbf{D}\widehat{u}_{LS}) : \mathbf{D}\widehat{w}\,dx.$$

This, the Cauchy–Schwarz inequality, the Young inequality for any parameter $\varepsilon > 0$, (13), and Lemma 2.4 conclude the proof. $\square$

3. Scaling

This section is devoted to an analysis of the norm equivalence constants of the homogeneous least-squares functional in (3), which in general are dependent on the size of the domain Ω. This scaling dependence is sometimes subject to a

general criticism on least-squares methods. However, the main result Theorem 3.1 presents a weighting of the least-squares functional and the norms of the function spaces that guarantee scaling independent equivalence constants.

Let $C_{\mathrm{L}} > 0$ denote the generic constant from the Ladyshenskaya lemma in [26, Thm. 3.1] and $C_{\mathrm{F}} > 0$ the Friedrichs constant [27, Thm. 1.5] which depend on the domain Ω and lead to the well-established estimates

$$\|v\|_{L^2(\Omega)} \leqslant C_{\mathrm{F}} \|\mathbf{D}v\|_{L^2(\Omega)} \quad \text{for all } v \in H^1_0(\Omega;\mathbb{R}^2)$$

and, since Ω is simply-connected, for all $q \in L^2_0(\Omega;\mathbb{R}^2)$, there exists $\boldsymbol{\tau} \in H^1_0(\Omega;\mathbb{R}^{2\times 2})$ such that $q = \operatorname{div}\boldsymbol{\tau}$ and

$$\|\mathbf{D}\boldsymbol{\tau}\|_{L^2(\Omega)} \leqslant C_{\mathrm{L}} \|q\|_{L^2(\Omega)}.$$

Given two parameters $\kappa, \mu > 0$, define a weighted least-squares functional, for $(\boldsymbol{\tau}, v) \in H(\operatorname{div}, \Omega;\mathbb{R}^{2\times 2}) \times H^1_0(\Omega;\mathbb{R}^2)$, by

$$LS(\kappa, \mu; \boldsymbol{\tau}, v) := \|\mathbf{dev}(\boldsymbol{\tau} - \mathbf{D}v)\|^2_{L^2(\Omega)} + \kappa C_{\mathrm{F}}^2 \|\operatorname{div}\boldsymbol{\tau}\|^2_{L^2(\Omega)} + \mu C_{\mathrm{L}}^2 \|\operatorname{div} v\|^2_{L^2(\Omega)} \tag{14}$$

and set

$$C_-(\kappa, \mu) := \frac{3 + \min\{\kappa, \mu\}}{4} - \frac{1}{4}\sqrt{(1 - \min\{\kappa, \mu\})^2 + 8} > 0, \tag{15}$$

$$C_+(\kappa, \mu) := \frac{3 + \min\{\kappa, \mu\}}{4} + \frac{1}{4}\sqrt{(1 - \min\{\kappa, \mu\})^2 + 8} > 0. \tag{16}$$

Theorem 3.1. *Any $(\boldsymbol{\tau}, v) \in \boldsymbol{\Sigma} \times H^1_0(\Omega;\mathbb{R}^2)$ satisfies*

$$\begin{aligned} &C_-(\kappa, \mu)\big(\|\mathbf{dev}\boldsymbol{\tau}\|^2_{L^2(\Omega)} + C_{\mathrm{F}}^2\|\operatorname{div}\boldsymbol{\tau}\|^2_{L^2(\Omega)} + \|\mathbf{dev}\mathbf{D}v\|^2_{L^2(\Omega)} + C_{\mathrm{L}}^2\|\operatorname{div} v\|^2_{L^2(\Omega)}\big) \\ &\quad\leqslant LS(\kappa, \mu; \boldsymbol{\tau}, v) \\ &\quad\leqslant C_+(\kappa, \mu)\big(\|\mathbf{dev}\boldsymbol{\tau}\|^2_{L^2(\Omega)} + C_{\mathrm{F}}^2\|\operatorname{div}\boldsymbol{\tau}\|^2_{L^2(\Omega)} + \|\mathbf{dev}\mathbf{D}v\|^2_{L^2(\Omega)} + C_{\mathrm{L}}^2\|\operatorname{div} v\|^2_{L^2(\Omega)}\big). \end{aligned}$$

Several remarks are in order before the proof of the Theorem 3.1 concludes this section.

First, the weights C_{L} and C_{F} scale differently with the size of Ω. In general, the results of least-squares calculation depend on $\operatorname{diam}(\Omega)$ owing to the different scaling of functions and their derivatives in $L^2(\Omega)$. It is known that C_{L} does *not* scale with $\operatorname{diam}(\Omega)$, while the first Dirichlet eigenvalue λ_1 of the Laplacian leads to $C_{\mathrm{F}}^2 = 1/\lambda_1$ and C_{F} is proportional to $\operatorname{diam}(\Omega)$. Moreover, it is well-known and follows with a one-dimensional Friedrichs inequality $\|f\|_{L^2(a,b)} \leqslant (b-a)/\pi \, \|f'\|_{L^2(a,b)}$ and Fubini's theorem for the width $\operatorname{width}(\Omega)$ of Ω, that

$$C_{\mathrm{F}} \leqslant \operatorname{width}(\Omega)/\pi.$$

The choice of the weights in (14)–(16) reflects this behaviour. This is why we introduced the values κ and μ as scaling invariant variables, which lead to equivalence constants $C_\pm(\kappa, \mu)$ independent of the domain Ω.

This is indeed different for the natural choice of the weights. In fact, $\kappa = 1/C_{\mathrm{F}}^2$ and $\mu = 1/(4C_{\mathrm{L}}^2)$ lead to (2) and, hence, $C(\kappa, \mu) = C(1/C_{\mathrm{F}}^2, 1/(4C_{\mathrm{L}}^2))$ in Theorem 3.1 scales with the size of Ω.

The proof of Theorem 3.1 employs the following special case of the well-known tr-dev-div lemma.

Lemma 3.2. *Any $\phi \in H^1(\Omega;\mathbb{R}^2)$ with $\operatorname{curl}\phi \in L^2_0(\Omega)$ satisfies*

$$\|\operatorname{curl}\phi\|_{L^2(\Omega)} \leqslant 2C_{\mathrm{L}} \|\mathbf{devCurl}\phi\|_{L^2(\Omega)}.$$

Proof. Since $\operatorname{curl}\phi \in L^2_0(\Omega)$, the Ladyshenskaya lemma [26, Thm. 3.1] yields existence of some $\psi \in H^1_0(\Omega;\mathbb{R}^2)$ with $\operatorname{div}\psi = \operatorname{curl}\phi$ and $|||\psi||| \leqslant C_{\mathrm{L}}\|\operatorname{curl}\phi\|_{L^2(\Omega)}$. Hence,

$$\begin{aligned} \frac{1}{2}\|\operatorname{curl}\phi\|^2_{L^2(\Omega)} &= \frac{1}{2}\int_\Omega \mathbf{D}\psi : (\operatorname{curl}\phi\, I_{2\times 2})\,\mathrm{d}x \\ &= \int_\Omega \mathbf{D}\psi : (\mathbf{Curl}\phi - \mathbf{devCurl}\phi)\,\mathrm{d}x \\ &\leqslant |||\psi|||\,\|\mathbf{devCurl}\phi\|_{L^2(\Omega)}. \end{aligned}$$

This concludes the proof. □

A technical lemma is required in the proof of Theorem 3.1.

Lemma 3.3. *Given $\kappa > 0$ and $X := \{y \in [0,\infty)^3 : |y|^2 := y_1^2 + y_2^2 + y_3^2 > 0\}$, it holds*

$$\min_{x\in X} \frac{x_1^2 + x_2^2 + \kappa\, x_3^2 - 2x_2 \min\{x_1, x_3\}}{|x|^2} = \frac{3+\kappa}{4} - \frac{1}{4}\sqrt{(1-\kappa)^2 + 8} > 0;$$
$$\max_{x\in X} \frac{x_1^2 + x_2^2 + \kappa\, x_3^2 + 2x_2 \min\{x_1, x_3\}}{|x|^2} = \frac{3+\kappa}{4} + \frac{1}{4}\sqrt{(1-\kappa)^2 + 8} > 0.$$

Proof. *Step* 1. For $x \in X$, set

$$f(x) := (x_1^2 + x_2^2 + \kappa\, x_3^2 - 2x_2 \min\{x_1, x_3\}) \big/ |x|^2. \tag{17}$$

This step establishes the claimed formula for

$$M := \inf_{0<t<\infty, 0\leqslant x_2} f(t, x_2, t).$$

Since $f(t, 0, t) = (1+\kappa)/2$ for all $0 < t < \infty$, it remains to consider $f(t, x_2, t)$ for $x_2 > 0$. A rescaling allows a reduction to $x_2 = 1$. Since

$$f(t, 1, t) = \frac{\kappa t^2 + (t-1)^2}{2t^2 + 1}$$

has a negative derivative $\partial f(t, 1, t)/\partial t$ for $0 < t < t^*$ and a positive derivative for $t^* < t$ with the unique zero

$$t^* = \frac{1-\kappa}{4} + \frac{1}{4}\sqrt{(1-\kappa)^2 + 8},$$

it follows

$$M = f(t^*, 1, t^*) = \frac{3+\kappa}{4} - \frac{1}{4}\sqrt{(1-\kappa)^2 + 8} > 0.$$

Step 2. Since $m := \min_{x\in X} f(x) \leqslant M$ is obvious, the remaining parts of the proof are devoted to the converse inequality.
Step 3. If $0 < \kappa \leqslant 1$ and $x \in X$, then $\min\{x_1^2, x_3^2\} + \kappa \max\{x_1^2, x_3^2\} \leqslant x_1^2 + \kappa x_3^2$ implies

$$f(\min\{x_1, x_3\}, x_2, \max\{x_1, x_3\}) \leqslant f(x).$$

(If $\kappa \geqslant 1$, the aforementioned inequality holds when min and max are interchanged.)
Step 4. If $0 < \kappa < 1$ and $x \in X$, then $f(x) \geqslant M$. To see this, consider $x \in X$ and utilize Step 3 to reduce to $0 \leqslant x_1 \leqslant x_3$ and $0 \leqslant x_2$. Since $\min\{x_1, x_3\} = x_1$, elementary algebra shows

$$f(x) = \kappa + (1-\kappa)(x_1^2 + x_2^2 - 2x_1x_2/(1-\kappa)) \big/ |x|^2. \tag{18}$$

The infimum of f will be smaller than κ only in case that $x_1^2 + x_2^2 - 2x_1x_2/(1-\kappa) < 0$ and then (with x_1, x_2 fixed), (18) shows that $f(x)$ is monotonically increasing with x_3. The infimum is therefore achieved for minimal $x_1 = x_3$. Thus $f(x) \geqslant M$.
Step 5. If $\kappa \geqslant 1$, the Step 3 shows, for all $x \in X$, that

$$f(\max\{x_1, x_3\}, x_2, \min\{x_1, x_3\}) \leqslant f(x).$$

This leads to the case $0 \leqslant x_3 \leqslant x_1$, $0 \leqslant x_2$, and

$$f(x) = 1 + ((\kappa-1)x_3^2 - 2x_2x_3) \big/ |x|^2.$$

With fixed x_2, x_3 and $(\kappa-1)x_3^2 - 2x_2x_3 \leqslant 0$, $f(x)$ is monotonically increasing in x_1. This shows $f(x) \geqslant f(x_1, x_2, x_1) \geqslant M$.
Step 6. Analogous arguments yield the upper bound with the only positive root

$$t^* = \frac{\kappa-1}{4} + \frac{1}{4}\sqrt{(\kappa-1)^2 + 8}$$

of the derivative $\partial f(t, 1, t)/\partial t$ of

$$f(t, 1, t) = \frac{\kappa t^2 + (t+1)^2}{2t^2 + 1}$$

and

$$\sup_{0<t<\infty} f(t, 1, t) = f(t^*, 1, t^*) = \frac{\kappa+3}{4} + \frac{1}{4}\sqrt{(\kappa-1)^2 + 8}. \quad \square$$

Proof of Theorem 3.1. *Step* 1. For $Z := \{\alpha \in H^1_0(\Omega;\mathbb{R}^2) : \operatorname{div}\alpha = 0\}$, the Helmholtz decomposition for a simply-connected domain Ω reads

$$L^2(\Omega;\mathbf{dev}\mathbb{R}^{2\times 2}) = \mathbf{D}Z \oplus (\mathbf{D}Z)^\perp \tag{19}$$

and is $L^2(\Omega)$ orthogonal with the orthogonal complement

$$\begin{aligned}(\mathbf{D}Z)^\perp &= \mathbf{dev}(\mathbf{Curl}H^1(\Omega;\mathbb{R}^2)/\mathbb{R}^3)\\ &:= \left\{\mathbf{devCurl}\phi : \phi \in H^1(\Omega;\mathbb{R}^2) \text{ with } \int_\Omega \phi\,\mathrm{d}x = 0 \text{ and } \int_\Omega \operatorname{curl}\phi\,\mathrm{d}x = 0\right\}.\end{aligned}$$

The authors expect that the above representation of $(\mathbf{D}Z)^\perp$ is well-known and deduce this from a discrete variant [28, Thm. 3.2] for the mesh-size in the underlying triangulation tends to zero. This implies existence of $\alpha, a \in Z$ and $\boldsymbol{\beta}, \mathbf{b} \in (\mathbf{D}Z)^\perp$ with

$$\mathbf{dev}\boldsymbol{\tau} = \mathbf{D}\alpha + \boldsymbol{\beta} \quad \text{and} \quad \mathbf{dev}\mathbf{D}v = \mathbf{D}a + \boldsymbol{b}.$$

Let $\phi \in H^1(\Omega;\mathbb{R}^2)$ such that $\boldsymbol{\beta} = \mathbf{devCurl}\phi = \mathbf{Curl}\phi - \operatorname{curl}\phi/2\, I_{2\times 2}$.

Step 2. The orthogonality from the Helmholtz decomposition (19), an integration by parts, the Cauchy–Schwarz inequality, and Lemma 3.2 yield

$$\begin{aligned}\int_\Omega \boldsymbol{b} : \boldsymbol{\beta}\,\mathrm{d}x &= \int_\Omega \boldsymbol{\beta} : \mathbf{dev}\mathbf{D}v\,\mathrm{d}x = \int_\Omega \mathbf{Curl}\phi : \mathbf{dev}\mathbf{D}v\,\mathrm{d}x\\ &= -\frac{1}{2}\int_\Omega \mathbf{Curl}\phi : (\operatorname{div} v I_{2\times 2})\,\mathrm{d}x \leqslant \frac{1}{2}\|\operatorname{curl}\phi\|_{L^2(\Omega)}\|\operatorname{div} v\|_{L^2(\Omega)}\\ &\leqslant C_{\mathrm{L}}\|\boldsymbol{\beta}\|_{L^2(\Omega)}\|\operatorname{div} v\|_{L^2(\Omega)}.\end{aligned}$$

This and the Cauchy–Schwarz inequality prove

$$\left|\int_\Omega \boldsymbol{b} : \boldsymbol{\beta}\,\mathrm{d}x\right| \leqslant \|\boldsymbol{\beta}\|_{L^2(\Omega)} \min\left\{\|\boldsymbol{b}\|_{L^2(\Omega)}, C_{\mathrm{L}}\|\operatorname{div} v\|_{L^2(\Omega)}\right\}.$$

Step 3. The orthogonality from the Helmholtz decomposition (19), an integration by parts, the Cauchy–Schwarz and Friedrichs inequality yield

$$\int_\Omega \mathbf{D}a : \mathbf{D}\alpha\,\mathrm{d}x = \int_\Omega \mathbf{D}a : \boldsymbol{\tau}\,\mathrm{d}x = -\int_\Omega a\cdot\operatorname{div}\boldsymbol{\tau}\,\mathrm{d}x \leqslant \|a\|_{L^2(\Omega)}\|\operatorname{div}\boldsymbol{\tau}\|_{L^2(\Omega)} \leqslant C_{\mathrm{F}}|||a|||\,\|\operatorname{div}\boldsymbol{\tau}\|_{L^2(\Omega)}.$$

This and the Cauchy–Schwarz inequality prove

$$\left|\int_\Omega \mathbf{D}a : \mathbf{D}\alpha\,\mathrm{d}x\right| \leqslant |||a||| \min\left\{|||\alpha|||, C_{\mathrm{F}}\|\operatorname{div}\boldsymbol{\tau}\|_{L^2(\Omega)}\right\}.$$

Step 4. Abbreviate

$$x := \left(|||\alpha|||, |||a|||, C_{\mathrm{F}}\|\boldsymbol{\tau}\|_{L^2(\Omega)}, \|\boldsymbol{b}\|_{L^2(\Omega)}, \|\boldsymbol{\beta}\|_{L^2(\Omega)}, C_{\mathrm{L}}\|\operatorname{div}\boldsymbol{\tau}\|_{L^2(\Omega)}\right) \in \mathbb{R}^6.$$

In this notation and with $f = f_\kappa$ from (17) and f_μ from (17) with κ replaced by μ, the combination of Step 2 and 3 lead to

$$\begin{aligned}f_\kappa(x_1,x_2,x_3) + f_\mu(x_4,x_5,x_6) &= x_1^2 + x_2^2 + \kappa x_3^2 - 2x_2\min\{x_1,x_3\} + x_4^2 + x_5^2 + \mu x_6^2 - 2x_5\min\{x_4,x_6\}\\ &\leqslant x_1^2 + x_2^2 + \kappa x_3^2 + x_4^2 + x_5^2 + \mu x_6^2 - 2\int_\Omega \mathbf{D}a : \mathbf{D}\alpha\,\mathrm{d}x - 2\int_\Omega \boldsymbol{b} : \boldsymbol{\beta}\,\mathrm{d}x\\ &= LS(\kappa,\mu;\boldsymbol{\tau},v).\end{aligned}$$

Lemma 3.3 results in $C_-(\kappa,\mu)(x_1^2+\cdots+x_6^2) \leqslant f_\kappa(x_1,x_2,x_3) + f_\mu(x_4,x_5,x_6)$.

Step 5. An analogous computation and the upper bound from Lemma 3.3 proves

$$\begin{aligned}LS(\kappa,\mu;\boldsymbol{\tau},v) &\leqslant x_1^2 + x_2^2 + \kappa x_3^2 + 2x_2\min\{x_1,x_3\} + x_4^2 + x_5^2 + \mu x_6^2 + 2x_5\min\{x_4,x_6\}\\ &\leqslant C_+(\kappa,\mu)(x_1^2+\cdots+x_6^2). \quad \square\end{aligned}$$

4. Quasi-optimal adaptive algorithm

4.1. Alternative a posteriori error estimator

For the solution $(\boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}})$ to the discrete equation (11) and for any subset $\mathcal{M} \subseteq \mathcal{T}$ of triangles in the triangulation $\mathcal{T}$ of the 2D polygonal Lipschitz domain Ω, define an a posteriori error estimator $\eta^2(\mathcal{T}, \mathcal{M}) := \sum_{T\in\mathcal{M}} \eta^2(\mathcal{T}, T)$ and $\eta^2(\mathcal{T}) := \eta^2(\mathcal{T}, \mathcal{T})$ with

$$\begin{aligned}\eta^2(\mathcal{T}, T) :=\ & |T| \big\| \operatorname{div}(\mathbf{dev}\boldsymbol{\sigma}_{\mathrm{LS}} - \mathbf{D}u_{\mathrm{LS}}) \big\|^2_{L^2(T)} + |T| \big\| \operatorname{curl} \mathbf{dev}(\boldsymbol{\sigma}_{\mathrm{LS}} - \mathbf{D}u_{\mathrm{LS}}) \big\|^2_{L^2(T)} \\ & + |T|^{1/2} \sum_{E\in\mathcal{E}(T)\cap\mathcal{E}(\Omega)} \big\| [\mathbf{dev}\boldsymbol{\sigma}_{\mathrm{LS}} - \mathbf{D}u_{\mathrm{LS}}]_E\ \nu_E \big\|^2_{L^2(E)} \\ & + |T|^{1/2} \sum_{E\in\mathcal{E}(T)} \big\| [\mathbf{dev}(\boldsymbol{\sigma}_{\mathrm{LS}} - \mathbf{D}u_{\mathrm{LS}})]_E\ \tau_E \big\|^2_{L^2(E)} \end{aligned} \tag{20}$$

with jumps along the edge $E \in \mathcal{E}$ with neighbours T_+ and where available T_- defined, for any discrete tensor $\boldsymbol{\tau}_{\mathrm{NC}} \in P_k(\mathcal{T}; \mathbb{R}^{2\times 2})$, by

$$[\boldsymbol{\tau}_{\mathrm{NC}}]_E := \begin{cases} (\boldsymbol{\tau}_{\mathrm{NC}})\big|_{T_+} - (\boldsymbol{\tau}_{\mathrm{NC}})\big|_{T_-} & \text{on } E \in \mathcal{E}(\Omega) \text{ with } E = \partial T_+ \cap \partial T_-, \\ (\boldsymbol{\tau}_{\mathrm{NC}})\big|_{T_+} & \text{on } E \in \mathcal{E}(\partial\Omega) \cap \mathcal{E}(T_+). \end{cases}$$

Theorem 4.1 (*Efficiency*). *For any admissible triangulation $\mathcal{T} \in \mathbb{T}$ with discrete solutions $(\boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}}) \in \boldsymbol{\Sigma}_{k-1}(\mathcal{T}) \times \mathcal{A}_k(\mathcal{T})$ to (11), it holds $\eta^2(\mathcal{T}) \lesssim LS(f; \boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}})$.*

Proof. The proof of [5, Thm. 3.5] relies on the discrete test function technology due to Verfürth [29] and applies literally to the situation at hand. □

Define the data approximation error $\mu^2(\mathcal{T}) := \sum_{T\in\mathcal{T}} \mu^2(T)$ with oscillations

$$\mu^2(T) := \big\| (1 - \Pi_k) f \big\|^2_{L^2(T)} + |T|^{1/2} \sum_{E\in\mathcal{E}(T)\cap\mathcal{E}(\partial\Omega)} \big\| (1 - \Pi_{k-1}) g' \big\|^2_{L^2(E)} \tag{21}$$

for all $T \in \mathcal{T}$.

4.2. Adaptive algorithm (ALS-FEM)

Input: Initial regular triangulation $\mathcal{T}_0$ with refinement edges of the polygonal domain Ω into triangles and parameters $0 < \theta \leqslant 1, 0 < \rho < 1, 0 < \kappa < \infty$.

for any level $\ell = 0, 1, 2, \ldots$ **do**
Solve LS-FEM with respect to regular triangulation $\mathcal{T}_\ell$ with solution $(\boldsymbol{\sigma}_\ell, u_\ell)$.
Compute $\eta_\ell(T) := \eta(\mathcal{T}_\ell, T)$ from (20) and $\mu(T)$ from (21) for all $T \in \mathcal{T}_\ell$ and set $\eta^2_\ell := \eta^2(\mathcal{T}_\ell)$ and $\mu^2_\ell := \mu^2(\mathcal{T}_\ell)$.
if $\mu^2_\ell \leqslant \kappa\eta^2_\ell$ (**CASE A**) **then**
Select a subset $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ of (almost) minimal cardinality $|\mathcal{M}_\ell|$ with

$$\theta\eta^2_\ell \leqslant \eta^2_\ell(\mathcal{M}_\ell) := \sum_{T\in\mathcal{M}_\ell} \eta^2_\ell(T).$$

Compute smallest regular refinement $\mathcal{T}_{\ell+1}$ of $\mathcal{T}_\ell$ with $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}$ by NVB.
else ($\kappa\eta^2_\ell < \mu^2_\ell$ **CASE B**)
Compute an admissible refinement $\mathcal{T}_{\ell+1}$ of $\mathcal{T}_\ell$ with (almost) minimal cardinality $|\mathcal{T}_{\ell+1}|$ and $\mu(\mathcal{T}_{\ell+1}) \leqslant \rho\mu_\ell$. **fi od**

Output: Sequence of discrete solutions $(\boldsymbol{\sigma}_\ell, u_\ell)_{\ell\in\mathbb{N}_0}$ and triangulations $(\mathcal{T}_\ell)_{\ell\in\mathbb{N}_0}$.

Remark 4.2 (*Case B*). Any (quasi-)optimal data approximation algorithm may be employed in the algorithm ALS-FEM and in the analysis. The Approx algorithm [16,30,7] consists of the thresholding second algorithm (TSA) of [31, Sect. 5] followed by a completion step and is one possible realization of the refinement in Case B.

4.3. Optimal convergence rates

The main result of this paper involves, for any given $0 < s < \infty$, the notion of approximation classes $\mathcal{A}_s$ which consists of all pairs $(u, f) \in \mathcal{A} \times L^2(\Omega; \mathbb{R}^2)$ such that

$$\big|(u, f)\big|^2_{\mathcal{A}_s} := \sup_{N\in\mathbb{N}} (N + 1)^{2s} E(u, f, g, N) < \infty$$

with the best possible error

$$E(u,f,g,N) := \min_{\mathcal{T}\in\mathbb{T}(N)} \min_{(\boldsymbol{\tau}_{\mathrm{LS}},v_{\mathrm{LS}})\in\boldsymbol{\Sigma}_{k-1}(\mathcal{T})\times\mathcal{A}(\mathcal{T})} \big(LS(f;\boldsymbol{\tau}_{\mathrm{LS}},v_{\mathrm{LS}}) + \mathrm{osc}^2(g',\mathcal{E}(\partial\Omega))\big).$$

Theorem 4.3. *There exists a maximal bulk parameter $0<\theta_0<1$ and a maximal separation parameter $0<\kappa_0<\infty$ (which depend exclusively on $\mathcal{T}_0$) such that for all $0<\theta<\theta_0$, for all $0<\kappa<\kappa_0$, for all $0<\rho<1$, and for all $0<s<\infty$, the output $(\boldsymbol{\sigma}_\ell,u_\ell)_{\ell\in\mathbb{N}}$ of ALS-FEM with $(u,f)\in\mathcal{A}_s$ satisfies*

$$\sup_{\ell\in\mathbb{N}} \big(|\mathcal{T}_\ell| - |\mathcal{T}_0| + 1\big)^s \big(LS(f;\boldsymbol{\sigma}_\ell,u_\ell) + \mathrm{osc}^2(g',\mathcal{E}_\ell(\partial\Omega))\big)^{1/2} \approx \big|(u,f)\big|_{\mathcal{A}_s}.$$

The equivalence constants depend only on the initial mesh $\mathcal{T}_0$, the polynomial degree k, and the parameters s, ρ, θ, and κ.

4.4. Axioms of adaptivity

This section summarizes the convergence analysis of [7] based on the axioms (A1)–(A4), (B1)–(B2), and (QM) for the proof of Theorem 4.3. It will be shown in the remaining part of this subsection and Section 5 that those conditions are satisfied for the problem at hand. The axioms (A1)–(A3) and (B2) concern an admissible refinement $\widehat{\mathcal{T}}\in\mathbb{T}$ of an arbitrary triangulation $\mathcal{T}\in\mathbb{T}$ and the associated discrete solutions $(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}},\widehat{u}_{\mathrm{LS}})$ and $(\boldsymbol{\sigma}_{\mathrm{LS}},u_{\mathrm{LS}})$ to (11) in the definition of the distance

$$\delta^2(\mathcal{T},\widehat{\mathcal{T}}) := \big\|\mathrm{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}-\boldsymbol{\sigma}_{\mathrm{LS}})\big\|^2_{L^2(\Omega)} + \big\|\mathbf{dev}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}-\boldsymbol{\sigma}_{\mathrm{LS}}) - \mathbf{D}(\widehat{u}_{\mathrm{LS}}-u_{\mathrm{LS}})\big\|^2_{L^2(\Omega)}. \tag{22}$$

Recall the estimator $\eta^2(\mathcal{T})$ from (20), $\mu^2(\mathcal{T})$ from (21), and their sum convention from the beginning of this section.

Theorem 4.4 (*Stability and Reduction*)**.** *Any admissible refinement $\widehat{\mathcal{T}}$ of $\mathcal{T}\in\mathbb{T}$ satisfies*

$$|\eta(\widehat{\mathcal{T}},\mathcal{T}\cap\widehat{\mathcal{T}}) - \eta(\mathcal{T},\mathcal{T}\cap\widehat{\mathcal{T}})| \leqslant \Lambda_1\delta(\widehat{\mathcal{T}},\mathcal{T}), \tag{A1}$$

$$\eta(\widehat{\mathcal{T}},\widehat{\mathcal{T}}\setminus\mathcal{T}) \leqslant 2^{-1/4}\eta(\mathcal{T},\mathcal{T}\setminus\widehat{\mathcal{T}}) + \Lambda_2\delta(\widehat{\mathcal{T}},\mathcal{T}). \tag{A2}$$

Proof. The proofs of (A1)–(A2) are straight-forward from [3,5,4] even for $k\geqslant 2$. □

The proof of the discrete reliability

$$\delta^2(\widehat{\mathcal{T}},\mathcal{T}) \leqslant \Lambda_3(\eta^2(\mathcal{T},\mathcal{T}\setminus\widehat{\mathcal{T}}) + \mu^2(\mathcal{T})) + \widehat{\Lambda}_3\eta^2(\widehat{\mathcal{T}}) \tag{A3}$$

is postponed to Section 5.

The quasi-orthogonality concerns the outcome $(\mathcal{T}_\ell)_{\ell\in\mathbb{N}_0}$ of the algorithm ALS-FEM and reads

$$\sum_{k=\ell}^{\infty}\delta^2(\mathcal{T}_{k+1},\mathcal{T}_k) \leqslant \Lambda_4(\eta^2(\mathcal{T}_\ell)+\mu^2(\mathcal{T}_\ell)). \tag{A4}$$

This follows directly from (A1)–(A2), the following theorem, and [7, Thm. 3.1].

Theorem 4.5 (*Quasi-Orthogonality with $\varepsilon>0$*)**.** *For any sequence of successive admissible refinements $\mathcal{T}_0,\mathcal{T}_1,\ldots\in\mathbb{T}$ and all positive ε, $\Lambda_4(\varepsilon) := C_{\mathrm{rel}} + C_{\mathrm{qo}}/\varepsilon < \infty$ satisfies*

$$\sum_{k=\ell}^{\ell+m}\big(\delta^2(\mathcal{T}_{k+1},\mathcal{T}_k) - \varepsilon LS(f;\boldsymbol{\sigma}_k,u_k)\big) \leqslant \Lambda_4(\varepsilon)\big(\eta^2(\mathcal{T}_\ell) + \|f-f_\ell\|^2_{L^2(\Omega)}\big).$$

Proof. For all $k=\ell,\ldots,\ell+m$ and positive ε, Proposition 2.8 proves

$$\begin{aligned}\delta^2(\mathcal{T}_{k+1},\mathcal{T}_k) - \varepsilon LS(f;\boldsymbol{\sigma}_{k+1},u_{k+1}) &\leqslant LS(f;\boldsymbol{\sigma}_k,u_k) - LS(f;\boldsymbol{\sigma}_{k+1},u_{k+1})\\ &\quad + C_{\mathrm{qo}}/\varepsilon\big(\mathrm{osc}^2(g',\mathcal{E}_k(\partial\Omega)) - \mathrm{osc}^2(g',\mathcal{E}_{k+1}(\partial\Omega))\big).\end{aligned}$$

The proof concludes with the telescoping sum of the aforementioned estimates over all $k=\ell,\ldots,\ell+m$ and the reliability from Corollary 5.2. □

The subsequent assumptions (B1)–(B2) transfer directly from [7] to the situation at hand in three components for the TSA plus completion (called Approx in) [30].

(B1) Rate s data approximation. $\forall\mathrm{Tol}>0$, $\mathcal{T}_{\mathrm{Tol}} := \textsc{Approx}(\mathrm{Tol},\mu(K):K\in\mathcal{T}_0)\in\mathbb{T}$ satisfies $|\mathcal{T}_{\mathrm{Tol}}|-|\mathcal{T}_0| \leqslant \Lambda_5\mathrm{Tol}^{-1/(2s)}$ and $\mu^2(\mathcal{T}_{\mathrm{Tol}})\leqslant\mathrm{Tol}$.

(B2) Quasi-monotonicity of μ. $\mu(\widehat{\mathcal{T}})\leqslant\Lambda_6\mu(\mathcal{T})$.

Since $\widehat{\Lambda}_3$ in (A3) may be large, the following result is required.

Theorem 4.6 (*Quasi-Monotonicity of $\eta + \mu$*). *Any admissible refinement $\widehat{\mathcal{T}} \in \mathbb{T}$ of $\mathcal{T} \in \mathbb{T}$ satisfies*

$$\eta(\widehat{\mathcal{T}}) + \mu(\widehat{\mathcal{T}}) \leqslant \Lambda_7(\eta(\mathcal{T}) + \mu(\mathcal{T})). \tag{QM}$$

Proof. The efficiency from Theorem 4.1 plus Proposition 2.8 and the reliability from Corollary 5.2 prove (QM). □

5. Discrete reliability

Throughout this section, let $\widehat{\mathcal{T}} \in \mathbb{T}$ denote an admissible refinement of $\mathcal{T} \in \mathbb{T}$ with respective discrete solutions $(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}, \widehat{u}_{\mathrm{LS}})$ and $(\boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}})$ to (11). Recall the definition of the distance $\delta(\widehat{\mathcal{T}}, \mathcal{T})$ from (22).

Theorem 5.1 (*Discrete Reliability*). *It holds*

$$\delta^2(\widehat{\mathcal{T}}, \mathcal{T}) \lesssim \eta^2(\mathcal{T}, \mathcal{T} \setminus \widehat{\mathcal{T}}) + \mu^2(\mathcal{T}) + LS(f; \widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}, \widehat{u}_{\mathrm{LS}}). \tag{23}$$

Proof. The proof follows the lines of [5, Sect. 4.2] for $k = 1$ and utilizes the auxiliary problem of Section 2.3 adapted for all $k \geqslant 1$. Moreover, the generic constants of the stability and approximation estimates of the Scott–Zhang quasi-interpolation [20] as well as the constant C_{ext} from Lemma 2.3 depend on the polynomial degree $k \geqslant 1$. The rest of the proof follows analogously and its key arguments are presented below.

The proof involves three discrete mixed FEM solutions $\widehat{\boldsymbol{\tau}}_{\mathrm{RT}}$, $\widehat{\boldsymbol{\tau}}^*_{\mathrm{RT}}$, and $\boldsymbol{\tau}_{\mathrm{RT}}$ to the Poisson model problem from Section 2.3 with respect to homogeneous boundary conditions $g \equiv 0$, the respective right-hand sides

$$-\operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}), \qquad -\Pi_{k-1}\operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}), \quad \text{and} \quad -\Pi_{k-1}\operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}),$$

and the triangulations $\widehat{\mathcal{T}}$, $\widehat{\mathcal{T}}$, and $\mathcal{T}$; in particular,

$$\operatorname{div}\widehat{\boldsymbol{\tau}}_{\mathrm{RT}} = \operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}) \quad \text{and} \quad \operatorname{div}\widehat{\boldsymbol{\tau}}^*_{\mathrm{RT}} = \Pi_{k-1}\operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}) = \operatorname{div}\boldsymbol{\tau}_{\mathrm{PS}}.$$

The stability of mixed FEM reads

$$\big\|\widehat{\boldsymbol{\tau}}_{\mathrm{RT}} - \widehat{\boldsymbol{\tau}}^*_{\mathrm{RT}}\big\|_{L^2(\Omega)} \lesssim \big\|(1 - \Pi_{k-1})\operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}})\big\|_{L^2(\Omega)} \quad \text{and} \quad \big\|\boldsymbol{\tau}_{\mathrm{PS}}\big\|_{L^2(\Omega)} \lesssim \big\|\Pi_{k-1}\operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}})\big\|_{L^2(\Omega)}. \tag{24}$$

The proof of [5, Lem. 4.5] with Lemma 2.3 and Proposition 2.7 plus elementary algebra shows in this paper that there exist some $\widehat{w}, \widehat{\beta} \in S^k(\widehat{\mathcal{T}}; \mathbb{R}^2)$ with

$$\begin{aligned}
&\widehat{w}|_{\partial\Omega} = (\widehat{I}_k - I_k)g, \qquad \big|\!\big|\!\big|\widehat{w}\big|\!\big|\!\big| \leqslant C_{\mathrm{ext}}\operatorname{osc}((\widehat{I}_k g)', \mathcal{E}(\partial\Omega)), \quad \text{and} \\
&\big\|\operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}})\big\|^2_{L^2(\Omega)} + \big\|\mathbf{dev}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}) - \mathbf{D}(\widehat{u}_{\mathrm{LS}} - u_{\mathrm{LS}})\big\|^2_{L^2(\Omega)} \\
&\quad = \big\|(1 - \Pi)\operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}})\big\|^2_{L^2(\Omega)} + \int_\Omega (\mathbf{dev}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}) - \mathbf{D}(\widehat{u}_{\mathrm{LS}} - u_{\mathrm{LS}})) : (\mathbf{dev}(\widehat{\boldsymbol{\tau}}_{\mathrm{PS}} - \widehat{\boldsymbol{\tau}}^*_{\mathrm{PS}}) - \mathbf{D}\widehat{w})\,\mathrm{d}x \\
&\qquad + \int_\Omega (\mathbf{dev}\boldsymbol{\sigma}_{\mathrm{LS}} - \mathbf{D}u_{\mathrm{LS}}) : \big(\mathbf{D}(\widehat{u}_{\mathrm{LS}} - u_{\mathrm{LS}} - \widehat{w}) - \mathbf{devCurl}\widehat{\beta}\big)\,\mathrm{d}x.
\end{aligned} \tag{25}$$

An analogous proof to [5, Eqn. (23)] using the Cauchy–Schwarz and triangle inequality plus the stability of mixed FEM from (24) and (25) shows

$$\begin{aligned}
&\int_\Omega (\mathbf{dev}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}) - \mathbf{D}(\widehat{u}_{\mathrm{LS}} - u_{\mathrm{LS}})) : \big(\mathbf{dev}(\widehat{\boldsymbol{\tau}}_{\mathrm{PS}} - \widehat{\boldsymbol{\tau}}^*_{\mathrm{PS}}) - \mathbf{D}\widehat{z}\big)\,\mathrm{d}x \\
&\quad \lesssim \big\|\mathbf{dev}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}) - \mathbf{D}(\widehat{u}_{\mathrm{LS}} - u_{\mathrm{LS}})\big\|_{L^2(\Omega)}\big(\big\|(1 - \Pi)\operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}})\big\|_{L^2(\Omega)} + \operatorname{osc}((\widehat{I}_k g)', \mathcal{E}(\partial\Omega))\big).
\end{aligned}$$

An analogous proof to [5, Lem. 4.6] with the discrete equation (11), a piecewise integration by parts, the Cauchy–Schwarz and trace inequality, the stability and approximation property of the Scott–Zhang quasi-interpolation operator [20], and a finite overlap of patches around triangles yields

$$\begin{aligned}
\int_\Omega (\mathbf{dev}\boldsymbol{\sigma}_{\mathrm{LS}} - \mathbf{D}u_{\mathrm{LS}}) : \mathbf{D}(\widehat{u}_{\mathrm{LS}} - u_{\mathrm{LS}} - \widehat{z})\,\mathrm{d}x \lesssim \big|\!\big|\!\big|\widehat{u}_{\mathrm{LS}} - u_{\mathrm{LS}} - \widehat{z}\big|\!\big|\!\big| \Bigg(&\sum_{T \in \mathcal{T} \setminus \widehat{\mathcal{T}}} \Big(|T| \, \big\|\operatorname{div}(\mathbf{dev}\boldsymbol{\sigma}_{\mathrm{LS}} - \mathbf{D}u_{\mathrm{LS}})\big\|^2_{L^2(T)} \\
&+ \sum_{E \in \mathcal{E}(T) \cap \mathcal{E}(\Omega)} |T|^{1/2} \big\|[\mathbf{dev}\boldsymbol{\sigma}_{\mathrm{LS}} - \mathbf{D}u_{\mathrm{LS}}]_E \, \nu_E\big\|^2_{L^2(E)} \Big) \Bigg)^{1/2}.
\end{aligned}$$

An analogous proof to [5, Lem. 4.7] with the discrete equation (11), a piecewise integration by parts, the Cauchy–Schwarz and trace inequality, the stability and approximation property of the Scott–Zhang quasi-interpolation operator [20], a finite overlap of patches around triangles plus the stability of mixed FEM from (24) and Proposition 2.7 yields

$$\int_\Omega (\mathbf{dev}\boldsymbol{\sigma}_{\mathrm{LS}} - \mathbf{D}u_{\mathrm{LS}}) : \mathbf{devCurl}\widehat{\beta}\,\mathrm{d}x \lesssim \big\|\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}\big\|_{H(\mathrm{div},\Omega)} \Bigg(\sum_{T\in\mathcal{T}\setminus\widehat{\mathcal{T}}} \Big(|T| \,\big\|\mathrm{curl}\,\mathbf{dev}\boldsymbol{\sigma}_{\mathrm{LS}}\big\|^2_{L^2(T)} + \sum_{E\in\mathcal{E}(T)} |T|^{1/2} \big\|[\mathbf{dev}(\boldsymbol{\sigma}_{\mathrm{LS}} - \mathbf{D}u_{\mathrm{LS}})]_E\,\tau_E\big\|^2_{L^2(E)}\Big)\Bigg)^{1/2}.$$

The combination of those estimates plus standard rearrangements conclude the proof as in the proof of [5, Thm. 4.3]. □

The discrete reliability and the plain convergence of the LS-FEM implies reliability of the error estimator $\eta(\mathcal{T})$ in the following sense.

Corollary 5.2 (*Reliability*). *Any admissible triangulation* $\mathcal{T} \in \mathbb{T}$ *with discrete solution* $(\boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}}) \in \boldsymbol{\Sigma}_{k-1}(\mathcal{T}) \times \mathcal{A}_k(\mathcal{T})$ *to* (11) *satisfies*

$$LS(f; \boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}}) \leqslant C_{\mathrm{rel}}\big(\eta^2(\mathcal{T}) + \mu^2(\mathcal{T})\big). \tag{26}$$

Proof. The proof of [5, Corollary 4.4] relies on the discrete reliability (23) with $\widehat{\mathcal{T}}$ replaced by successive uniform refinements of $\mathcal{T}$ and applies literally to the situation at hand. The convergence of the LS-FEM in the limit as the maximal mesh-sizes tend to zero proves (26). □

Proof of (A3). The combination of (23) and (26) with respect to $\widehat{\mathcal{T}}$ proves (A3) from Section 4.4. □

6. Numerical experiments

This section presents some numerical experiments with an implementation of the ALS-FEM algorithm from Section 4.2 for the lowest-order case $k = 1$.

6.1. Natural adaptive refinement

The following adaptive algorithm with the natural error indicator is used for comparison.

Input: Initial regular triangulation $\mathcal{T}_0$ with refinement edges of the polygonal domain Ω into triangles and bulk parameter $0 < \theta \leqslant 1$.

for any level $\ell = 0, 1, 2, \ldots$ **do**

Solve LS-FEM with respect to regular triangulation $\mathcal{T}_\ell$ with solution $(\boldsymbol{\sigma}_\ell, u_\ell)$.

Compute $LS(f; \boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}}; T) := \|f + \mathrm{div}\,\boldsymbol{\sigma}_{\mathrm{LS}}\|^2_{L^2(T)} + \|\mathbf{dev}\boldsymbol{\sigma}_{\mathrm{LS}} - \mathbf{D}u_{\mathrm{LS}}\|^2_{L^2(T)}$ for all $T \in \mathcal{T}_\ell$.

Mark a subset $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ of (almost) minimal cardinality $|\mathcal{M}_\ell|$ with

$$\theta\, LS(f; \boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}}) \leqslant LS(f; \boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}}; \mathcal{M}_\ell) := \sum_{T\in\mathcal{M}_\ell} LS(f; \boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}}; T).$$

Compute smallest regular refinement $\mathcal{T}_{\ell+1}$ of $\mathcal{T}_\ell$ with $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}$ by NVB.

Output: Sequence of discrete solutions $(\boldsymbol{\sigma}_\ell, u_\ell)_{\ell\in\mathbb{N}_0}$ and triangulations $(\mathcal{T}_\ell)_{\ell\in\mathbb{N}_0}$.

6.2. Preliminary remarks

The main output of the experiments displayed are the convergence history plots for errors and error estimators as functions of numbers of degrees of freedom (ndof). The grey scale value and the shape of the marker determine the specific terms which are listed in the associated legend. The style of the line and the brightened grey scale value of the marker show whether a uniform refinement (dotted line), the natural adaptive refinement (dashed line) with error indicator $LS(f; \boldsymbol{\sigma}_\ell, u_\ell)$, or the alternative adaptive refinement (solid line) with error indicator $\eta^2(\mathcal{T}_\ell)$ is used for comparison. The letters 'A' and 'B' indicate the case in the adaptive algorithm from Section 4.2 for the computation on the current level.

The degrees of the Gaussian quadrature are chosen in such a way that the integration is exact if possible (for polynomial data) or 10 otherwise. If not stated otherwise, the chosen parameters read $\theta = 0.5$, $\rho = 0.3$, and $\kappa = 0.1$.

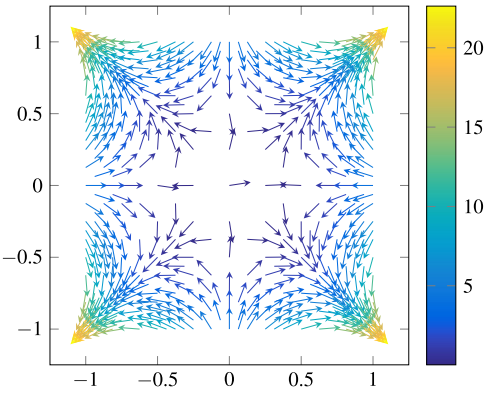

Fig. 1. Velocity plot for the colliding flow example from Section 6.3 for adaptively refined mesh with 1048 triangles (ndof = 4 195).

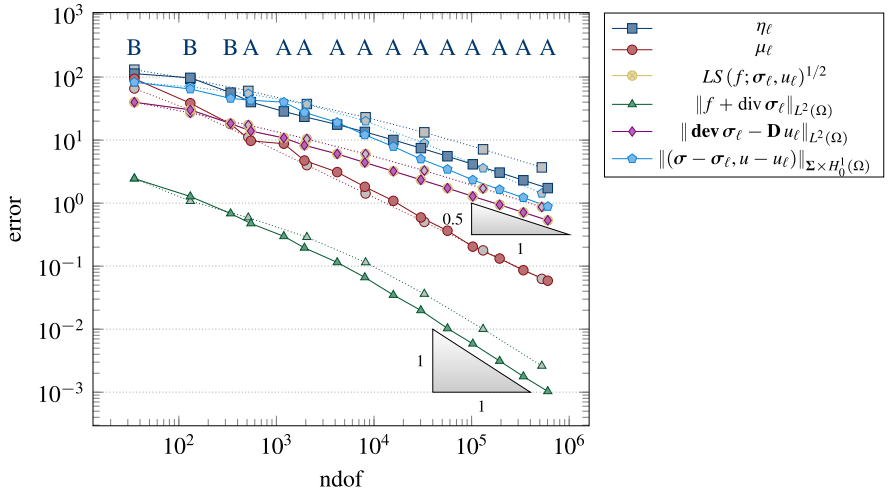

Fig. 2. Convergence history plot for the colliding flow example from Section 6.3 for uniform (dotted lines) and adaptive (solid lines) mesh-refinement.

6.3. Colliding flow example

The first benchmark problem employs $f \equiv 0$ on the square domain $\Omega := (-1, 1)^2$. Let the exact solution

$$u(x) := (20x_1x_2^4 - 4x_1^5, 20x_1^4x_2 - 4x_2^5)^\top \quad \text{and} \quad p(x) := 120x_1^2x_2^2 - 20x_1^4 - 20x_2^4 - 16/3$$

prescribe the inhomogeneous Dirichlet boundary data $g := u|_{\partial\Omega}$. Fig. 1 presents the discrete solution u_ℓ for $\ell = 6$ with ndof = 4 195.

On convex domains, the regularity theory of elliptic operators [32, Ch. 5] yields the optimal convergence rate 0.5 independently from the chosen refinement strategy. The convergence history plot in Fig. 2 confirms this theoretical result for the uniform and adaptive case. The equilibrium residual $\|f + \operatorname{div}\sigma_\ell\|_{L^2(\Omega)}$ in the least-squares functional is of higher order and attains between 1% and 10% of the constitutive residual $\|\mathbf{dev}\sigma_\ell - \mathbf{D}u_\ell\|_{L^2(\Omega)}$.

6.4. Example on L-shaped domain

The second benchmark solves the ALS-FEM with the right-hand side $f \equiv 0$ on the L-shaped domain $\Omega := (-1, 1)^2 \setminus ((0, 1) \times (-1, 0))$. With $\alpha = 856399/1572864$ and $\omega = 3\pi/2$, an exact solution from [33] reads in polar coordinates $(r, \vartheta) \in [0, \infty) \times [0, 3\pi/2]$ as

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

$$p(r,\vartheta) := -r^{\alpha-1}\big((1+\alpha)^2w'(\vartheta) + w'''(\vartheta)\big)\big/(1-\alpha),$$

$$w(\vartheta) := 1/(\alpha+1)\sin((\alpha+1)\vartheta)\cos(\alpha\omega) - \cos((\alpha+1)\vartheta) + 1/(\alpha-1)\sin((\alpha-1)\vartheta)\cos(\alpha\omega) + \cos((\alpha-1)\vartheta).$$

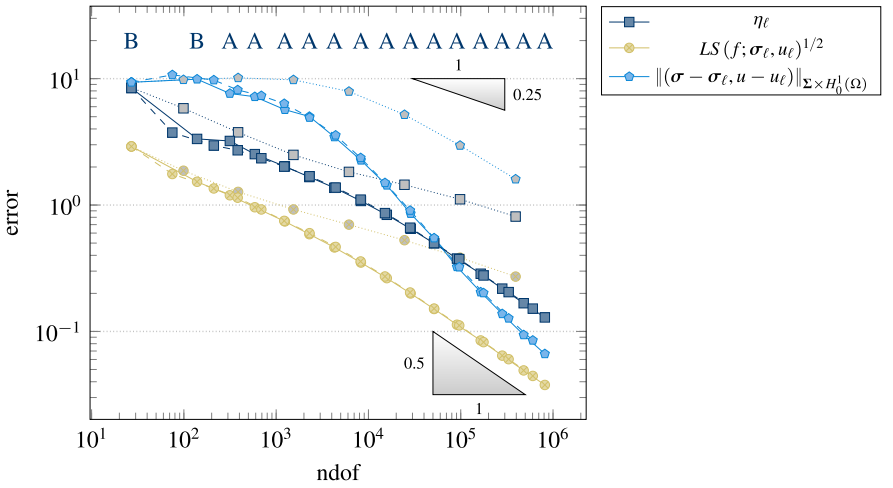

Fig. 3. Convergence history plot for the L-shaped domain example from Section 6.4 for uniform (dotted lines), natural adaptive (dashed lines), and alternative adaptive (solid lines) mesh-refinement.

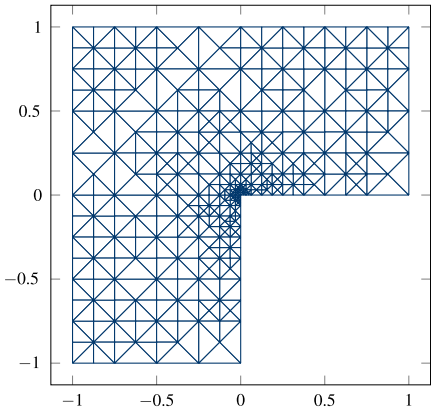

Fig. 4. Triangulation plot for the L-shaped domain example from Section 6.4 of adaptively refined mesh with 572 triangles (ndof = 4275).

Let the exact solution prescribe the inhomogeneous Dirichlet data $g := u|_{\partial\Omega}$. Note that g vanishes at the two edges of the reentrant corner for $\vartheta = 0$ and $\vartheta = 3\pi/2$ and, thus, is smooth on the whole boundary $\partial\Omega$. Therefore, the boundary data oscillations are of higher order also for this example.

The convergence history plot in Fig. 3 confirms the superiority of the adaptive mesh-refinement for the natural error indicator in the spirit of [6] as well as the alternative error indicator from Section 4.1 with the separate marking algorithm from Section 4.2. Both converge with the optimal rate of 0.5 starting at about 1000 degrees of freedom, while uniform refinement shows an empirical convergence rate 0.25.

The plot of the adaptively generated mesh in Fig. 4 shows the expected increased refinement at the re-entrant corner. This avoids unnecessary computational costs by reducing the number of degrees of freedom in the less refined parts of the domain.

6.5. Backward facing step example

The third benchmark concerns a flow in a pipe $\Omega := ((-2, 8) \times (-1, 1)) \setminus ((-2, 0) \times (-1, 0))$ with a bottleneck and a polynomial in- and outflow, for $x \in \partial\Omega$,

$$g(x) := \begin{cases} 1/10\,(-x_2(x_2-1), 0)^\top & \text{for } x_1 = -2, \\ 1/80\,(-(x_2-1)(x_2+1), 0)^\top & \text{for } x_1 = 8, \\ 0 & \text{otherwise.} \end{cases}$$

An exact solution to this problem with $f \equiv 0$ is unknown but the errors are equivalent to the estimator.

Fig. 5 presents the convergence history plot for $k = 1$ and supports the qualitative equivalence of the natural and the alternative mesh-refinement.

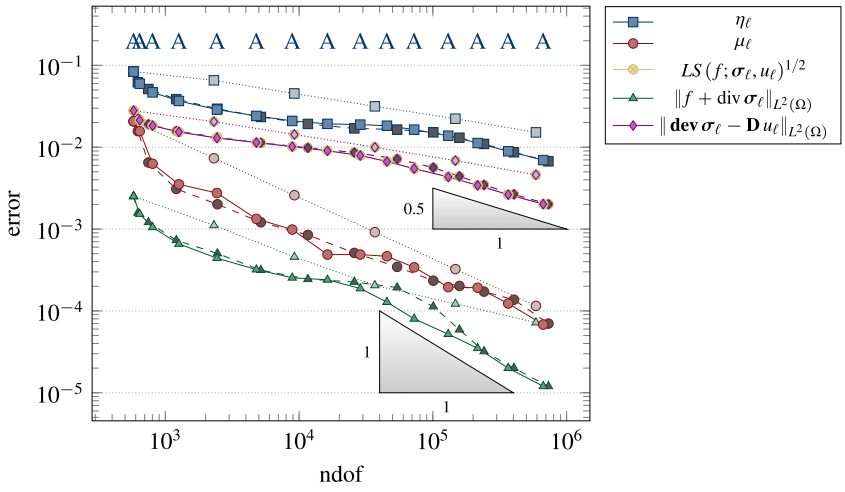

Fig. 5. Convergence history plot for the backward facing step example from Section 6.5 for uniform (dotted lines), natural adaptive (dashed lines), and alternative adaptive (solid lines) mesh-refinement.

6.6. Conclusions

In all examples with a known exact solution, the displayed error estimator by $LS^{1/2}$ appears to be equivalent to the error—even in Section 6.4 we expect that the convergence rates will be equal for finer and finer triangulations.

In all examples, the adaptive algorithm leads to optimal convergence rates $1/2$ for the lowest-order discretization with $k = 1$ even for a relatively large bulk parameter $\theta = 0.5$. This is not theoretically founded but a numerical observation deduced from all the convergence history plots.

The separate marking strategy differs the two cases A and B displayed in the convergence history plots. It appears that this dominates the pre-asymptotic rates only which is expected as for all examples with smooth Dirichlet data. In case g is globally continuous and edgewise in H^2, the oscillations are of higher order and hence may appear negligible for quasi-uniform and fine triangulations. Thus, the separate marking algorithm runs the case B solely in the first few loops.

References

[1] C. Carstensen, D. Gallistl, M. Schedensack, Quasi-optimal adaptive pseudostress approximation of the stokes equations, SIAM J. Numer. Anal. 51 (3) (2013) 1715–1734. http://dx.doi.org/10.1137/110852346.
[2] P.B. Bochev, M.D. Gunzburger, Least-Squares Finite Element Methods, in: Applied Mathematical Sciences, vol. 166, Springer, New York, 2009, http://dx.doi.org/10.1007/b13382.
[3] C. Carstensen, M. Feischl, M. Page, D. Praetorius, Axioms of adaptivity, Comput. Math. Appl. 67 (6) (2014) 1195–1253. http://dx.doi.org/10.1016/j.camwa.2013.12.003.
[4] C. Carstensen, E.-J. Park, Convergence and optimality of adaptive least squares finite element methods, SIAM J. Numer. Anal. 53 (1) (2015) 43–62. http://dx.doi.org/10.1137/130949634.
[5] P. Bringmann, C. Carstensen, An adaptive least-squares FEM for the Stokes equations with optimal convergence rates, Numer. Math. 135 (2) (2017) 459–492. http://dx.doi.org/10.1007/s00211-016-0806-1.
[6] C. Carstensen, E.-J. Park, P. Bringmann, Convergence of natural adaptive least squares finite element methods, (in press) 2017.
[7] C. Carstensen, H. Rabus, Axioms of adaptivity for separate marking, 2016. arXiv:1606.02165.
[8] Z. Cai, B. Lee, P. Wang, Least-squares methods for incompressible Newtonian fluid flow: linear stationary problems, SIAM J. Numer. Anal. 42 (2) (2004) 843–859. http://dx.doi.org/10.1137/S0036142903422673 (electronic).
[9] D.N. Arnold, R.S. Falk, R. Winther, Preconditioning in $H(\mathrm{div})$ and applications, Math. Comp. 66 (219) (1997) 957–984. http://dx.doi.org/10.1090/S0025-5718-97-00826-0.
[10] F. Brezzi, M. Fortin, R. Stenberg, Error analysis of mixed-interpolated elements for Reissner-Mindlin plates, Math. Models Methods Appl. Sci. 1 (2) (1991) 125–151. http://dx.doi.org/10.1142/S0218202591000083.
[11] P. Bochev, M. Gunzburger, On least-squares finite element methods for the Poisson equation and their connection to the Dirichlet and Kelvin principles, SIAM J. Numer. Anal. 43 (1) (2005) 340–362. http://dx.doi.org/10.1137/S003614290443353X.
[12] D.N. Arnold, R.S. Falk, R. Winther, Finite element exterior calculus, homological techniques, and applications, Acta Numer. 15 (2006) 1–155. http://dx.doi.org/10.1017/S0962492906210018.
[13] P. Bringmann, C. Carstensen, G. Starke, An adaptive least-squares FEM for linear elasticity with optimal convergence rates (submitted for publication) 2016.
[14] M. Aurada, M. Feischl, J. Kemetmüller, M. Page, D. Praetorius, Each $H^{1/2}$-stable projection yields convergence and quasi-optimality of adaptive FEM with inhomogeneous Dirichlet data in $\mathbb{R}^d$, ESAIM Math. Model. Numer. Anal. 47 (4) (2013) 1207–1235. http://dx.doi.org/10.1051/m2an/2013069.
[15] W.F. Mitchell, A comparison of adaptive refinement techniques for elliptic problems, ACM Trans. Math. Software 15 (4) (1989) 326–347. http://dx.doi.org/10.1145/76909.76912 (1990).
[16] P. Binev, W. Dahmen, R. DeVore, Adaptive finite element methods with convergence rates, Numer. Math. 97 (2) (2004) 219–268. http://dx.doi.org/10.1007/s00211-003-0492-7.
[17] R. Stevenson, Optimality of a standard adaptive finite element method, Found. Comput. Math. 7 (2) (2007) 245–269. http://dx.doi.org/10.1007/s10208-005-0183-0.
[18] R. Stevenson, The completion of locally refined simplicial partitions created by bisection, Math. Comp. 77 (261) (2008) 227–241. http://dx.doi.org/10.1090/S0025-5718-07-01959-X (electronic).

[19] A.H. Turetskii, The bounding of polynomials prescribed at equally distributed points, Proc. Pedag. Inst. Vitebsk 3 (1940) 117–121 (in Russian).
[20] L.R. Scott, S. Zhang, Finite element interpolation of nonsmooth functions satisfying boundary conditions, Math. Comp. 54 (190) (1990) 483–493. http://dx.doi.org/10.2307/2008497.
[21] C. Carstensen, An a posteriori error estimate for a first-kind integral equation, Math. Comp. 66 (217) (1997) 139–155. http://dx.doi.org/10.1090/S0025-5718-97-00790-4.
[22] P. Grisvard, Elliptic Problems in Nonsmooth Domains, in: Monographs and Studies in Mathematics, vol. 24, Pitman (Advanced Publishing Program), Boston, MA, 1985.
[23] M. Dauge, Smoothness and asymptotics of solutions, in: Elliptic Boundary Value Problems on Corner Domains, in: Lecture Notes in Mathematics, vol. 1341, Springer-Verlag, Berlin, 1988.
[24] S.A. Nazarov, B.A. Plamenevsky, Elliptic Problems in Domains with Piecewise Smooth Boundaries, in: de Gruyter Expositions in Mathematics, vol. 13, Walter de Gruyter & Co., Berlin, 1994, http://dx.doi.org/10.1515/9783110848915.525.
[25] D. Boffi, F. Brezzi, M. Fortin, Mixed Finite Element Methods and Applications, in: Springer Series in Computational Mathematics, vol. 44, Springer, Heidelberg, 2013, http://dx.doi.org/10.1007/978-3-642-36519-5.
[26] D.N. Arnold, L.R. Scott, M. Vogelius, Regular inversion of the divergence operator with Dirichlet boundary conditions on a polygon, Ann. Sc. Norm. Super. Pisa Cl. Sci. (4) 15 (2) (1988) 169–192 (1989).
[27] D. Braess, Theory, fast solvers, and applications in elasticity theory, in: Finite Elements, third ed., Cambridge University Press, Cambridge, 2007, http://dx.doi.org/10.1017/CBO9780511618635.
[28] C. Carstensen, D. Peterseim, H. Rabus, Optimal adaptive nonconforming FEM for the Stokes problem, Numer. Math. 123 (2) (2013) 291–308. http://dx.doi.org/10.1007/s00211-012-0490-8.
[29] R. Verfürth, A posteriori error estimation techniques for finite element methods, in: Numerical Mathematics and Scientific Computation, Oxford University Press, Oxford, 2013, http://dx.doi.org/10.1093/acprof:oso/9780199679423.001.0001.
[30] C. Carstensen, H. Rabus, An optimal adaptive mixed finite element method, Math. Comp. 80 (274) (2011) 649–667. http://dx.doi.org/10.1090/S0025-5718-2010-02397-X.
[31] P. Binev, R. DeVore, Fast computation in adaptive tree approximation, Numer. Math. 97 (2) (2004) 193–217. http://dx.doi.org/10.1007/s00211-003-0493-6.
[32] S.C. Brenner, L.R. Scott, The Mathematical Theory of Finite Element Methods, third ed., in: Texts in Applied Mathematics, vol. 15, Springer, New York, 2008, http://dx.doi.org/10.1007/978-0-387-75934-0.
[33] R. Verfürth, A posteriori error estimators for the Stokes equations, Numer. Math. 55 (3) (1989) 309–325. http://dx.doi.org/10.1007/BF01390056.