Numer. Math.
DOI 10.1007/s00211-016-0806-1

Numerische Mathematik

# An adaptive least-squares FEM for the Stokes equations with optimal convergence rates

**P. Bringmann**[1] · **C. Carstensen**[1]

Received: 7 May 2015 / Revised: 5 January 2016


**Abstract** This paper introduces the first adaptive least-squares finite element method (LS-FEM) for the Stokes equations with optimal convergence rates based on the newest vertex bisection with lowest-order Raviart-Thomas and conforming $P_1$ discrete spaces for the divergence least-squares formulation in 2D. Although the least-squares functional is a reliable and efficient error estimator, the novel refinement indicator stems from an alternative explicit residual-based a posteriori error control with exact solve. Particular interest is on the treatment of the data approximation error which requires a separate marking strategy. The paper proves linear convergence in terms of the levels and optimal convergence rates in terms of the number of unknowns relative to the notion of a non-linear approximation class. It extends and generalizes the approach of Carstensen and Park (SIAM J. Numer. Anal. 53:43–62 2015) from the Poisson model problem to the Stokes equations.

**Mathematics Subject Classification** 65N12 · 65N15 · 65N30 · 65N50 · 65Y20 · 76D07

## 1 Introduction

The universality of the least-squares finite element method (LS-FEM) and its built-in a posteriori error control has enjoyed some ongoing attention over the years; cf. [8] for a general monograph and [1,5] for details on adaptive LS-FEMs. A competitive formulation for the Stokes equations (prototypical in computational fluid dynamics) is the divergence LS-FEM in comparison to the pseudostress mixed finite element

✉ C. Carstensen
cc@math.hu-berlin.de

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

Published online: 12 May 2016

method (PS-FEM) and the non-conforming Crouzeix-Raviart finite element method. The LS-FEM has moderately more degrees of freedom but allows for some immediate a posteriori error estimator even for discrete approximations which do not solve the discrete equations exactly through the computable least-squares functional. Unlike the aforementioned competitors [4,18,21,25], the convergence of an adaptive LS-FEM is an open and not too immediate problem.

From the practical point of view, it appears natural to drive an adaptive mesh-refining with the local contribution from the least-squares functional. From the point of view of the general theory on optimal convergence rates [17], the reduction property is seemingly unavailable simply because the error estimator contributions from the least-squares functional do not involve any mesh-size as a factor that reduces under refinement. It is therefore necessary to base the adaptive mesh-design on some novel a posteriori error terms as it is suggested in [20] for the Poisson model problem with homogeneous Dirichlet boundary conditions. This paper contributes the proof of optimal convergence rates of an adaptive LS-FEM for the Stokes equations in an abstract framework (geared to the four axioms of adaptivity [17] but self-contained) with a detailed analysis of non-homogeneous Dirichlet boundary conditions.

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

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

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

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

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 $LS(0; \boldsymbol{\tau}, v)$ to the natural norm of the underlying function space $\boldsymbol{\Sigma} \times H^1_0(\Omega; \mathbb{R}^2)$ [12, Theorem 4.2] leads to efficiency and reliability of the a posteriori error estimator $LS(f; \boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}})$ for some discrete minimizer $(\boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}})$. Since the contributions to the estimator do not contain any powers of the mesh-size, the known arguments for the proof of the estimator reduction do not apply to the situation at hand; cf. [17] for a state-of-the-art survey on the convergence of adaptive finite element methods. A major contribution of this paper is the statement of an equivalent a posteriori error estimator $\eta$ in Sect. 3.1 with the volume contributions

$$|T| \|\operatorname{div} \operatorname{dev} \boldsymbol{\sigma}_{\mathrm{LS}}\|_{L^2(T)} + |T| \|\operatorname{curl} \operatorname{dev} \boldsymbol{\sigma}_{\mathrm{LS}}\|_{L^2(T)}$$

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

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

for an interior edge $E$ plus terms on the boundary which include Dirichlet data oscillations.

It satisfies the axioms of adaptivity, namely stability, reduction, and discrete reliability, as proven in Sect. 4. The discrete reliability, however, includes some additional term $\|f-\Pi f\|_{L^2(\Omega)}$, which requires the reduction of the data approximation error with the piecewise constant $L^2$ best-approximation $\Pi f$ of $f$ by some separate marking strategy in the adaptive algorithm [22]. The main loop on the level $\ell$ with some regular triangulation $\mathcal{T}_\ell$ in the adaptive LS-FEM with separate marking computes the discrete solution $(\boldsymbol{\sigma}_\ell, u_\ell)$ and the estimator $\eta_\ell$ and reads (for parameters $\kappa, \rho, \theta$) as follows.

**(ALS-FEM)** In **Case A** $\|f-f_\ell\|_{L^2(\Omega)}\le\kappa\eta_\ell$ with $f_\ell:=\Pi_\ell f$, compute $\mathcal{T}_{\ell+1}$ with Dörfler marking for $\eta_\ell(T)$ and newest-vertex bisection (NVB).
In **Case B** (i.e. $\|f-f_\ell\|_{L^2(\Omega)}>\kappa\eta_\ell$), compute optimal approximation $f_{\ell+1}$ of $f$ by refinement $\mathcal{T}_{\ell+1}$ of $\mathcal{T}_\ell$ with

$$\left\|f-f_{\ell+1}\right\|_{L^2(\Omega)}\le\rho\left\|f-f_\ell\right\|_{L^2(\Omega)}.$$

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

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

with the non-linear approximation class

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

and the best possible error

$$E(u,f,N):=\min_{\mathcal{T}\in\mathbb{T}(N)}\min_{(\boldsymbol{\tau}_{\mathrm{LS}},v_{\mathrm{LS}})\in\boldsymbol{\Sigma}(\mathcal{T})\times S^1(\mathcal{T};\mathbb{R}^2)}\left(LS(f;\boldsymbol{\tau}_{\mathrm{LS}},v_{\mathrm{LS}})+\operatorname{osc}^2(g',\mathcal{E}(\Gamma))\right).$$

The proofs require an adopted Helmholtz decomposition [21] for piecewise constant matrix-valued functions and, thus, the analysis is restricted to the lowest-order case. Moreover, this paper establishes a medius analysis of the LS-FEM as well as a novel reliable and efficient a posteriori error control thereof. The pseudostress method [12,14,19] serves as a related mixed discretization and allows the discrete reliability analysis.

The paper is organized as follows: Sect. 2 introduces the notation employed for triangulations, finite element function spaces, and the approximation of the Dirichlet boundary data. It recalls the involved PS-FEM and LS-FEM and concludes with a medius analysis of the LS-FEM, a discrete Helmholtz decomposition, and the tr-dev-div lemma. Section 3 defines a reliable and efficient alternative a posteriori error estimator and presents the associated adaptive algorithm with separate marking. Section 4 covers the proof of the four axioms of adaptivity and concludes with the proof of the main result.

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

To keep the notation and technical overhead minimal and this first paper on ALS-FEM for the Stokes equations short, this paper is restricted to the 2D case although most of the arguments carry over to 3D as well. However, the remaining modifications for 3D concern the discrete Helmholtz decomposition in 2D, which can be circumvented with the observation, that the divergence-free Raviart-Thomas function is the curl of a Nédélec edge-element function on some fine level which is approximated on a coarse level plus a discrete regular split as in [30]. The modification of the Dirichlet data approximation may follow the paper [2] for 3D.

## 2 Notation and preliminaries

### 2.1 Standard notation

Let tr and dev denote the trace operator and the deviatoric part of a matrix $M \in \mathbb{R}^{2\times 2}$, i.e.,

$$\mathrm{tr}\, M := M_{11} + M_{22} \quad \text{and} \quad \mathrm{dev}\, M := M - \mathrm{tr}(M)/2\; I_{2\times 2}.$$

Define $\mathbb{R}^{2\times 2}_{\mathrm{dev}}$ as the space of trace-free $2 \times 2$ matrices. For $M, N \in \mathbb{R}^{2\times 2}$, $M : N := \mathrm{tr}(M^\top N)$ abbreviates the Euclidian scalar product in $\mathbb{R}^{2\times 2}$.

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

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

### 2.2 Triangulations and finite element function spaces

Given an initial shape-regular triangulation $\mathcal{T}_0$ into triangles of the polygonal Lipschitz domain $\Omega$ with some initial condition on the refinement edges, the set of admissible triangulations is defined as

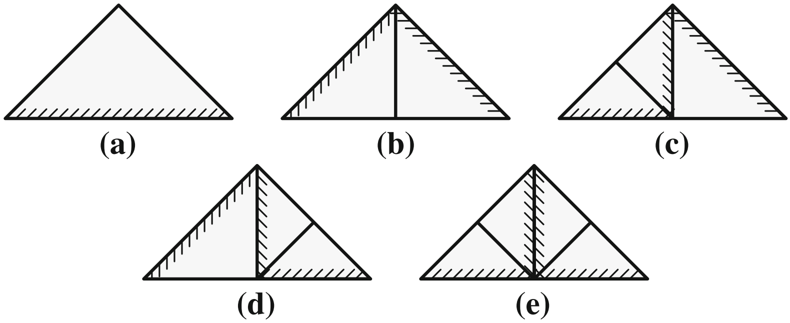


**Fig. 1** One-level refinements of a triangle $K$ in the NVB with refinement edge ⁄⁄⁄⁄⁄⁄⁄. (a) Triangle $K$, (b) green, (c) blue-left, (d) blue-right, (e) bisec3

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

For any natural number $N \in \mathbb{N}$, set

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

All triangulations in this paper are admissible, when generated with NVB as depicted in Fig. 1. This implies shape-regularity of all $\mathcal{T} \in \mathbb{T}$ in the sense that only a finite number of angles appear in $\bigcup \mathbb{T}$. The reader is referred to [6,27] for details on mesh-refining.

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

For any triangulation $\mathcal{T} \in \mathbb{T}$, $\mathcal{N}$ denotes the set of nodes and $\mathcal{E}$ the set of edges and the corresponding sets $\mathcal{N}(\Gamma)$ and $\mathcal{E}(\Gamma)$ on the boundary $\Gamma$, $\mathcal{N}(\Omega)$ and $\mathcal{E}(\Omega)$ in the interior $\Omega$. 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. For the node $z \in \mathcal{N}$ and the edge $E \in \mathcal{E}$, define $\omega_z \subseteq \Omega$ and $\omega_E \subseteq \Omega$ by

$$\omega_z := \operatorname{int}\left(\bigcup_{T \in \mathcal{T}, z \in \mathcal{N}(T)} T\right) \quad \text{and} \quad \omega_E := \operatorname{int}\left(\bigcup_{T \in \mathcal{T}, E \subseteq T} T\right).$$

Let $P_k(\mathcal{T})$ and $P_k(\mathcal{T}; \mathbb{R}^2)$ (resp. $P_k(\mathcal{T}; \mathbb{R}^{2\times 2})$) denote the space of piecewise polynomials of degree at most $k \in \mathbb{N}_0$ for vector-valued (resp. matrix-valued) functions. Let the piecewise averages $f_\mathcal{T} := \Pi f \in P_0(\mathcal{T})$ be the orthogonal projection of an $L^2$-function $f$ onto $P_0(\mathcal{T})$ and analogously for every component of vector-valued or matrix-valued functions. The oscillations

$$\operatorname{osc}(f, \mathcal{T}) := \left\| h_\mathcal{T}(f - f_\mathcal{T}) \right\|_{L^2(\Omega)}$$

of $f$ on the triangulation $\mathcal{T}$ are weighted with the piecewise constant mesh-size function $h_{\mathcal{T}} \in P_0(\mathcal{T})$ defined by $h_{\mathcal{T}}|_T := h_T := \operatorname{diam}(T)$ for $T \in \mathcal{T}$.

The Courant finite element function spaces read

$$S^1(\mathcal{T};\mathbb{R}^2) := P_1(\mathcal{T};\mathbb{R}^2) \cap C(\overline{\Omega};\mathbb{R}^2) \subseteq H^1(\Omega;\mathbb{R}^2),$$
$$S^1_0(\mathcal{T};\mathbb{R}^2) := S^1(\mathcal{T};\mathbb{R}^2) \cap H^1_0(\mathcal{T};\mathbb{R}^2) \subseteq V := H^1_0(\Omega;\mathbb{R}^2).$$

The discrete approximation of row-wise $H(\operatorname{div},\Omega)$-functions in $\boldsymbol{\Sigma} := \{\boldsymbol{\tau} \in H(\operatorname{div},\Omega;\mathbb{R}^{2\times 2}) : \operatorname{tr}\boldsymbol{\tau} \in L^2_0(\Omega)\}$ employs the space of row-wise Raviart-Thomas functions [9–11]

$$RT_0(\mathcal{T}) := \left\{q_{\mathrm{RT}} \in H(\operatorname{div},\Omega) : \forall T \in \mathcal{T}\ \exists a,b,c \in \mathbb{R},\ q_{\mathrm{RT}}\big|_T = (a,b) + cx^\top\right\},$$
$$\boldsymbol{\Sigma}(\mathcal{T}) := \left\{\boldsymbol{\tau}_{\mathrm{RT}} = (\boldsymbol{\tau}_{jk})_{j,k=1,2} \in \boldsymbol{\Sigma} : \forall j = 1,2,\ (\boldsymbol{\tau}_{j1},\boldsymbol{\tau}_{j2}) \in RT_0(\mathcal{T})\right\}.$$

### 2.3 Approximation of Dirichlet boundary data

Given some initial triangulation $\mathcal{T}$, let $H^1(\mathcal{E}(\Gamma);\mathbb{R}^2)$ consist of all boundary functions $g \in L^2(\Gamma;\mathbb{R}^2)$ with square-integrable arc-length derivative $g' = \partial g/\partial s \in L^2(\Gamma;\mathbb{R}^2)$ along the boundary edges $\mathcal{E}(\Gamma)$. Let $P_k(\mathcal{E}(\Gamma))$ denote the space of piecewise polynomials of degree at most $k \in \mathbb{N}_0$ on the boundary. For any function $g \in H^1(\mathcal{E}(\Gamma);\mathbb{R}^2) \cap C(\Gamma;\mathbb{R}^2)$, let $Ig \in S^1(\mathcal{E}(\Gamma);\mathbb{R}^2) := P_1(\mathcal{E}(\Gamma);\mathbb{R}^2) \cap C(\Gamma;\mathbb{R}^2)$ denote the nodal interpolation defined by linear interpolation of the nodal values, for $z \in \mathcal{N}(\Gamma)$, $(Ig)(z) := g(z)$. Let $\Pi g'$ denote the $L^2(\Gamma)$-orthogonal projection of $g'$ onto $P_0(\mathcal{E}(\Gamma);\mathbb{R}^2)$ and $h_{\mathcal{E}} \in P_0(\mathcal{E})$ the piecewise constant function with $h_{\mathcal{E}}\big|_E \equiv \operatorname{diam}(\omega_E)$ for every $E \in \mathcal{E}$ to define the Dirichlet data oscillation

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

(Cf. [2,3,24] for details on the approximation of Dirichlet boundary data.)

**Lemma 2.1** *Given any boundary data $g \in H^1(\Gamma;\mathbb{R}^2)$, there exists some extension $w \in H^1(\Omega;\mathbb{R}^2)$ with*

$$w\big|_\Gamma = (1-I)g \quad \textit{and} \quad \|\!|w|\!\| \lesssim \operatorname{osc}(g',\mathcal{E}(\Gamma)).$$

*If $\widehat{g} \in S^1(\widehat{\mathcal{E}}(\Gamma);\mathbb{R}^2)$ for any admissible refinement $\widehat{\mathcal{T}}$ of $\mathcal{T}$, this even holds for some discrete extension $\widehat{w} \in S^1(\widehat{\mathcal{T}};\mathbb{R}^2)$ in that*

$$\widehat{w}\big|_\Gamma = (1-I)\widehat{g} \quad \textit{and} \quad \|\!|\widehat{w}|\!\| \lesssim \operatorname{osc}(\widehat{g}',\mathcal{E}(\Gamma)).$$

*Proof Step 1:* Set $y := (1-I)g$ and let $w \in H^1(\Omega;\mathbb{R}^2)$ solve the Dirichlet problem

$$-\Delta w + w = 0 \quad \text{in } \Omega \quad \text{and} \quad w = y \quad \text{on } \Gamma. \tag{3}$$

The weak solution $w$ solves the minimization problem

$$\|y\|_{H^{1/2}(\Gamma)} := \min_{Y\in H^1(\Omega;\mathbb{R}^2), Y|_\Gamma = y} \|Y\|_{H^1(\Omega)} = \|w\|_{H^1(\Omega)}.$$

Since $y$ vanishes in $\mathcal{N}(\Gamma)$ and the triangulation $\mathcal{T}$ is shape-regular, [16, Theorem 1] implies

$$\|y\|_{H^{1/2}(\Gamma)} \lesssim \|h_{\mathcal{E}}^{1/2} y'\|_{L^2(\Gamma)}.$$

Notice that various definitions of the $H^{1/2}$-norm in $H^{1/2}(\Gamma)$ are equivalent and the universal equivalent constants solely depend on $\Omega$. The combination of the last two displayed formulas and the definition of $y \equiv (1-I)g$ prove

$$|||w||| \le \|w\|_{H^1(\Omega)} \lesssim \|h_{\mathcal{E}}^{1/2}\partial\big((1-I)g\big)/\partial s\|_{L^2(\Gamma)} = \operatorname{osc}(g', \mathcal{E}(\Gamma)). \tag{4}$$

*Step 2:* If $\widehat{g} \in S^1(\widehat{\mathcal{E}}(\Gamma);\mathbb{R}^2)$ for some admissible refinement $\widehat{\mathcal{T}}$ of $\mathcal{T}$, Step 1 leads to $w \in H^1(\Omega;\mathbb{R}^2)$ with (4). The Scott-Zhang quasi-interpolation [26], which is carefully defined in [2] with respect to the edges on the boundary, leads to $\widehat{w} := \widehat{J}w$ in $S^1(\widehat{\mathcal{T}};\mathbb{R}^2)$ with $\widehat{w} = (1-I)\widehat{g}$ on $\Gamma$. It is known [26, Theorem 3.1] that this quasi-interpolation operator is $H^1$-stable in the sense that $|||\widehat{w}||| \lesssim |||w|||$. The combination with (4) leads to $|||\widehat{w}||| \lesssim \operatorname{osc}(\widehat{g}', \mathcal{E}(\Gamma))$ and concludes the proof. □

Along the polygonal one-dimensional boundary $\Gamma$, the nodal interpolation $Ig$ of $g$ allows the following well-known orthogonality of the arc-length derivative $\partial \bullet /\partial s$ which is stated and proved here for convenient reading.

**Lemma 2.2** *Any admissible refinement $\widehat{\mathcal{T}}$ of $\mathcal{T}$ with corresponding approximations $\widehat{I}g$ and $Ig$ of the boundary data satisfies, for every $E \in \mathcal{E}(\Gamma)$, that*

$$\int_E (\widehat{\Pi} - \Pi)g' \cdot (1-\widehat{\Pi})g' \,\mathrm{d}s = 0. \tag{5}$$

*In particular, this implies*

$$\operatorname{osc}^2(\widehat{\Pi}g', \mathcal{E}(\Gamma)) + \operatorname{osc}^2(g', \widehat{\mathcal{E}}(\Gamma)) \le \operatorname{osc}^2(g', \mathcal{E}(\Gamma)). \tag{6}$$

*Proof* The assertion (5) is the orthogonality of the operator $\widehat{\Pi}$ and the conformity of the finite element spaces. The fundamental theorem of calculus on $E = \operatorname{conv}\{A, B\} \in \mathcal{E}(\Gamma)$ and nodal interpolation of $A, B \in \mathcal{N}(\Gamma)$ shows that

$$\int_E \partial\big((1-I)g\big)/\partial s \,\mathrm{d}s = \big((1-I)g\big)(B) - \big((1-I)g\big)(A) = 0.$$

This proves $\Pi g' = \partial(Ig)/\partial s$. This, the estimate $h_{\widehat{\mathcal{E}}} \le h_{\mathcal{E}}$ a.e., and the Pythagoras theorem imply (6). □

**Corollary 2.3** *Any sequence of successive refinements* $\mathcal{T}_\ell, \dots, \mathcal{T}_{\ell+m+1} \in \mathbb{T}$ *with corresponding approximations* $g_\ell, \dots, g_{\ell+m+1}$ *satisfies*

$$\sum_{j=\ell}^{\ell+m} \operatorname{osc}^2(g'_{j+1}, \mathcal{E}_j(\Gamma)) \le \operatorname{osc}^2(g'_{\ell+m+1}, \mathcal{E}_\ell(\Gamma)).$$

*Proof* This follows from Lemma 2.2. Since $g'_{j+1} - g'_j$ is orthogonal to $g'_{k+1} - g'_k$ in $L^2(\Gamma; \mathbb{R}^2)$ for all $\ell \le j < k$, the Pythagoras theorem leads to

$$\begin{aligned}\sum_{j=\ell}^{\ell+m} \operatorname{osc}^2(g'_{j+1}, \mathcal{E}_j(\Gamma)) &\le \Big\| h_\ell^{1/2} \sum_{j=\ell}^{\ell+m} (g'_{j+1} - g'_j) \Big\|_{L^2(\Gamma)}^2 \\ &= \| h_\ell^{1/2} (g'_{\ell+m+1} - g'_\ell) \|_{L^2(\Gamma)}^2 = \operatorname{osc}^2(g'_{\ell+m+1}, \mathcal{E}_\ell(\Gamma)).\end{aligned}$$

□

### 2.4 Pseudostress approximation

Given some right-hand side $f \in L^2(\Omega; \mathbb{R}^2)$ and Dirichlet boundary data $g \in H^1(\Gamma; \mathbb{R}^2)$ with $\int_\Gamma g \cdot \nu \, \mathrm{d}s = 0$, the weak formulation of (1) seeks $(\boldsymbol{\sigma}, u) \in \boldsymbol{\Sigma} \times L^2(\Omega; \mathbb{R}^2)$ such that, for all $(\boldsymbol{\tau}, v) \in \boldsymbol{\Sigma} \times L^2(\Omega; \mathbb{R}^2)$,

$$\begin{aligned}\int_\Omega \boldsymbol{\sigma} : \operatorname{dev} \boldsymbol{\tau} \, \mathrm{d}x + \int_\Omega u \cdot \operatorname{div} \boldsymbol{\tau} \, \mathrm{d}x &= \langle g, \boldsymbol{\tau} \nu \rangle_\Gamma, \\ \int_\Omega v \cdot \operatorname{div} \boldsymbol{\sigma} \, \mathrm{d}x &= -\int_\Omega f \cdot v \, \mathrm{d}x.\end{aligned} \tag{7}$$

The PS-FEM seeks $(\boldsymbol{\sigma}_{\mathrm{PS}}, u_{\mathrm{PS}}) \in \boldsymbol{\Sigma}(\mathcal{T}) \times P_0(\mathcal{T}; \mathbb{R}^2)$ such that, for all $(\boldsymbol{\tau}_{\mathrm{PS}}, v_{\mathrm{PS}}) \in \boldsymbol{\Sigma}(\mathcal{T}) \times P_0(\mathcal{T}; \mathbb{R}^2)$,

$$\begin{aligned}\int_\Omega \boldsymbol{\sigma}_{\mathrm{PS}} : \operatorname{dev} \boldsymbol{\tau}_{\mathrm{PS}} \mathrm{d}x + \int_\Omega u_{\mathrm{PS}} \cdot \operatorname{div} \boldsymbol{\tau}_{\mathrm{PS}} \mathrm{d}x &= \langle g, \boldsymbol{\tau}_{\mathrm{PS}} \nu \rangle_\Gamma, \\ \int_\Omega v_{\mathrm{PS}} \cdot \operatorname{div} \boldsymbol{\sigma}_{\mathrm{PS}} \mathrm{d}x &= -\int_\Omega f \cdot v_{\mathrm{PS}} \mathrm{d}x.\end{aligned} \tag{8}$$

The papers [12,14,18,19] outline a detailed analysis of this first-order method.

### 2.5 Least-squares FEM

The LS-FEM approximates the system (1) by minimizing the residual functional $LS(f; \bullet)$ defined, for any $(\boldsymbol{\tau}, v) \in \boldsymbol{\Sigma} \times H^1(\Omega; \mathbb{R}^2)$, by

$$LS(f; \boldsymbol{\tau}, v) := \| f + \operatorname{div} \boldsymbol{\tau} \|_{L^2(\Omega)}^2 + \| \operatorname{dev} \boldsymbol{\tau} - \mathrm{D}\, v \|_{L^2(\Omega)}^2.$$

The associated bilinear form $\mathcal{B} : (\boldsymbol{\Sigma} \times H^1(\Omega;\mathbb{R}^2)) \times (\boldsymbol{\Sigma} \times H^1(\Omega;\mathbb{R}^2)) \to \mathbb{R}$ of the least-squares functional $LS$ and the linear functional $F : \boldsymbol{\Sigma} \to \mathbb{R}$ read, for $\boldsymbol{\sigma}, \boldsymbol{\tau} \in \boldsymbol{\Sigma}$ and $u, v \in H^1(\Omega;\mathbb{R}^2)$,

$$\mathcal{B}(\boldsymbol{\sigma}, u; \boldsymbol{\tau}, v) := \int_\Omega \operatorname{div}\boldsymbol{\sigma} : \operatorname{div}\boldsymbol{\tau}\,\mathrm{d}x + \int_\Omega (\operatorname{dev}\boldsymbol{\sigma} - \mathrm{D}\,u) : (\operatorname{dev}\boldsymbol{\tau} - \mathrm{D}\,v)\mathrm{d}x,$$
$$F(\boldsymbol{\tau}) := -\int_\Omega f \cdot \operatorname{div}\boldsymbol{\tau}\,\mathrm{d}x.$$

The Euler-Lagrange equations for the minimization of $LS(f;\bullet)$ lead to the weak problem: Seek $(\boldsymbol{\sigma}, u) \in \boldsymbol{\Sigma} \times \mathcal{A}$ such that, for all $(\boldsymbol{\tau}, v) \in \boldsymbol{\Sigma} \times V$,

$$\mathcal{B}(\boldsymbol{\sigma}, u; \boldsymbol{\tau}, v) = F(\boldsymbol{\tau}). \tag{9}$$

The well-established equivalence [12, Theorem 4.2] of the natural norm in $\boldsymbol{\Sigma} \times V$ with the homogeneous least-squares functional reads

$$\mathcal{B}(\boldsymbol{\tau}, v; \boldsymbol{\tau}, v) = LS(0; \boldsymbol{\tau}, v) \approx \|\boldsymbol{\tau}\|^2_{H(\operatorname{div},\Omega)} + |||v|||^2 \quad \text{for all } (\boldsymbol{\tau}, v) \in \boldsymbol{\Sigma} \times V. \tag{10}$$

This leads to the uniqueness of solutions $(\boldsymbol{\sigma}, u) \in \boldsymbol{\Sigma} \times \mathcal{A}$ to (9) with arbitrary Dirichlet boundary data. The existence of a solution follows from the standard existence proof for the Stokes equations and the Ladyzhenskaya lemma.

**Lemma 2.4** *For* $(\boldsymbol{\tau}, v) \in \boldsymbol{\Sigma} \times H^1(\Omega;\mathbb{R}^2)$*, any extension* $z \in v + V \subseteq H^1(\Omega;\mathbb{R}^2)$ *of the boundary data* $v|_\Gamma$ *satisfies*

$$LS(0; \boldsymbol{\tau}, v) + |||z|||^2 \approx \|\boldsymbol{\tau}\|^2_{H(\operatorname{div},\Omega)} + |||v|||^2 + |||z|||^2.$$

*Proof* This follows from elementary calculations with the Cauchy-Schwarz and the Young inequality. □

Recall the set $\mathcal{A} := \{v \in H^1(\Omega;\mathbb{R}^2) : v = g \text{ on } \Gamma\}$ of admissible displacements and the nodal interpolation $Ig$ from Sect. 2.3 and define the space of discrete admissible velocity functions

$$\mathcal{A}(\mathcal{T}) := \{v \in S^1(\mathcal{T};\mathbb{R}^2) : v = Ig \text{ on } \Gamma\}$$

on a regular triangulation $\mathcal{T}$ of $\Omega$. A conforming discretization seeks $(\boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}}) \in \boldsymbol{\Sigma}(\mathcal{T}) \times \mathcal{A}(\mathcal{T})$ such that, for all $(\boldsymbol{\tau}_{\mathrm{LS}}, v_{\mathrm{LS}}) \in \boldsymbol{\Sigma}(\mathcal{T}) \times S^1_0(\mathcal{T};\mathbb{R}^2)$,

$$\mathcal{B}(\boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}}; \boldsymbol{\tau}_{\mathrm{LS}}, v_{\mathrm{LS}}) = F(\boldsymbol{\tau}_{\mathrm{LS}}) = -\int_\Omega f_{\mathcal{T}} \cdot \operatorname{div}\boldsymbol{\tau}_{\mathrm{LS}}\mathrm{d}x. \tag{11}$$

The equivalence (10) proves that $\|\bullet\|_{\mathcal{B}} := \mathcal{B}(\bullet,\bullet)^{1/2}$ is an equivalent norm on $\boldsymbol{\Sigma} \times V$. However, the expression

$$\mathcal{B}(\boldsymbol{\tau}, v; \boldsymbol{\tau}, v) = \|\operatorname{div}\boldsymbol{\tau}\|^2_{L^2(\Omega)} + \|\operatorname{dev}\boldsymbol{\tau} - \mathrm{D}\,v\|^2_{L^2(\Omega)}$$

is non-negative for all $\boldsymbol{\tau} \in \boldsymbol{\Sigma}$ and $v \in H^1(\Omega; \mathbb{R}^2)$. This enables the subsequent definition of $\delta(\widehat{\mathcal{T}}, \mathcal{T})$.

**Definition 2.5** Given any admissible refinement $\widehat{\mathcal{T}} \in \mathbb{T}$ of an admissible triangulation $\mathcal{T} \in \mathbb{T}$, let $\widehat{g} := \widehat{I} g$ be the nodal interpolation of the boundary data $g$ and let $(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}, \widehat{u}_{\mathrm{LS}})$ and $(\boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}})$ solve the discrete equation (11) with respect to $\widehat{\mathcal{T}}$ and $\mathcal{T}$, respectively. Define

$$\delta^2(\widehat{\mathcal{T}}, \mathcal{T}) := \left\|(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}, \widehat{u}_{\mathrm{LS}} - u_{\mathrm{LS}})\right\|_{\mathcal{B}}^2 + \mathrm{osc}^2(\widehat{g}', \mathcal{E}(\Gamma)).$$

### 2.6 Medius analysis of LS-FEM

Let $(\boldsymbol{\sigma}, u) \in \boldsymbol{\Sigma} \times L^2(\Omega; \mathbb{R}^2)$ be the exact solution to the continuous pseudostress equation (7) with right-hand side $f$ and Dirichlet boundary data $g$. Let $(\boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}}) \in \boldsymbol{\Sigma}(\mathcal{T}) \times \mathcal{A}(\mathcal{T})$ denote the discrete solution to (11) and let $Gu \in \mathcal{A}(\mathcal{T})$ be the Galerkin projection $G$ of $u$ onto $\mathcal{A}(\mathcal{T})$ with

$$|||u - Gu||| = \min_{v_{\mathrm{C}} \in \mathcal{A}(\mathcal{T})} |||u - v_{\mathrm{C}}|||.$$

Let $\Pi_{\mathrm{RT}}$ denote the $L^2$ projection of $\boldsymbol{\sigma}$ onto $\boldsymbol{\Sigma}(\mathcal{T})$, i.e., $\Pi_{\mathrm{RT}}\boldsymbol{\sigma} \in \boldsymbol{\Sigma}(\mathcal{T})$ with

$$\left\|\boldsymbol{\sigma} - \Pi_{\mathrm{RT}}\boldsymbol{\sigma}\right\|_{L^2(\Omega)} = \min_{\boldsymbol{\tau}_{\mathrm{RT}} \in \boldsymbol{\Sigma}(\mathcal{T})} \left\|\boldsymbol{\sigma} - \boldsymbol{\tau}_{\mathrm{RT}}\right\|_{L^2(\Omega)}.$$

**Theorem 2.6** *It holds that*

$$\begin{aligned} &LS(f; \boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}}) + \mathrm{osc}^2(g', \mathcal{E}(\Gamma)) \\ &\quad \approx \left\|\boldsymbol{\sigma} - \Pi_{\mathrm{RT}}\boldsymbol{\sigma}\right\|_{L^2(\Omega)}^2 + |||u - Gu|||^2 + \mathrm{osc}^2(g', \mathcal{E}(\Gamma)) + \left\|f - f_{\mathcal{T}}\right\|_{L^2(\Omega)}^2. \end{aligned}$$

*Proof* The proof of the estimate "$\lesssim$" starts with the $L^2(\Omega; \mathbb{R}^2)$-orthogonality $f - f_{\mathcal{T}} \perp P_0(\mathcal{T}; \mathbb{R}^2)$ and the Pythagoras theorem

$$LS(f; \boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}}) = \left\|f - f_{\mathcal{T}}\right\|_{L^2(\Omega)}^2 + LS(f_{\mathcal{T}}; \boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}}).$$

Since $(\boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}})$ is a discrete minimizer of $LS(f_{\mathcal{T}}; \bullet)$,

$$LS(f; \boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}}) \leq \left\|f - f_{\mathcal{T}}\right\|_{L^2(\Omega)}^2 + LS(f_{\mathcal{T}}; \boldsymbol{\sigma}_{\mathrm{PS}}, Gu).$$

The second discrete equation in (8) shows $f_{\mathcal{T}} = \Pi f = -\operatorname{div} \boldsymbol{\sigma}_{\mathrm{PS}}$. Hence,

$$LS(f; \boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}}) \leq \left\|f - f_{\mathcal{T}}\right\|_{L^2(\Omega)}^2 + \left\|\operatorname{dev} \boldsymbol{\sigma}_{\mathrm{PS}} - \mathrm{D}\, Gu\right\|_{L^2(\Omega)}^2.$$

The solution $(\boldsymbol{\sigma}, u)$ to (7) solves (1) with $u \in \mathcal{A}$ and $\operatorname{dev} \boldsymbol{\sigma} = \mathrm{D}\, u$. Therefore,

$$\left\|\operatorname{dev} \boldsymbol{\sigma}_{\mathrm{PS}} - \mathrm{D}\, Gu\right\|_{L^2(\Omega)} \lesssim \left\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_{\mathrm{PS}}\right\|_{L^2(\Omega)} + |||u - Gu|||.$$

A medius analysis shows that the $L^2$ best-approximation of the pseudostress [15, Theorem 5.3] holds in the sense that

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_{\mathrm{PS}}\|_{L^2(\Omega)} \lesssim \|\boldsymbol{\sigma} - \Pi_{\mathrm{RT}}\boldsymbol{\sigma}\|_{L^2(\Omega)} + \operatorname{osc}(f, \mathcal{T}).$$

This and the estimate $\operatorname{osc}(f, \mathcal{T}) \lesssim \|f - f_{\mathcal{T}}\|_{L^2(\Omega)}$ conclude the proof of "$\lesssim$".

The proof of the converse estimate "$\gtrsim$" employs $f + \operatorname{div}\boldsymbol{\sigma} = 0$, $\operatorname{div}\boldsymbol{\sigma}_{\mathrm{LS}} \perp f - f_{\mathcal{T}}$, and the Cauchy-Schwarz estimate to show

$$\|f - f_{\mathcal{T}}\|_{L^2(\Omega)} \leq \|\operatorname{div}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_{\mathrm{LS}})\|_{L^2(\Omega)}.$$

The definition of $\Pi_{\mathrm{RT}}\boldsymbol{\sigma}$ and that of $Gu$ imply

$$\|\boldsymbol{\sigma} - \Pi_{\mathrm{RT}}\boldsymbol{\sigma}\|_{L^2(\Omega)} + |||u - Gu||| \leq \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_{\mathrm{LS}}\|_{L^2(\Omega)} + |||u - u_{\mathrm{LS}}|||.$$

The sum of the two previously displayed estimates leads to

$$\begin{aligned} &\|\boldsymbol{\sigma} - \Pi_{\mathrm{RT}}\boldsymbol{\sigma}\|_{L^2(\Omega)} + |||u - Gu||| + \|f - f_{\mathcal{T}}\|_{L^2(\Omega)} \\ &\quad \leq \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_{\mathrm{LS}}\|_{H(\operatorname{div},\Omega)} + |||u - u_{\mathrm{LS}}|||. \end{aligned} \tag{12}$$

Lemma 2.1 proves the existence of some $z \in H^1(\Omega; \mathbb{R}^2)$ with

$$u - u_{\mathrm{LS}} - z \in V \quad \text{and} \quad |||z||| \lesssim \operatorname{osc}(g', \mathcal{E}(\Gamma)).$$

This, Lemma 2.4 with $\boldsymbol{\tau} \equiv \boldsymbol{\sigma} - \boldsymbol{\sigma}_{\mathrm{LS}}$, $v \equiv u - u_{\mathrm{LS}}$, $z$ as above, and (1) imply

$$\begin{aligned} \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_{\mathrm{LS}}\|^2_{H(\operatorname{div},\Omega)} + |||u - u_{\mathrm{LS}}|||^2 &\lesssim LS(0; \boldsymbol{\sigma} - \boldsymbol{\sigma}_{\mathrm{LS}}, u - u_{\mathrm{LS}}) + |||z|||^2 \\ &\lesssim LS(f; \boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}}) + \operatorname{osc}^2(g', \mathcal{E}(\Gamma)). \end{aligned} \tag{13}$$

The combination of (12)–(13) concludes the proof of "$\gtrsim$". □

### 2.7 Helmholtz decomposition

Recall that $\mathbb{R}^{2\times 2}_{\mathrm{dev}}$ denotes the space of trace-free $2 \times 2$-matrices and define

$$\begin{aligned} Z(\mathcal{T}) &:= \{v_{\mathrm{CR}} \in CR^1_0(\mathcal{T}; \mathbb{R}^2) \,:\, \operatorname{div}_{\mathrm{NC}} v_{\mathrm{CR}} = 0 \text{ a.e. in } \Omega\} \quad \text{and} \\ X(\mathcal{T}) &:= \left\{v_{\mathrm{C}} \in S^1(\mathcal{T}; \mathbb{R}^2) \,:\, \int_\Omega v_{\mathrm{C}} \mathrm{d}x = 0 \text{ and } \int_\Omega \operatorname{curl} v_{\mathrm{C}} \mathrm{d}x = 0\right\}. \end{aligned}$$

For the simply connected domain $\Omega$, the discrete Helmholtz decomposition of [21] leads to the $L^2(\Omega; \mathbb{R}^{2\times 2})$-orthogonal split

$$P_0(\mathcal{T}; \mathbb{R}^{2\times 2}{}_{\mathrm{dev}}) = \mathrm{D}_{\mathrm{NC}} Z(\mathcal{T}) \oplus \operatorname{devCurl} X(\mathcal{T}). \tag{14}$$

### 2.8 tr-dev-div Lemma

There exists some constant $C > 0$ (which depends solely on $\Omega$) such that every $\boldsymbol{\tau} \in \boldsymbol{\Sigma}$ satisfies

$$\|\boldsymbol{\tau}\|_{L^2(\Omega)} \le C \left( \|\operatorname{dev} \boldsymbol{\tau}\|_{L^2(\Omega)} + \|\operatorname{div} \boldsymbol{\tau}\|_{L^2(\Omega)} \right). \tag{15}$$

The proof of (15) follows as in [9, Proposition 9.1.1].

## 3 Alternative a posteriori error control

### 3.1 A posteriori error estimator

For the solution $(\boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}})$ to the discrete equation (11), define an a posteriori error estimator $\eta^2(\mathcal{T}) := \sum_{T\in\mathcal{T}} \eta^2(\mathcal{T}, T)$ by

$$\begin{aligned} \eta^2(\mathcal{T}, T) := |T| \Big( &\|\operatorname{div} \operatorname{dev} \boldsymbol{\sigma}_{\mathrm{LS}}\|^2_{L^2(T)} + \|\operatorname{curl} \operatorname{dev} \boldsymbol{\sigma}_{\mathrm{LS}}\|^2_{L^2(T)} \Big) \\ &+ |T|^{1/2} \sum_{E\in\mathcal{E}(T)\cap\mathcal{E}(\Omega)} \|[\operatorname{dev} \boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\, u_{\mathrm{LS}}]_E\, \nu_E\|^2_{L^2(E)} \\ &+ |T|^{1/2} \sum_{E\in\mathcal{E}(T)} \|[\operatorname{dev}(\boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\, u_{\mathrm{LS}})]_E\, \tau_E\|^2_{L^2(E)} \\ &+ |T|^{1/2} \sum_{E\in\mathcal{E}(T)\cap\mathcal{E}(\Gamma)} \|(1-\Pi)g'\|^2_{L^2(E)} \end{aligned} \tag{16}$$

for any $T \in \mathcal{T}$ and with jumps along the edge $E \in \mathcal{E}$ defined, for any discrete tensor $\boldsymbol{\tau}_{\mathrm{NC}} \in P_1(\mathcal{T}; \mathbb{R}^{2\times 2})$, by

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

For any interior edge $E \in \mathcal{E}(\Omega)$, let $T_+, T_- \in \mathcal{T}$ denote the two neighbouring triangles according to Fig. 2. For $E \in \mathcal{E}(\Gamma)$, let $T_+ \in \mathcal{T}$ denote the only adjacent triangle to $E$. The error estimator $\eta(\mathcal{T})$ is reliable and efficient in that

$$LS(f; \boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}}) \lesssim \eta^2(\mathcal{T}) + \|f - f_{\mathcal{T}}\|^2_{L^2(\Omega)} \lesssim LS(f; \boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}}) + \operatorname{osc}^2(g', \mathcal{E}(\Gamma))$$

from Theorem 3.5 in Sect. 3.4 and Corollary 4.4 in Sect. 4.2 with data oscillation terms $\operatorname{osc}^2(g', \mathcal{E}(\Gamma))$ from Sect. 2.3.

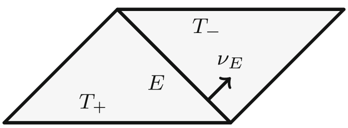


**Fig. 2** Edge patch $\omega_E$

### 3.2 Adaptive algorithm (ALS-FEM)

**Input:** Initial regular triangulation $\mathcal{T}_0$ with refinement edges of the polygonal domain $\Omega$ into triangles and parameters $0 < \theta \leq 1, 0 < \rho < 1, 0 < \kappa < \infty$.
**for** any level $\ell = 0, 1, 2, \ldots$ **do**
**Solve** LS-FEM with respect to regular triangulation $\mathcal{T}_\ell$ with solution $(\boldsymbol{\sigma}_\ell, u_\ell)$ and $f_\ell := \Pi_\ell f$.
**Compute** $(\eta_\ell(T), T \in \mathcal{T}_\ell)$ with $\eta_\ell(\bullet) := \eta(\mathcal{T}_\ell, \bullet)$ from (16).
**if CASE A** $\|f - f_\ell\|^2_{L^2(\Omega)} \leq \kappa \eta_\ell^2$ **then**
**Mark** a subset $\mathcal{M}_\ell$ of $\mathcal{T}_\ell$ of (almost) minimal cardinality $|\mathcal{M}_\ell|$ with

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

**Refine.** Compute the smallest regular refinement $\mathcal{T}_{\ell+1}$ of $\mathcal{T}_\ell$ with $\mathcal{M} \subseteq \mathcal{T}_\ell \backslash \mathcal{T}_{\ell+1}$ by NVB;
**else (CASE B** $\kappa \eta_\ell^2 < \|f - f_\ell\|^2_{L^2(\Omega)}$)
Compute an admissible refinement $\mathcal{T}_{\ell+1}$ of $\mathcal{T}_\ell$ with (almost) minimal cardinality $|\mathcal{T}_{\ell+1}|$ and

$$\|f - f_{\ell+1}\|_{L^2(\Omega)} \leq \rho \|f - f_\ell\|_{L^2(\Omega)}. \quad \textbf{fi od}$$

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

*Remark 3.1* (NVB) The NVB requires an initial condition on the refinement edges in $\mathcal{T}_0$. With reference to [28] for the suppressed details, this is assumed throughout this paper in the definition of $\mathbb{T}$ for refinement control and existence of overlays as summarized in [17, Section 2.4] with further references.

*Remark 3.2* (Case B) The thresholding second algorithm (TSA) of [7, Section 5] is one possible realisation of an optimal refinement in Case B of ALS-FEM. Any other (quasi-)optimal algorithm for the data error reduction may be employed in the algorithm and in the analysis.

### 3.3 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

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

with the best possible error

$$E(u, f, N) := \min_{\mathcal{T} \in \mathbb{T}(N)} \min_{(\boldsymbol{\tau}_{\mathrm{LS}}, v_{\mathrm{LS}}) \in \boldsymbol{\Sigma}(\mathcal{T}) \times S^1(\mathcal{T}; \mathbb{R}^2)} \left( LS(f; \boldsymbol{\tau}_{\mathrm{LS}}, v_{\mathrm{LS}}) + \mathrm{osc}^2(g', \mathcal{E}(\Gamma)) \right).$$

**Theorem 3.3** *There exists a maximal bulk parameter $0 < \theta_0 < 1$ and maximal separation parameter $0 < \kappa_0 < \infty$ which depend exclusively on $\mathcal{T}_0$ such that for all $0 < \theta \leq \theta_0$, for all $0 < \kappa \leq \kappa_0$, for all $0 < \rho < 1$, and for all $0 < s < \infty$, the output $(\boldsymbol{\sigma}_\ell, u_\ell)_\ell$ of ALS-FEM with $(u, f) \in \mathcal{A}_s$ satisfies*

$$\sup_{\ell\in\mathbb{N}} \left(|\mathcal{T}_\ell| - |\mathcal{T}_0|\right)^s \left(LS(f;\boldsymbol{\sigma}_\ell, u_\ell) + \operatorname{osc}^2(g', \mathcal{E}_\ell(\Gamma))\right)^{1/2} \leq C_{\mathrm{qopt}} \left|(u, f)\right|_{\mathcal{A}_s}.$$

*The constant $C_{\mathrm{qopt}} < \infty$ depends only on the initial mesh $\mathcal{T}_0$ the constant $s$ and the parameters $\rho, \theta$, and $\kappa$.*

The proof of Theorem 3.3 will be given in Sect. 4.5. The converse inequality "$\gtrsim$" stated in (2) is elementary.

*Remark 3.4* The equivalence from Theorem 2.6 proves the equivalence of $\mathcal{A}_s$ to the approximation class $\widetilde{\mathcal{A}}_s$ defined as all pairs $(u, f) \in \mathcal{A} \times L^2(\Omega; \mathbb{R}^2)$ with

$$\left|(u, f)\right|^2_{\widetilde{\mathcal{A}}_s} := \sup_{N\in\mathbb{N}} N^{2s} \widetilde{E}(u, f, N) < \infty$$

for the best-approximation error

$$\begin{aligned}\widetilde{E}(u, f, N) := \min_{\mathcal{T}\in\mathbb{T}(N)} \Big(&\|\boldsymbol{\sigma} - \Pi_{\mathrm{RT}}\boldsymbol{\sigma}\|^2_{L^2(\Omega)} + |||u - Gu|||^2 \\ &+ \operatorname{osc}^2(g', \mathcal{E}(\Gamma)) + \|f - f_{\mathcal{T}}\|^2_{L^2(\Omega)}\Big).\end{aligned}$$

Hence, Theorem 3.3 implies (with a different constant $C_{\mathrm{qopt}}$) that

$$\begin{aligned}\sup_{\ell\in\mathbb{N}} \left(|\mathcal{T}_\ell| - |\mathcal{T}_0|\right)^s \Big(&\|\boldsymbol{\sigma} - \Pi_{\mathrm{RT}(\ell)}\boldsymbol{\sigma}\|^2_{L^2(\Omega)} + |||u - G_\ell u|||^2 \\ &+ \operatorname{osc}^2(g', \mathcal{E}_\ell(\Gamma)) + \|f - f_\ell\|^2_{L^2(\Omega)}\Big)^{1/2} \leq C_{\mathrm{qopt}} \left|(u, f)\right|_{\widetilde{\mathcal{A}}_s}.\end{aligned}$$

### 3.4 Efficiency

The discrete test function technology due to Verfürth [29] leads to efficiency of the estimator $\eta$ from Sect. 3.1 in the following sense.

**Theorem 3.5** (efficiency) *The error estimator $\eta^2(\mathcal{T}) := \sum_{T\in\mathcal{T}} \eta^2(\mathcal{T}, T)$ from* (16) *satisfies*

$$\eta^2(\mathcal{T}) + \|f - f_{\mathcal{T}}\|^2_{L^2(\Omega)} \lesssim LS(f; \boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}}) + \operatorname{osc}^2(g', \mathcal{E}(\Gamma)).$$

*Proof* Since $\mathrm{D}\, u_{\mathrm{LS}}\big|_T$ is constant on $T \in \mathcal{T}$, an inverse estimate proves

$$\begin{aligned}&\|\operatorname{div}\operatorname{dev}\boldsymbol{\sigma}_{\mathrm{LS}}\|_{L^2(T)} + \|\operatorname{curl}\operatorname{dev}\boldsymbol{\sigma}_{\mathrm{LS}}\|_{L^2(T)}\\ &\quad = \|\operatorname{div}(\operatorname{dev}\boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\, u_{\mathrm{LS}})\|_{L^2(T)} + \|\operatorname{curl}(\operatorname{dev}\boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\, u_{\mathrm{LS}})\|_{L^2(T)}\\ &\quad \lesssim |T|^{-1/2}\|\operatorname{dev}\boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\, u_{\mathrm{LS}}\|_{L^2(T)}.\end{aligned}$$

Let $E = \partial T_+ \cap \partial T_- \in \mathcal{E}(\Omega)$ and $\omega_E := \operatorname{int}(T_+ \cup T_-)$ as depicted in Fig. 2. A triangle and a trace inequality plus an inverse estimate in the end prove

$$\|[\operatorname{dev}\boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\, u_{\mathrm{LS}}]_E\|_{L^2(E)} \lesssim |E|^{-1/2}\|\operatorname{dev}\boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\, u_{\mathrm{LS}}\|_{L^2(\omega_E)}.$$

The deviatoric part satisfies

$$\|\operatorname{dev}(\boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\, u_{\mathrm{LS}})\|_{L^2(T_\pm)} \le \|\operatorname{dev}\boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\, u_{\mathrm{LS}}\|_{L^2(T_\pm)}.$$

The aforegoing inequalities prove local efficiency of all volume terms on $T$ and all jump terms on interior edges $E \in \mathcal{E}(T) \cap \mathcal{E}(\Omega)$. For any boundary edge $E \in \mathcal{E}(T) \cap \mathcal{E}(\Gamma)$, the trace inequality and an inverse estimate prove

$$\|[\operatorname{dev}(\boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\, u_{\mathrm{LS}})]_E\, \tau_E\|_{L^2(E)} \lesssim |T|^{-1/4}\|\operatorname{dev}\boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\, u_{\mathrm{LS}}\|_{L^2(T)}.$$

The estimate $\|f - f_{\mathcal{T}}\|_{L^2(\Omega)} \le LS(f; \boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}})^{1/2}$ concludes the proof. □

## 4 Convergence analysis of ALS-FEM

This section is devoted to the proof of Theorem 3.3.

### 4.1 Stability and reduction

Let $\widehat{\mathcal{T}}$ be any admissible refinement of a regular triangulation $\mathcal{T}$ with the respective LS-FEM solutions $(\boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}})$ and $(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}, \widehat{u}_{\mathrm{LS}})$. Recall the a posteriori error estimator $\eta^2(\mathcal{T}, \bullet)$ from Sect. 3.1 and $\delta^2(\widehat{\mathcal{T}}, \mathcal{T})$ from Definition 2.5. Abbreviate the contributions of any subset $\mathcal{M} \subseteq \mathcal{T}$ of the triangulation $\mathcal{T}$ as

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

The estimator $\eta$ and the distances $\delta$ satisfy the first two axioms of adaptivity from [17] with generic constants $C_{\mathrm{stab}} \approx 1 \approx C_{\mathrm{red}}$ and $0 < \rho_{\mathrm{red}} < 1$.

**Theorem 4.1** (stability) *There exists $C_{\mathrm{stab}} \approx 1$ such that*

$$\left|\eta(\widehat{\mathcal{T}}, \widehat{\mathcal{T}} \cap \mathcal{T}) - \eta(\mathcal{T}, \widehat{\mathcal{T}} \cap \mathcal{T})\right| \le C_{\mathrm{stab}}\delta(\widehat{\mathcal{T}}, \mathcal{T}).$$

*Proof* The proof of the stability of the volume and edge contributions

$$\begin{aligned}
&|T|\Big(\|\operatorname{div}\operatorname{dev}\boldsymbol{\sigma}_{\mathrm{LS}}\|^2_{L^2(T)}+\|\operatorname{curl}\operatorname{dev}\boldsymbol{\sigma}_{\mathrm{LS}}\|^2_{L^2(T)}\Big)\\
&\quad+|T|^{1/2}\sum_{E\in\mathcal{E}(T)\cap\mathcal{E}(\Omega)}\|[\operatorname{dev}\boldsymbol{\sigma}_{\mathrm{LS}}-\mathrm{D}\,u_{\mathrm{LS}}]_E\,\nu_E\|^2_{L^2(E)}\\
&\quad+|T|^{1/2}\sum_{E\in\mathcal{E}(T)}\|[\operatorname{dev}(\boldsymbol{\sigma}_{\mathrm{LS}}-\mathrm{D}\,u_{\mathrm{LS}})]_E\,\tau_E\|^2_{L^2(E)}
\end{aligned}\tag{17}$$

follows the lines of that in [23, Corollary 3.4] and in [27, Proposition 4.6]. Details are therefore omitted here.

Since $T\in\mathcal{T}\cap\widehat{\mathcal{T}}$, the remaining contributions of boundary data oscillations coincide, $\|(1-\widehat{\Pi})g'\|_{L^2(\partial T\cap\Gamma)}=\|(1-\Pi)g'\|_{L^2(\partial T\cap\Gamma)}$. This concludes the proof. □

**Theorem 4.2** (reduction) *There exist* $0<\rho_{\mathrm{red}}<1$ *and* $C_{\mathrm{red}}\approx 1$ *such that*

$$\eta^2(\widehat{\mathcal{T}},\widehat{\mathcal{T}}\backslash\mathcal{T})\le\rho_{\mathrm{red}}\,\eta^2(\mathcal{T},\mathcal{T}\backslash\widehat{\mathcal{T}})+C_{\mathrm{red}}\,\delta^2(\widehat{\mathcal{T}},\mathcal{T}).$$

*Proof* The proof of the reduction of the volume and edge contributions (17) relies on the fact that each term is weighted with a corresponding power of the mesh-size $|T|$ (which is reduced at least by a factor 2), cf. the proof of [23, Corollary 3.4] for details. This leads to the reduction constants

$$\widetilde{\rho}_{\mathrm{red}}:=(1+\lambda)\,2^{-1/2}\quad\text{and}\quad\widetilde{C}_{\mathrm{red}}:=(1+1/\lambda)\big(2C_{\mathrm{inv}}+48C_{\mathrm{tr}}(1+C_{\mathrm{inv}})\big)$$

with generic constants $C_{\mathrm{inv}}<\infty$ from an inverse estimate, $C_{\mathrm{tr}}<\infty$ from the trace inequality, and for any parameter $0<\lambda$. Choose $0<\lambda$ sufficiently small to guarantee $\widetilde{\rho}_{\mathrm{red}}<1$.

The reduction of the remaining boundary data oscillations follows directly from Lemma 2.2, for $K\in\mathcal{T}\backslash\widehat{\mathcal{T}}$ and $T\in\widehat{\mathcal{T}}(K)$,

$$\begin{aligned}
&|T|^{1/2}\|(1-\widehat{\Pi})g'\|^2_{L^2(\Gamma\cap\partial T)}\\
&\quad\le\big(|K|/2\big)^{1/2}\Big(\|(1-\widehat{\Pi})g'\|^2_{L^2(\Gamma\cap\partial T)}+\|(\widehat{\Pi}-\Pi)g'\|^2_{L^2(\Gamma\cap\partial T)}\Big)\\
&\quad=\big(|K|/2\big)^{1/2}\|(1-\Pi)g'\|^2_{L^2(\Gamma\cap\partial T)}.
\end{aligned}$$

The sum over all $K\in\mathcal{T}\backslash\widehat{\mathcal{T}}$ and $T\in\widehat{\mathcal{T}}(K)$ leads to

$$\operatorname{osc}^2(g',\widehat{\mathcal{T}}\backslash\mathcal{T})\le 2^{-1/2}\operatorname{osc}^2(g',\mathcal{T}\backslash\widehat{\mathcal{T}}).$$

This concludes the proof with the constants $0<\rho_{\mathrm{red}}:=\widetilde{\rho}_{\mathrm{red}}<1$ and $C_{\mathrm{red}}:=\widetilde{C}_{\mathrm{red}}<\infty$. □

### 4.2 Discrete reliability

The reliability of the error estimator (16) is the key to the analysis and requires a modification by the extra term $\|(1-\Pi)\operatorname{div}\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}\|_{L^2(\Omega)}$.

**Theorem 4.3** (discrete reliability) *There exists some constant $C_{\mathrm{drel}} \approx 1$ such that any admissible refinement $\widehat{\mathcal{T}}$ of $\mathcal{T}$ in $\mathbb{T}$ with discrete solutions $(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}, \widehat{u}_{\mathrm{LS}})$ and $(\boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}})$ to* (11) *with respect to and the error estimator $\eta(\mathcal{T}, \bullet)$ from* (16) *satisfy*

$$\delta^2(\widehat{\mathcal{T}}, \mathcal{T}) \le C_{\mathrm{drel}} \left(\eta^2(\mathcal{T}, \mathcal{T}\backslash\widehat{\mathcal{T}}) + \left\|(1-\Pi)\operatorname{div}\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}\right\|^2_{L^2(\Omega)}\right).$$

The last term gives rise to reliability in the following sense.

**Corollary 4.4** (reliability) *Given an admissible triangulation $\mathcal{T} \in \mathbb{T}$ with discrete solutions $(\boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}}) \in \boldsymbol{\Sigma}(\mathcal{T}) \times \mathcal{A}(\mathcal{T})$ to* (11)*, the error estimator $\eta(\mathcal{T}, \bullet)$ is reliable in the sense that*

$$LS(f; \boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}}) \lesssim \eta^2(\mathcal{T}) + \|f - f_{\mathcal{T}}\|^2_{L^2(\Omega)}. \tag{18}$$

*Proof* (Proof of Corollary 4.4) Define the sequence $(\mathcal{T}_j)_{j\in\mathbb{N}}$ of successive uniform one-level refinements $\mathcal{T}_j := \mathrm{bisec3}^{(j)}(\mathcal{T})$ with discrete solutions $(\boldsymbol{\sigma}_j, u_j) \in \boldsymbol{\Sigma}(\mathcal{T}_j) \times \mathcal{A}(\mathcal{T}_j)$ to (11). This design ensures uniform convergence of the mesh-sizes $h_j$ as $j \to \infty$,

$$\lim_{j\to\infty} \|h_j\|_{L^\infty(\Omega)} = 0.$$

The convergence of the LS-FEM yields

$$\begin{aligned}
&\lim_{j\to\infty} \delta^2(\mathcal{T}_j, \mathcal{T}) \\
&\quad = \lim_{j\to\infty} \left(\|\operatorname{div}(\boldsymbol{\sigma}_j - \boldsymbol{\sigma}_{\mathrm{LS}})\|^2_{L^2(\Omega)} + \|\operatorname{dev}(\boldsymbol{\sigma}_j - \boldsymbol{\sigma}_{\mathrm{LS}}) - \mathrm{D}(u_j - u_{\mathrm{LS}})\|^2_{L^2(\Omega)}\right) \\
&\quad = LS(f; \boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}})
\end{aligned} \tag{19}$$

and

$$\lim_{j\to\infty} \|(1-\Pi)\operatorname{div}\boldsymbol{\sigma}_j\|^2_{L^2(\Omega)} = \|f - f_{\mathcal{T}}\|^2_{L^2(\Omega)}. \tag{20}$$

Theorem 4.3 implies, for every $j \in \mathbb{N}$, that

$$\delta^2(\mathcal{T}_j, \mathcal{T}) \le C_{\mathrm{drel}}\, \eta^2(\mathcal{T}) + \|(1-\Pi)\operatorname{div}\boldsymbol{\sigma}_j\|^2_{L^2(\Omega)}.$$

This and (19)–(20) conclude the proof for $j \to \infty$. □

The remainder of this subsection is devoted to the proof of Theorem 4.3. Recall that $\Pi$ denotes the $L^2(\Omega)$-orthogonal projection onto the piecewise constant functions $P_0(\mathcal{T})$. The following proofs involve three PS-FEM solutions $\widehat{\boldsymbol{\tau}}_{\mathrm{PS}}, \widehat{\boldsymbol{\tau}}^*_{\mathrm{PS}}$, and $\boldsymbol{\tau}_{\mathrm{PS}}$, which allow the split of the term $\operatorname{dev}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}})$ into a divergence-free part and a remaining part $\widehat{\boldsymbol{\tau}}_{\mathrm{PS}} - \widehat{\boldsymbol{\tau}}^*_{\mathrm{PS}}$ in Lemma 4.5. This lemma mainly consists of algebraic rearrangements in such a way that the resulting terms can be treated by integration by parts in combination with a Scott-Zhang quasi-interpolation in the Lemmas 4.6–4.7.

Let $\widehat{\boldsymbol{\tau}}_{\mathrm{PS}}$, $\widehat{\boldsymbol{\tau}}^*_{\mathrm{PS}}$, and $\boldsymbol{\tau}_{\mathrm{PS}}$ solve the PS-FEM of Sect. 2.4 with homogeneous boundary conditions $g \equiv 0$, the right-hand sides

$$-\operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}), \quad -\Pi \operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}), \quad \text{and} \quad -\Pi \operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}})$$

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

$$\begin{aligned} \operatorname{div} \widehat{\boldsymbol{\tau}}_{\mathrm{PS}} &= \operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}) \quad \text{and} \\ \operatorname{div} \widehat{\boldsymbol{\tau}}^*_{\mathrm{PS}} &= \Pi \operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}) = \operatorname{div} \boldsymbol{\tau}_{\mathrm{PS}}. \end{aligned} \tag{21}$$

Recall the function spaces $X(\mathcal{T})$ and $X(\widehat{\mathcal{T}})$ from Sect. 2.7. The proof of the discrete reliability in Theorem 4.3 uses the following three lemmas. Their extensive and technical proofs are postponed to the appendix to improve readability of this section.

**Lemma 4.5** *There exist some $\widehat{z} \in S^1(\widehat{\mathcal{T}}; \mathbb{R}^2)$ and $\widehat{\beta} \in X(\widehat{\mathcal{T}})$ with*

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

**Lemma 4.6** *It holds that*

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

**Lemma 4.7** *It holds that*

$$\begin{aligned}&\int_\Omega (\operatorname{dev}\boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\,u_{\mathrm{LS}}) : \operatorname{dev}\operatorname{Curl}\widehat{\beta}\,\mathrm{d}x\\ &\quad\lesssim \|\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}\|_{H(\mathrm{div},\Omega)}\Big(\sum_{T\in\mathcal{T}\setminus\widehat{\mathcal{T}}}\Big(|T|\,\|\operatorname{curl}\operatorname{dev}\boldsymbol{\sigma}_{\mathrm{LS}}\|_{L^2(T)}^2\\ &\qquad + \sum_{E\in\mathcal{E}(T)}|T|^{1/2}\|[\operatorname{dev}(\boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\,u_{\mathrm{LS}})]_E\,\tau_E\|_{L^2(E)}^2\Big)\Big)^{1/2}.\end{aligned}$$

*Proof* (Proof of Theorem 4.3) For $\widehat{z}$ from Lemma 4.5, the design of $\widehat{\Pi}g$ and $\Pi g$ and the orthogonality from Lemma 2.2 yield

$$\begin{aligned}|||\widehat{z}|||^2 &\lesssim \operatorname{osc}^2(\widehat{\Pi}g', \mathcal{E}(\Gamma)) = \sum_{E\in\mathcal{E}(\Gamma)\setminus\widehat{\mathcal{E}}(\Gamma)}\|h_{\mathcal{E}}^{1/2}(\widehat{\Pi} - \Pi)g'\|_{L^2(E)}^2\\ &\le \sum_{E\in\mathcal{E}(\Gamma)\setminus\widehat{\mathcal{E}}(\Gamma)}\|h_{\mathcal{E}}^{1/2}(1 - \Pi)g'\|_{L^2(E)}^2 \lesssim \eta^2(\mathcal{T}, \mathcal{T}\setminus\widehat{\mathcal{T}}).\end{aligned}\tag{22}$$

Recall (53) from the proof of Lemma 4.7 and deduce

$$\|\operatorname{dev}(\widehat{\boldsymbol{\tau}}_{\mathrm{PS}} - \widehat{\boldsymbol{\tau}}_{\mathrm{PS}}^*)\|_{L^2(\Omega)} \le \|\widehat{\boldsymbol{\tau}}_{\mathrm{PS}} - \widehat{\boldsymbol{\tau}}_{\mathrm{PS}}^*\|_{L^2(\Omega)} \lesssim \|(1 - \Pi)\operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}})\|_{L^2(\Omega)}.$$

The Cauchy-Schwarz inequality, the triangle inequality, Lemma 2.4 with $\boldsymbol{\tau} \equiv \widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}$, $v \equiv \widehat{u}_{\mathrm{LS}} - u_{\mathrm{LS}}$, and $w \equiv \widehat{z}$ plus the previous estimate imply

$$\begin{aligned}&\int_\Omega (\operatorname{dev}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}) - \mathrm{D}(\widehat{u}_{\mathrm{LS}} - u_{\mathrm{LS}})) : \big(\operatorname{dev}(\widehat{\boldsymbol{\tau}}_{\mathrm{PS}} - \widehat{\boldsymbol{\tau}}_{\mathrm{PS}}^*) - \mathrm{D}\widehat{z}\big)\mathrm{d}x\\ &\quad\lesssim \|\operatorname{dev}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}) - \mathrm{D}(\widehat{u}_{\mathrm{LS}} - u_{\mathrm{LS}})\|_{L^2(\Omega)}\\ &\qquad\times\big(\|(1 - \Pi)\operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}})\|_{L^2(\Omega)} + |||\widehat{z}|||\big)\\ &\quad\lesssim \big(\|\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}\|_{H(\mathrm{div},\Omega)} + |||\widehat{u}_{\mathrm{LS}} - u_{\mathrm{LS}}||| + |||\widehat{z}|||\big)\\ &\qquad\times\big(\|(1 - \Pi)\operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}})\|_{L^2(\Omega)} + |||\widehat{z}|||\big).\end{aligned}\tag{23}$$

The converse estimate from Lemma 2.4 reads

$$\begin{aligned}\|\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}\|_{H(\mathrm{div},\Omega)}^2 + |||\widehat{u}_{\mathrm{LS}} - u_{\mathrm{LS}}|||^2 &\lesssim \|\operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}})\|_{L^2(\Omega)}^2\\ &\quad + \|\operatorname{dev}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}) - \mathrm{D}(\widehat{u}_{\mathrm{LS}} - u_{\mathrm{LS}})\|_{L^2(\Omega)}^2 + |||\widehat{z}|||^2.\end{aligned}$$

The combination of this with Lemma 4.5–4.7 and (23) shows

$$
\begin{aligned}
&\|\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}-\boldsymbol{\sigma}_{\mathrm{LS}}\|_{H(\mathrm{div},\Omega)}^2+|||\widehat{u}_{\mathrm{LS}}-u_{\mathrm{LS}}|||^2\\
&\quad\lesssim \|(1-\Pi)\operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}-\boldsymbol{\sigma}_{\mathrm{LS}})\|_{L^2(\Omega)}^2\\
&\qquad+\big(\|\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}-\boldsymbol{\sigma}_{\mathrm{LS}}\|_{H(\mathrm{div},\Omega)}+|||\widehat{u}_{\mathrm{LS}}-u_{\mathrm{LS}}|||+|||\widehat{z}|||\big)\\
&\qquad\times\big(\|(1-\Pi)\operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}-\boldsymbol{\sigma}_{\mathrm{LS}})\|_{L^2(\Omega)}+|||\widehat{z}|||\big)\\
&\qquad+\eta(\mathcal{T},\mathcal{T}\setminus\widehat{\mathcal{T}})\big(|||\widehat{u}_{\mathrm{LS}}-u_{\mathrm{LS}}|||+|||\widehat{z}|||+\|\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}-\boldsymbol{\sigma}_{\mathrm{LS}}\|_{H(\mathrm{div},\Omega)}\big)+|||\widehat{z}|||^2.
\end{aligned}
$$

This, (22), and some standard rearrangements conclude the proof. □

### 4.3 Quasi-orthogonality

Recall $\delta^2(\widehat{\mathcal{T}},\mathcal{T})$ from the Definition 2.5.

**Theorem 4.8** (quasi-orthogonality) *Any regular triangulation $\mathcal{T}$ with admissible refinement $\widehat{\mathcal{T}}$, the corresponding solutions $(\boldsymbol{\sigma}_{\mathrm{LS}},u_{\mathrm{LS}})$ and $(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}},\widehat{u}_{\mathrm{LS}})$ to the discrete equation (11) with respect to $\mathcal{T}$ and $\widehat{\mathcal{T}}$, and any $0<\mu$ satisfy*

$$
\begin{aligned}
\delta^2(\widehat{\mathcal{T}},\mathcal{T})\le{}& LS(f;\boldsymbol{\sigma}_{\mathrm{LS}},u_{\mathrm{LS}})-(1-\mu)LS(f;\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}},\widehat{u}_{\mathrm{LS}})\\
&+(1+C_{\mathrm{osc}}/\mu)\big(\operatorname{osc}^2(g',\mathcal{E}(\Gamma))-\operatorname{osc}^2(g',\widehat{\mathcal{E}}(\Gamma))\big).
\end{aligned}
$$

*Remark 4.9* The assertion in Theorem 4.8 refers to axiom (B3a) from [17] with $\mu(\mathcal{T}):=\operatorname{osc}(g',\mathcal{E}(\Gamma))$ and this implies axiom (B3b) therein. The reliability from Corollary 4.4 and [17, Lemma 3.7] prove, for any sequence of successive admissible refinements $\mathcal{T}_0,\mathcal{T}_1,\dots$ and all $\varepsilon_{\mathrm{qo}}>0$, quasi-orthogonality in the generalized sense that

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

*Proof* (Proof of Theorem 4.8) Abbreviate the exact solution $X:=(\boldsymbol{\sigma},u)$ to the continuous least-squares problem (9) and the discrete solutions $X_{\mathrm{LS}}:=(\boldsymbol{\sigma}_{\mathrm{LS}},u_{\mathrm{LS}})$ and $\widehat{X}_{\mathrm{LS}}:=(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}},\widehat{u}_{\mathrm{LS}})$ to (11). Lemma 2.1 with $g$ replaced by $(\widehat{I}-I)g\in S^1(\widehat{\mathcal{E}}(\Gamma);\mathbb{R}^2)$ yields the existence of some generic constant $C_{\mathrm{osc}}\approx 1$ and some $\widehat{w}\in S^1(\widehat{\mathcal{T}};\mathbb{R}^2)$ with $\widehat{u}_{\mathrm{LS}}-u_{\mathrm{LS}}-\widehat{w}\in S_0^1(\widehat{\mathcal{T}};\mathbb{R}^2)$ and

$$
|||\widehat{w}|||^2\le C_{\mathrm{osc}}\operatorname{osc}^2(\widehat{g}',\mathcal{E}(\Gamma)). \tag{24}
$$

Because of the non-homogeneous Dirichlet boundary data, the Galerkin orthogonality holds in general exclusively for velocity test functions that vanish on the boundary. Hence,

$$
\mathcal{B}(X-\widehat{X}_{\mathrm{LS}};\widehat{X}_{\mathrm{LS}}-X_{\mathrm{LS}}-(0,\widehat{w}))=0.
$$

This plus elementary algebra with the symmetric bilinear form $\mathcal{B}$ prove

$$\begin{aligned}
&\mathcal{B}(\widehat{X}_{\mathrm{LS}} - X_{\mathrm{LS}}; \widehat{X}_{\mathrm{LS}} - X_{\mathrm{LS}}) \\
&\quad = \mathcal{B}(X - X_{\mathrm{LS}}; X - X_{\mathrm{LS}}) - \mathcal{B}(X - \widehat{X}_{\mathrm{LS}}; X - \widehat{X}_{\mathrm{LS}}) - 2\mathcal{B}(X - \widehat{X}_{\mathrm{LS}}; 0, \widehat{w}).
\end{aligned} \tag{25}$$

The rewriting in terms of the least-squares functional yields

$$\begin{aligned}
\delta^2(\widehat{\mathcal{T}}, \mathcal{T}) = LS(f; \boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}}) - LS(f; \widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}, \widehat{u}_{\mathrm{LS}}) \\
+ \operatorname{osc}^2(\widehat{g}', \mathcal{E}(\Gamma)) - 2\mathcal{B}(X - \widehat{X}_{\mathrm{LS}}; 0, \widehat{w}).
\end{aligned}$$

Lemma 2.2 implies

$$\operatorname{osc}^2(\widehat{g}', \mathcal{E}(\Gamma)) \leq \operatorname{osc}^2(g', \mathcal{E}(\Gamma)) - \operatorname{osc}^2(g', \widehat{\mathcal{E}}(\Gamma)).$$

The combination of the two previously displayed formulas, the Cauchy-Schwarz inequality, the Young inequality, and (24) imply the assertion for any parameter $\mu > 0$. $\square$

### 4.4 Contraction property

Recall the output $(\mathcal{T}_\ell)_{\ell\in\mathbb{N}}$ and $(\boldsymbol{\sigma}_\ell, u_\ell)_{\ell\in\mathbb{N}}$ of ALS-FEM from Sect. 3.2.

**Theorem 4.10** (contraction) *For all $0 < \theta < 1$, $0 < \kappa < \infty$, and $0 < \rho < 1$ from the input of the adaptive algorithm in Sect. 3.2, there exist constants $\Lambda_{\mathrm{con}}, \Lambda_{\mathrm{osc}} \approx 1$, and $0 < \rho_{\mathrm{con}} < 1$ such that*

$$\xi_\ell^2 := LS(f; \boldsymbol{\sigma}_\ell, u_\ell) + \|f - f_\ell\|^2_{L^2(\Omega)} + \Lambda_{\mathrm{osc}} \operatorname{osc}^2(g', \mathcal{E}_\ell(\Gamma)) + \Lambda_{\mathrm{con}} \eta_\ell^2 \tag{26}$$

*satisfies*

$$\xi_{\ell+1}^2 \leq \rho_{\mathrm{con}} \xi_\ell^2 \quad \text{for all } \ell \in \mathbb{N}_0. \tag{27}$$

*Proof* *Step 1:* The Theorems 4.1–4.2 motivate the additive split

$$\eta_{\ell+1}^2 = \eta_{\ell+1}^2(\mathcal{T}_\ell \cap \mathcal{T}_{\ell+1}) + \eta_{\ell+1}^2(\mathcal{T}_{\ell+1} \backslash \mathcal{T}_\ell). \tag{28}$$

For any $0 < \lambda$ with $\Lambda'_{\mathrm{red}} := ((1 + 1/\lambda)C^2_{\mathrm{stab}} + C_{\mathrm{red}})$, the Theorem 4.1 for $\mathcal{T}_\ell \cap \mathcal{T}_{\ell+1}$, and Theorem 4.2 for $\mathcal{T}_\ell \backslash \mathcal{T}_{\ell+1}$ with $\widehat{\mathcal{T}} \equiv \mathcal{T}_{\ell+1}$, $\mathcal{T} \equiv \mathcal{T}_\ell$ as well as (28) imply

$$\begin{aligned}
\eta_{\ell+1}^2 &\leq (1 + \lambda)\eta_\ell^2(\mathcal{T}_\ell \cap \mathcal{T}_{\ell+1}) + \rho_{\mathrm{red}} \eta_\ell^2(\mathcal{T}_\ell \backslash \mathcal{T}_{\ell+1}) + \Lambda'_{\mathrm{red}} \delta^2(\mathcal{T}_{\ell+1}, \mathcal{T}_\ell) \\
&= (1 + \lambda)\eta_\ell^2 - (1 + \lambda - \rho_{\mathrm{red}})\eta_\ell^2(\mathcal{T}_\ell \backslash \mathcal{T}_{\ell+1}) + \Lambda'_{\mathrm{red}} \delta^2(\mathcal{T}_{\ell+1}, \mathcal{T}_\ell).
\end{aligned} \tag{29}$$

For Case A, the Dörfler marking guarantees

$$\theta \eta_\ell^2 \leq \eta_\ell^2(\mathcal{M}_\ell) \leq \eta_\ell^2(\mathcal{T}_\ell \backslash \mathcal{T}_{\ell+1}).$$

For sufficiently small $0 < \lambda$ with $0 < \rho_{\text{red}}(\ell) := (1+\lambda)(1-(1-\rho_{\text{red}})\theta) < 1$ and $\Lambda'_{\text{red}} \approx 1 + 1/\lambda$, the two previously displayed formulas lead to

$$\eta_{\ell+1}^2 \leq \rho_{\text{red}}(\ell)\eta_\ell^2 + \Lambda'_{\text{red}}\delta^2(\mathcal{T}_{\ell+1}, \mathcal{T}_\ell). \tag{30}$$

In Case A, the data approximation is possibly not strictly reduced. The minimization in the definition of $f_\ell$ and the inclusion $P_0(\mathcal{T}_\ell; \mathbb{R}^2) \subseteq P_0(\mathcal{T}_{\ell+1}; \mathbb{R}^2)$ imply, for $\rho(\ell) := 1$, that

$$\|f - f_{\ell+1}\|_{L^2(\Omega)}^2 \leq \rho(\ell)\|f - f_\ell\|_{L^2(\Omega)}^2. \tag{31}$$

For Case B, however, (29) directly implies (30) with $\rho_{\text{red}}(\ell) := 1+\lambda$ for any $0 < \lambda$ and (31) holds for $0 < \rho(\ell) := \rho < 1$.

*Step 2:* Theorem 4.8 with $\mathcal{T} \equiv \mathcal{T}_\ell$, $\widehat{\mathcal{T}} \equiv \mathcal{T}_{\ell+1}$, and $0 < \mu < 1$ proves

$$\begin{aligned}\delta^2(\mathcal{T}_{\ell+1}, \mathcal{T}_\ell) \leq{}& LS(f; \boldsymbol{\sigma}_\ell, u_\ell) - (1-\mu)LS(f; \boldsymbol{\sigma}_{\ell+1}, u_{\ell+1})\\ &+ (1 + C_{\text{osc}}/\mu)\big(\operatorname{osc}^2(g', \mathcal{E}_\ell(\Gamma)) - \operatorname{osc}^2(g', \mathcal{E}_{\ell+1}(\Gamma))\big).\end{aligned}$$

The previous estimate and the estimator reduction (30) imply

$$\begin{aligned}\eta_{\ell+1}^2 \leq{}& \rho_{\text{red}}(\ell)\eta_\ell^2 + \Lambda'_{\text{red}}\Big(LS(f; \boldsymbol{\sigma}_\ell, u_\ell) - (1-\mu)LS(f; \boldsymbol{\sigma}_{\ell+1}, u_{\ell+1})\\ &+ (1 + C_{\text{osc}}/\mu)\big(\operatorname{osc}^2(g', \mathcal{E}_\ell(\Gamma)) - \operatorname{osc}^2(g', \mathcal{E}_{\ell+1}(\Gamma))\big)\Big).\end{aligned}$$

Hence, $\Lambda_{\text{con}} := 1/((1-\mu)\Lambda'_{\text{red}})$ and $\Lambda_{\text{osc}} := (1 + C_{\text{osc}}/\mu)/(1-\mu)$ satisfy

$$\begin{aligned}&LS(f; \boldsymbol{\sigma}_{\ell+1}, u_{\ell+1}) + \Lambda_{\text{osc}}\operatorname{osc}^2(g', \mathcal{E}_{\ell+1}(\Gamma)) + \Lambda_{\text{con}}\eta_{\ell+1}^2\\ &\quad\leq 1/(1-\mu)\, LS(f; \boldsymbol{\sigma}_\ell, u_\ell) + \Lambda_{\text{osc}}\operatorname{osc}^2(g', \mathcal{E}_\ell(\Gamma)) + \rho_{\text{red}}(\ell)\Lambda_{\text{con}}\eta_\ell^2.\end{aligned} \tag{32}$$

For $0 < \mu < \min\{\varepsilon, \varepsilon/\Lambda_{\text{osc}}, \varepsilon/\Lambda_{\text{con}}\}$, set

$$\begin{aligned}\rho_{\text{con}}(\varepsilon) &:= \max\{(1-\varepsilon)/(1-\mu), 1 - \varepsilon/\Lambda_{\text{osc}}, 1 - \varepsilon/\Lambda_{\text{con}}\} < 1,\\ B(\varepsilon) &:= \varepsilon/(1-\mu)\, LS(f; \boldsymbol{\sigma}_\ell, u_\ell) + \|f - f_{\ell+1}\|_{L^2(\Omega)}^2 - (1-\varepsilon)\|f - f_\ell\|_{L^2(\Omega)}^2\\ &\quad + \varepsilon\operatorname{osc}^2(g', \mathcal{E}_\ell(\Gamma)) + \big(\varepsilon + \Lambda_{\text{con}}(\rho_{\text{red}}(\ell) - 1)\big)\eta_\ell^2.\end{aligned}$$

The combination with (32) leads to

$$\xi_{\ell+1}^2 \leq \rho_{\text{con}}(\varepsilon)\xi_\ell^2 + B(\varepsilon). \tag{33}$$

To estimate $B(\varepsilon) \leq 0$ in Case A and B, notice that $\operatorname{osc}^2(g', \mathcal{E}_\ell(\Gamma)) \leq \eta_\ell^2$ and the reliability of the estimator $\eta_\ell$ from Corollary 4.4 with generic constant $C_{\text{rel}} \approx 1$ imply

$$LS(f; \boldsymbol{\sigma}_\ell, u_\ell) + \operatorname{osc}^2(g', \mathcal{E}_\ell(\Gamma)) \leq (1 + C_{\text{rel}})\eta_\ell^2 + C_{\text{rel}}\|f - f_\ell\|_{L^2(\Omega)}^2. \tag{34}$$

*Step 3 (Case A):* Since $\|f - f_{\ell+1}\|^2_{L^2(\Omega)} \le \|f - f_\ell\|^2_{L^2(\Omega)} \le \kappa\eta_\ell^2, 0 < \rho_{\mathrm{red}}(\ell) < 1$ and (34) yield

$$\begin{aligned} B(\varepsilon) &\le \varepsilon(1 + C_{\mathrm{rel}}/(1-\mu))\eta_\ell^2 + \varepsilon(1 + C_{\mathrm{rel}}/(1-\mu))\big\|f - f_\ell\big\|^2_{L^2(\Omega)} \\ &\qquad + \big(\varepsilon + \Lambda_{\mathrm{con}}(\rho_{\mathrm{red}}(\ell) - 1)\big)\eta_\ell^2 \\ &\le \big(\varepsilon(1 + C_{\mathrm{rel}}/(1-\mu))(1+\kappa) + \varepsilon + \Lambda_{\mathrm{con}}(\rho_{\mathrm{red}}(\ell) - 1)\big)\eta_\ell^2. \end{aligned}$$

Since $\rho_{\mathrm{red}}(\ell) - 1 < 0$, it is possible to choose $0 < \varepsilon$ sufficiently small such that $B(\varepsilon) \le 0$. This and (33) conclude the proof of (27) in Case A.

*Step 4 (Case B):* Recall, for any $0 < \lambda$, that $\rho_{\mathrm{red}}(\ell) := 1 + \lambda$,

$$\big\|f - f_{\ell+1}\big\|^2_{L^2(\Omega)} \le \rho\big\|f - f_\ell\big\|^2_{L^2(\Omega)}, \quad \text{and} \quad \eta_\ell^2 \le 1/\kappa\ \big\|f - f_\ell\big\|^2_{L^2(\Omega)}.$$

This plus (34) prove

$$\begin{aligned} B(\varepsilon) &\le \varepsilon(1 + C_{\mathrm{rel}}/(1-\mu))\eta_\ell^2 + \big(\varepsilon(1 + C_{\mathrm{rel}}/(1-\mu)) + \rho - 1\big)\big\|f - f_\ell\big\|^2_{L^2(\Omega)} \\ &\qquad + (\varepsilon + \lambda\Lambda_{\mathrm{con}})\eta_\ell^2 \\ &\le \big(\varepsilon(1 + C_{\mathrm{rel}}/(1-\mu))(1 + 1/\kappa) + (\varepsilon + \lambda\Lambda_{\mathrm{con}})/\kappa + \rho - 1\big)\big\|f - f_\ell\big\|^2_{L^2(\Omega)}. \end{aligned}$$

Since $\rho < 1$, for sufficiently small $0 < \varepsilon$ and $0 < \lambda$, it follows that

$$\big(\varepsilon(1 + C_{\mathrm{rel}}/(1-\mu))(1 + 1/\kappa) + \varepsilon/\kappa + \lambda\Lambda_{\mathrm{con}}/\kappa + \rho - 1\big) < 0.$$

Hence, $B(\varepsilon) \le 0$. This and (33) conclude the proof of (27) in Case B. $\square$

### 4.5 Optimal convergence rates

The proof of the main result Theorem 3.3 follows the arguments of [20], but involves additional estimates for the non-homogeneous boundary data.

*Proof* (Proof of Theorem 3.3) *Step 1:* Let $\ell \in \mathbb{N}$. Recall the definition of $\xi_\ell$ from (26). For $\varepsilon(\ell) := \tau\xi_\ell$ with a parameter $0 < \tau < \big|(u, f)\big|_{\mathcal{A}_s}/\xi_0$, an argument in [20, page 58] leads to $N(\ell) \in \mathbb{N}$ with

$$2 \le N(\ell) \le 2\,\big|(u, f)\big|^{1/s}_{\mathcal{A}_s}\varepsilon(\ell)^{-1/s}. \tag{35}$$

*Step 2:* The definition of $E(u, f, N(\ell))$ implies the existence of an optimal admissible triangulation $\widetilde{\mathcal{T}}_\ell \in \mathbb{T}(N(\ell))$ with solution $(\widetilde{\boldsymbol{\sigma}}_\ell, \widetilde{u}_\ell)$ to the discrete equation (11) on $\widetilde{\mathcal{T}}_\ell$, Dirichlet boundary data approximation $\widetilde{g}_\ell$, and

$$E(u, f, N(\ell)) = LS(f; \widetilde{\boldsymbol{\sigma}}_\ell, \widetilde{u}_\ell) + \mathrm{osc}^2(g', \widetilde{\mathcal{E}}_\ell(\Gamma)).$$

This, the supremum in the definition of $\left|(u, f)\right|_{\mathcal{A}_s}$, and the choice of $N(\ell)$ imply

$$\begin{aligned} LS(f;\widetilde{\boldsymbol{\sigma}}_\ell,\widetilde{u}_\ell)+\operatorname{osc}^2(g',\widetilde{\mathcal{E}}_\ell(\Gamma)) &= E(u,f,N(\ell)) \\ &\le N(\ell)^{-2s}\left|(u,f)\right|^2_{\mathcal{A}_s} \le \varepsilon(\ell)^2=\tau^2\xi_\ell^2. \end{aligned} \tag{36}$$

The smallest common refinement $\widehat{\mathcal{T}}_\ell := \mathcal{T}_\ell \otimes \widetilde{\mathcal{T}}_\ell \in \mathbb{T}$, called overlay of $\mathcal{T}_\ell$ and $\widetilde{\mathcal{T}}_\ell$, is an admissible refinement of $\widetilde{\mathcal{T}}_\ell$ and satisfies [23, Lemma 3.7]

$$\left|\mathcal{T}_\ell\backslash\widehat{\mathcal{T}}_\ell\right| \le \left|\widehat{\mathcal{T}}_\ell\right|-\left|\mathcal{T}_\ell\right| \le \left|\widetilde{\mathcal{T}}_\ell\right|-\left|\mathcal{T}_0\right| \le N(\ell).$$

The combination with (35) reads

$$\left|\mathcal{T}_\ell\backslash\widehat{\mathcal{T}}_\ell\right| \le 2\left|(u,f)\right|^{1/s}_{\mathcal{A}_s}\varepsilon(\ell)^{-1/s}. \tag{37}$$

Let $(\widehat{\boldsymbol{\sigma}}_\ell,\widehat{u}_\ell)\in\boldsymbol{\Sigma}(\widehat{\mathcal{T}}_\ell)\times\mathcal{A}(\widehat{\mathcal{T}}_\ell)$ solve the discrete equation (11) with respect to $\widehat{\mathcal{T}}_\ell$ and let $\widehat{g}_\ell$ denote the corresponding Dirichlet boundary data approximation. Lemmas 2.1 and 2.2 lead to the existence of some $\widehat{w}_\ell\in S^1(\widehat{\mathcal{T}}_\ell;\mathbb{R}^2)$ and $C_{\mathrm{osc}}\approx 1$ with

$$\widehat{w}_\ell\big|_\Gamma=\widehat{g}_\ell-\widetilde{g}_\ell \quad\text{and}\quad \left|\!\left|\!\left|\widehat{w}_\ell\right|\!\right|\!\right|^2 \le C_{\mathrm{osc}}\operatorname{osc}^2(\widehat{g}_\ell';\widetilde{\mathcal{E}}_\ell(\Gamma)) \le C_{\mathrm{osc}}\operatorname{osc}^2(g',\widetilde{\mathcal{E}}_\ell(\Gamma)).$$

The combination of this with $\widetilde{\boldsymbol{\sigma}}_\ell\in\boldsymbol{\Sigma}(\widehat{\mathcal{T}}_\ell)$, $\widetilde{u}_\ell+\widehat{w}_\ell\in\mathcal{A}(\widetilde{\mathcal{T}}_\ell)$, and the Young inequality implies

$$\begin{aligned} LS(f;\widehat{\boldsymbol{\sigma}}_\ell,\widehat{u}_\ell) &\le \left\|f+\operatorname{div}\widetilde{\boldsymbol{\sigma}}_\ell\right\|^2_{L^2(\Omega)}+\left\|\operatorname{dev}\widetilde{\boldsymbol{\sigma}}_\ell-\mathrm{D}(\widetilde{u}_\ell+\widehat{w}_\ell)\right\|^2_{L^2(\Omega)} \\ &\le \left\|f+\operatorname{div}\widetilde{\boldsymbol{\sigma}}_\ell\right\|^2_{L^2(\Omega)}+2\left\|\operatorname{dev}\widetilde{\boldsymbol{\sigma}}_\ell-\mathrm{D}\,\widetilde{u}_\ell\right\|^2_{L^2(\Omega)}+2\left|\!\left|\!\left|\widehat{w}_\ell\right|\!\right|\!\right|^2 \\ &\le 2LS(f;\widetilde{\boldsymbol{\sigma}}_\ell,\widetilde{u}_\ell)+2C_{\mathrm{osc}}\operatorname{osc}^2(g',\widetilde{\mathcal{E}}_\ell(\Gamma)). \end{aligned}$$

The boundary data oscillations are smaller on finer meshes, whence

$$\operatorname{osc}^2(g',\widehat{\mathcal{E}}_\ell(\Gamma)) \le \operatorname{osc}^2(g',\widetilde{\mathcal{E}}_\ell(\Gamma)).$$

The two previously displayed formulas and (36) imply, for $C_0 := \max\{2, 1+2C_{\mathrm{osc}}\}$, that

$$\begin{aligned} &LS(f;\widehat{\boldsymbol{\sigma}}_\ell,\widehat{u}_\ell)+\operatorname{osc}^2(g',\widehat{\mathcal{E}}_\ell(\Gamma)) \\ &\quad\le 2LS(f;\widetilde{\boldsymbol{\sigma}}_\ell,\widetilde{u}_\ell)+(1+2C_{\mathrm{osc}})\operatorname{osc}^2(g',\widetilde{\mathcal{E}}_\ell(\Gamma)) \le C_0\tau^2\xi_\ell^2. \end{aligned} \tag{38}$$

*Step 3 (Case A):* Lemmas 2.1 and 2.2 lead to the existence of some $\widehat{w}_\ell\in S^1(\widehat{\mathcal{T}}_\ell;\mathbb{R}^2)$ and $C_{\mathrm{osc}}\approx 1$ with

$$\widehat{w}_\ell\big|_\Gamma=\widehat{g}_\ell-g_\ell \quad\text{and}\quad \left|\!\left|\!\left|\widehat{w}_\ell\right|\!\right|\!\right|^2 \le C_{\mathrm{osc}}\operatorname{osc}^2(\widehat{g}_\ell',\mathcal{E}_\ell(\Gamma)). \tag{39}$$

The arguments for the proof of (25) apply literally to the situation at hand. For the exact solution $(\boldsymbol{\sigma}, u)$ to (9), this leads to

$$
\begin{aligned}
&LS(f; \boldsymbol{\sigma}_\ell, u_\ell) = LS(0; \widehat{\boldsymbol{\sigma}}_\ell - \boldsymbol{\sigma}_\ell, \widehat{u}_\ell - u_\ell) + LS(f; \widehat{\boldsymbol{\sigma}}_\ell, \widehat{u}_\ell)\\
&\quad +2\mathcal{B}(\boldsymbol{\sigma} - \widehat{\boldsymbol{\sigma}}_\ell, u - \widehat{u}_\ell; 0, \widehat{w}_\ell).
\end{aligned}
$$

The Cauchy-Schwarz inequality, the Young inequality, and (39) result in $C_1 := \max\{1, C_{\mathrm{osc}}\}$ and

$$
\begin{aligned}
LS(f; \boldsymbol{\sigma}_\ell, u_\ell) &\leq LS(0; \widehat{\boldsymbol{\sigma}}_\ell - \boldsymbol{\sigma}_\ell, \widehat{u}_\ell - u_\ell) + 2LS(f; \widehat{\boldsymbol{\sigma}}_\ell, \widehat{u}_\ell) + C_{\mathrm{osc}} \operatorname{osc}^2(\widehat{g}_\ell, \mathcal{E}_\ell(\Gamma))\\
&\leq C_1 \delta^2(\widehat{\mathcal{T}}_\ell, \mathcal{T}_\ell) + 2LS(f; \widehat{\boldsymbol{\sigma}}_\ell, \widehat{u}_\ell).
\end{aligned}
$$

The discrete reliability Theorem 4.3, the triangle inequality, and the Young inequality yield

$$
\begin{aligned}
\delta^2(\widehat{\mathcal{T}}_\ell, \mathcal{T}_\ell)/C_{\mathrm{drel}} &\leq \eta_\ell^2(\mathcal{T}_\ell \backslash \widehat{\mathcal{T}}_\ell) + \left\|(1 - \Pi_\ell) \operatorname{div} \widehat{\boldsymbol{\sigma}}_\ell\right\|_{L^2(\Omega)}^2\\
&\leq \eta_\ell^2(\mathcal{T}_\ell \backslash \widehat{\mathcal{T}}_\ell) + 2\left\|f - f_\ell\right\|_{L^2(\Omega)}^2 + 2\left\|(1 - \Pi_\ell)(f + \operatorname{div} \widehat{\boldsymbol{\sigma}}_\ell)\right\|_{L^2(\Omega)}^2.
\end{aligned}
$$

The orthogonality of the projection $\Pi_\ell$ and the definition of the least-squares functional imply

$$
\left\|(1 - \Pi_\ell)(f + \operatorname{div} \widehat{\boldsymbol{\sigma}}_\ell)\right\|_{L^2(\Omega)}^2 \leq \left\|f + \operatorname{div} \widehat{\boldsymbol{\sigma}}_\ell\right\|_{L^2(\Omega)}^2 \leq LS(f; \widehat{\boldsymbol{\sigma}}_\ell, \widehat{u}_\ell).
$$

Since $h_{\mathcal{E}}\big|_{\partial T \cap \Gamma} \approx |T|^{1/2}$, there exists a constant $C_2 \approx 1$ such that

$$
\begin{aligned}
\operatorname{osc}^2(g', \mathcal{E}_\ell(\Gamma)) &= \sum_{T \in \mathcal{T}_\ell \backslash \widehat{\mathcal{T}}_\ell} \left\|h_\ell^{1/2}(g' - g'_\ell)\right\|_{L^2(\partial T \cap \Gamma)}^2\\
&\quad + \sum_{T \in \mathcal{T}_\ell \cap \widehat{\mathcal{T}}_\ell} \left\|h_\ell^{1/2}(g' - g'_\ell)\right\|_{L^2(\partial T \cap \Gamma)}^2\\
&\leq C_2 \eta_\ell^2(\mathcal{T}_\ell \backslash \widehat{\mathcal{T}}_\ell) + \operatorname{osc}^2(g', \widehat{\mathcal{E}}_\ell(\Gamma)).
\end{aligned}
$$

For $C_3 := \max\{C_2 + C_1 C_{\mathrm{drel}}, 2(1 + C_1 C_{\mathrm{drel}})\}$, a combination of the four previously displayed formulas in Step 3 plus some rearrangements result in

$$
\begin{aligned}
\big(LS(f; \boldsymbol{\sigma}_\ell, u_\ell) + \operatorname{osc}^2(g', \mathcal{E}_\ell(\Gamma))\big)/C_3 &\leq \eta_\ell^2(\mathcal{T}_\ell \backslash \widehat{\mathcal{T}}_\ell) + \left\|f - f_\ell\right\|_{L^2(\Omega)}^2\\
&\quad + LS(f; \widehat{\boldsymbol{\sigma}}_\ell, \widehat{u}_\ell) + \operatorname{osc}^2(g', \widehat{\mathcal{E}}_\ell(\Gamma)).
\end{aligned}
$$

For $C_4 := C_3(\max\{1, \Lambda_{\text{osc}}\} + C_{\text{eff}} \max\{1, \Lambda_{\text{con}}\})$, this and the efficiency from Theorem 3.5 prove

$$\begin{aligned}\xi_\ell^2 &= LS(f; \boldsymbol{\sigma}_\ell, u_\ell) + \|f - f_\ell\|_{L^2(\Omega)}^2 + \Lambda_{\text{osc}} \operatorname{osc}^2(g', \mathcal{E}_\ell(\Gamma)) + \Lambda_{\text{con}} \eta_\ell^2 \\ &\le (\max\{1, \Lambda_{\text{osc}}\} + C_{\text{eff}} \max\{1, \Lambda_{\text{con}}\})\big(LS(f; \boldsymbol{\sigma}_\ell, u_\ell) + \operatorname{osc}^2(g', \mathcal{E}_\ell(\Gamma))\big) \\ &\le C_4\big(LS(f; \widehat{\boldsymbol{\sigma}}_\ell, \widehat{u}_\ell) + \operatorname{osc}^2(g', \widehat{\mathcal{E}}_\ell(\Gamma)) + \eta_\ell^2(\mathcal{T}_\ell \backslash \widehat{\mathcal{T}}_\ell) + \|f - f_\ell\|_{L^2(\Omega)}^2\big).\end{aligned}$$

Case A and the definition of $\xi_\ell$ imply

$$\|f - f_\ell\|_{L^2(\Omega)}^2 \le \kappa_0 \eta_\ell^2 \le \kappa_0 / \Lambda_{\text{con}}\, \xi_\ell^2.$$

The combination of the two previously displayed estimates with (38) results in

$$\xi_\ell^2 / C_4 \le \eta_\ell^2(\mathcal{T}_\ell \backslash \widehat{\mathcal{T}}_\ell) + (C_0 \tau^2 + \kappa_0 / \Lambda_{\text{con}}) \xi_\ell^2.$$

Every choice of $\tau^2 \le 1/(4C_0C_4)$ and $\kappa_0 \le \Lambda_{\text{con}}/(4C_4)$ leads to

$$\Lambda_{\text{con}} \eta_\ell^2 \le \xi_\ell^2 \le 2C_4 \eta_\ell^2(\mathcal{T}_\ell \backslash \widehat{\mathcal{T}}_\ell).$$

For $\theta \le \theta_0 := \Lambda_{\text{con}}/(2C_4)$, the Dörfler marking in Case A of the adaptive algorithm from Sect. 3.2 computes a subset $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ of (almost) minimal cardinality with

$$\theta \eta_\ell^2 \le \eta_\ell^2(\mathcal{M}_\ell) \quad \text{and} \quad |\mathcal{M}_\ell| \lesssim |\mathcal{T}_\ell \backslash \widehat{\mathcal{T}}_\ell|.$$

The estimate (37) implies

$$|\mathcal{M}_\ell|^{2s} \lesssim |\mathcal{T}_\ell \backslash \widehat{\mathcal{T}}_\ell|^{2s} \lesssim |(u, f)|_{\mathcal{A}_s}^2 \varepsilon(\ell)^{-2}. \tag{40}$$

*Step 4 (Case B):* Let $\mathcal{T} \in \mathbb{T}$ be the output of the TSA from [7, Section 5] with tolerance $\text{Tol} := \rho^{1/2} \|f - f_\ell\|_{L^2(\Omega)}$ such that

$$\|f - f_{\mathcal{T}}\|_{L^2(\Omega)} \le \text{Tol} \quad \text{and} \quad |\mathcal{T}| - |\mathcal{T}_0| \lesssim |(u, f)|_{\mathcal{A}_s}^{1/s} \text{Tol}^{-1/s}. \tag{41}$$

Let $\mathcal{T}_{\ell+1} := \mathcal{T} \otimes \mathcal{T}_\ell$ denote the overlay of $\mathcal{T}$ and the triangulation $\mathcal{T}_\ell$ on level $\ell$. The embed oscillation control algorithm from [22, Section 3] ensures the existence of a finite sequence of disjoint sets $\mathcal{M}_\ell^{(0)}, \mathcal{M}_\ell^{(1)}, \dots, \mathcal{M}_\ell^{(K(\ell))}$ of marked triangles realizing a successive one-level refinement with $\mathcal{T}_\ell^{(0)} := \mathcal{T}_\ell$,

$$\mathcal{T}_\ell^{(k+1)} = \text{Refine}(\mathcal{T}_\ell^{(k)}, \mathcal{M}_\ell^{(k)}) \quad \text{for} \quad k = 0, \dots, K(\ell),$$

and $\mathcal{T}_\ell^{(K(\ell)+1)} = \mathcal{T}_{\ell+1}$. The disjoint union of those sets $\mathcal{M}_\ell := \mathcal{M}_\ell^{(0)} \cup \cdots \cup \mathcal{M}_\ell^{(K(\ell))}$ satisfies [22, Theorem 3.3]

$$|\mathcal{M}_\ell| = \sum_{k=0}^{K(\ell)} |\mathcal{M}_\ell^{(k)}| \leq |\mathcal{T}| - |\mathcal{T}_0|.$$

The combination with (41) shows

$$|\mathcal{M}_\ell|^{2s} \lesssim |(u, f)|^2_{\mathcal{A}_s} \mathrm{Tol}^{-2}. \tag{42}$$

Recall that $\mathrm{osc}^2(g', \mathcal{E}_\ell(\Gamma)) \leq \eta_\ell^2 < \|f - f_\ell\|^2_{L^2(\Omega)}/\kappa$ in Case B. This and the reliability from Corollary 4.4 imply

$$\begin{aligned}\xi_\ell^2 &= LS(f; \boldsymbol{\sigma}_\ell, u_\ell) + \|f - f_\ell\|^2_{L^2(\Omega)} + \Lambda_{\mathrm{osc}}\,\mathrm{osc}^2(g', \mathcal{E}_\ell(\Gamma)) + \Lambda_{\mathrm{con}}\eta_\ell^2\\ &\lesssim \eta_\ell^2 + \|f - f_\ell\|^2_{L^2(\Omega)} \lesssim \|f - f_\ell\|^2_{L^2(\Omega)} \lesssim \mathrm{Tol}^2.\end{aligned}$$

The combination with (42) reads

$$|\mathcal{M}_\ell|^{2s} \lesssim |(u, f)|^2_{\mathcal{A}_s} \xi_\ell^{-2}.$$

This concludes the proof of an estimate like (40) in Case B.

*Step 5 (Finish of the proof):* As in [20], the overhead control of [6, Theorem 2.4] or [27], the contraction property from Theorem 4.10, and (40) lead to

$$|\mathcal{T}_\ell| - |\mathcal{T}_0| \lesssim |(u, f)|^{1/s}_{\mathcal{A}_s} \xi_\ell^{-1/s} \sum_{k=0}^{\ell-1} \rho_{\mathrm{con}}^{(\ell-k)/(2s)} \lesssim |(u, f)|^{1/s}_{\mathcal{A}_s} \xi_\ell^{-1/s}.$$

This and $\big(LS(f; \boldsymbol{\sigma}_\ell, u_\ell) + \mathrm{osc}^2(g', \mathcal{E}_\ell(\Gamma))\big)^{1/2} \leq \xi_\ell$ conclude the proof of

$$\big(|\mathcal{T}_\ell| - |\mathcal{T}_0|\big)^{2s} \big(LS(f; \boldsymbol{\sigma}_\ell, u_\ell) + \mathrm{osc}^2(g', \mathcal{E}_\ell(\Gamma))\big) \lesssim |(u, f)|^2_{\mathcal{A}_s} \quad \text{for any } \ell \in \mathbb{N}_0.$$

□

**Acknowledgments** This work was supported by Deutsche Forschungsgemeinschaft (DFG) SPP 1748.

## Appendix: Proofs of Lemma 4.5–4.7

*Proof* (Proof of Lemma 4.5) For $g$ replaced by $\widehat{I}g \in S^1(\widehat{\mathcal{E}}(\Gamma); \mathbb{R}^2)$, Lemma 2.1 guarantees the existence of some $\widehat{z} \in S^1(\widehat{\mathcal{T}}; \mathbb{R}^2)$ with

$$\widehat{z}|_\Gamma = (\widehat{I} - I)g \quad \text{and} \quad \|\!|\widehat{z}|\!\| \lesssim \mathrm{osc}(\widehat{\Pi}g', \mathcal{E}(\Gamma)).$$

The discrete equation (11) with respect to $\widehat{\mathcal{T}}$ with test functions $\boldsymbol{\tau}_{\mathrm{LS}} = \widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}} \in \boldsymbol{\Sigma}(\widehat{\mathcal{T}})$ and $v_{\mathrm{LS}} = \widehat{u}_{\mathrm{LS}} - u_{\mathrm{LS}} - \widehat{z} \in S^1_0(\widehat{\mathcal{T}};\mathbb{R}^2)$ imply that

$$
\begin{aligned}
&-\int_\Omega (\operatorname{dev}\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \mathrm{D}\,\widehat{u}_{\mathrm{LS}}) : (\operatorname{dev}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}) - \mathrm{D}(\widehat{u}_{\mathrm{LS}} - u_{\mathrm{LS}} - \widehat{z}))\mathrm{d}x\\
&\quad = \int_\Omega (f_{\widehat{\mathcal{T}}} + \operatorname{div}\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}) \cdot \operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}})\mathrm{d}x\\
&\quad = \int_\Omega (f_{\widehat{\mathcal{T}}} - f_{\mathcal{T}}) \cdot \operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}})\mathrm{d}x\\
&\qquad + \int_\Omega (f_{\mathcal{T}} + \operatorname{div}\boldsymbol{\sigma}_{\mathrm{LS}}) \cdot \operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}})\mathrm{d}x + \|\operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}})\|^2_{L^2(\Omega)}.
\end{aligned}
$$

The discrete equation (11) with respect to $\boldsymbol{\tau}_{\mathrm{LS}} = \boldsymbol{\tau}_{\mathrm{PS}}$ and the triangulation $\mathcal{T}$ and (21) result in

$$
\begin{aligned}
&-\int_\Omega (\operatorname{dev}\boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\,u_{\mathrm{LS}}) : \operatorname{dev}\boldsymbol{\tau}_{\mathrm{PS}}\mathrm{d}x\\
&\quad = \int_\Omega (f_{\mathcal{T}} + \operatorname{div}\boldsymbol{\sigma}_{\mathrm{LS}}) \cdot \operatorname{div}\boldsymbol{\tau}_{\mathrm{PS}}\mathrm{d}x = \int_\Omega (f_{\mathcal{T}} + \operatorname{div}\boldsymbol{\sigma}_{\mathrm{LS}}) \cdot \operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}})\mathrm{d}x.
\end{aligned}
$$

The combination of the preceding two displayed formulas proves

$$
\begin{aligned}
&\|\operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}})\|^2_{L^2(\Omega)}\\
&\quad = -\int_\Omega (f_{\widehat{\mathcal{T}}} - f_{\mathcal{T}}) \cdot \operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}})\mathrm{d}x + \int_\Omega (\operatorname{dev}\boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\,u_{\mathrm{LS}}) : \operatorname{dev}\boldsymbol{\tau}_{\mathrm{PS}}\mathrm{d}x\\
&\qquad - \int_\Omega (\operatorname{dev}\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \mathrm{D}\,\widehat{u}_{\mathrm{LS}}) : \big(\operatorname{dev}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}) - \mathrm{D}(\widehat{u}_{\mathrm{LS}} - u_{\mathrm{LS}} - \widehat{z})\big)\mathrm{d}x.
\end{aligned}
$$

This plus some elementary algebra leads to

$$
\begin{aligned}
&\|\operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}})\|^2_{L^2(\Omega)} + \|\operatorname{dev}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}) - \mathrm{D}(\widehat{u}_{\mathrm{LS}} - u_{\mathrm{LS}})\|^2_{L^2(\Omega)}\\
&\quad = -\int_\Omega (f_{\widehat{\mathcal{T}}} - f_{\mathcal{T}}) \cdot \operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}})\mathrm{d}x\\
&\qquad - \int_\Omega (\operatorname{dev}\boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\,u_{\mathrm{LS}}) : \big(\operatorname{dev}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\tau}_{\mathrm{PS}}) - \mathrm{D}(\widehat{u}_{\mathrm{LS}} - u_{\mathrm{LS}} - \widehat{z})\big)\mathrm{d}x\\
&\qquad - \int_\Omega (\operatorname{dev}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}) - \mathrm{D}(\widehat{u}_{\mathrm{LS}} - u_{\mathrm{LS}})) : \mathrm{D}\,\widehat{z}\mathrm{d}x.
\end{aligned}
\tag{43}
$$

The remaining analysis in this proof concerns the split

$$
\operatorname{dev}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\tau}_{\mathrm{PS}}) = \operatorname{dev}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}} - \widehat{\boldsymbol{\tau}}_{\mathrm{PS}} + \widehat{\boldsymbol{\tau}}^*_{\mathrm{PS}} - \boldsymbol{\tau}_{\mathrm{PS}}) + \operatorname{dev}(\widehat{\boldsymbol{\tau}}_{\mathrm{PS}} - \widehat{\boldsymbol{\tau}}^*_{\mathrm{PS}}).
$$

The Eq. (21) imply that the Raviart-Thomas function

$$
\widehat{\boldsymbol{\rho}} := \widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}} - \widehat{\boldsymbol{\tau}}_{\mathrm{PS}} + \widehat{\boldsymbol{\tau}}^*_{\mathrm{PS}} - \boldsymbol{\tau}_{\mathrm{PS}} \tag{44}
$$

is divergence-free and so piecewise constant. The Helmholtz decomposition (14),

$$\operatorname{dev}\widehat{\boldsymbol{\rho}} = \mathrm{D}_{\mathrm{NC}}\,\widehat{\alpha} + \operatorname{dev}\operatorname{Curl}\widehat{\beta} \in P_0(\widehat{\mathcal{T}};\mathbb{R}^{2\times 2}_{\mathrm{dev}}),$$

leads to some $\widehat{\alpha} \in Z(\widehat{\mathcal{T}})$ and $\widehat{\beta} \in X(\widehat{\mathcal{T}})$. The orthogonality and a piecewise integration by parts show

$$|||\widehat{\alpha}|||^2_{\mathrm{NC}} = \int_\Omega \operatorname{dev}\widehat{\boldsymbol{\rho}} : \mathrm{D}_{\mathrm{NC}}\,\widehat{\alpha}\mathrm{d}x = \int_\Omega \widehat{\boldsymbol{\rho}} : \mathrm{D}_{\mathrm{NC}}\,\widehat{\alpha}\mathrm{d}x = \sum_{E\in\widehat{\mathcal{E}}(\Omega)} \int_E [\widehat{\boldsymbol{\rho}}\nu_E \cdot \widehat{\alpha}]_E\,\mathrm{d}s.$$

Recall that the Raviart-Thomas function $\widehat{\boldsymbol{\rho}}$ is continuous in its normal components and that $\widehat{\boldsymbol{\rho}}\nu_E$ is constant along $E \in \widehat{\mathcal{E}}(\Omega)$. Since the jump $[\widehat{\alpha}]_E$ of $E$ has integral mean zero along $E$,

$$\int_E [\widehat{\boldsymbol{\rho}}\nu_E \cdot \widehat{\alpha}]_E\,\mathrm{d}s = 0.$$

Hence, $\widehat{\alpha} \equiv 0$ and the Helmholtz decomposition reduces to

$$\operatorname{dev}\widehat{\boldsymbol{\rho}} = \operatorname{dev}\operatorname{Curl}\widehat{\beta} \tag{45}$$

for the divergence-free test function $\operatorname{Curl}\widehat{\beta} \in RT_0(\widehat{\mathcal{T}};\mathbb{R}^{2\times 2})$. One term in (43) reads

$$\begin{aligned}
&-\int_\Omega (\operatorname{dev}\boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\,u_{\mathrm{LS}}) : \operatorname{dev}(\widehat{\boldsymbol{\tau}}_{\mathrm{PS}} - \widehat{\boldsymbol{\tau}}^*_{\mathrm{PS}})\mathrm{d}x \\
&\quad= \int_\Omega (\operatorname{dev}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}) - \mathrm{D}(\widehat{u}_{\mathrm{LS}} - u_{\mathrm{LS}})) : \operatorname{dev}(\widehat{\boldsymbol{\tau}}_{\mathrm{PS}} - \widehat{\boldsymbol{\tau}}^*_{\mathrm{PS}})\mathrm{d}x \\
&\qquad - \int_\Omega (\operatorname{dev}\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \mathrm{D}\,\widehat{u}_{\mathrm{LS}}) : \operatorname{dev}(\widehat{\boldsymbol{\tau}}_{\mathrm{PS}} - \widehat{\boldsymbol{\tau}}^*_{\mathrm{PS}})\mathrm{d}x.
\end{aligned} \tag{46}$$

The discrete equation (11) with respect to the triangulation $\widehat{\mathcal{T}}$, $\boldsymbol{\tau}_{\mathrm{LS}} = \widehat{\boldsymbol{\tau}}_{\mathrm{PS}} - \widehat{\boldsymbol{\tau}}^*_{\mathrm{PS}}$, and $v_{\mathrm{LS}} \equiv 0$, and the combination with (21) plus elementary algebra with the $L^2$-projection $\Pi$ onto $P_0(\mathcal{T};\mathbb{R}^{2\times 2})$ show

$$\begin{aligned}
&-\int_\Omega (\operatorname{dev}\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \mathrm{D}\,\widehat{u}_{\mathrm{LS}}) : \operatorname{dev}(\widehat{\boldsymbol{\tau}}_{\mathrm{PS}} - \widehat{\boldsymbol{\tau}}^*_{\mathrm{PS}})\mathrm{d}x \\
&= \int_\Omega (f_{\widehat{\mathcal{T}}} + \operatorname{div}\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}) \cdot \operatorname{div}(\widehat{\boldsymbol{\tau}}_{\mathrm{PS}} - \widehat{\boldsymbol{\tau}}^*_{\mathrm{PS}})\mathrm{d}x \\
&= \int_\Omega (f_{\widehat{\mathcal{T}}} + \operatorname{div}\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}) \cdot \big((1-\Pi)\operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}})\big)\mathrm{d}x \\
&= \int_\Omega (f_{\widehat{\mathcal{T}}} - f_{\mathcal{T}}) \cdot \operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}})\mathrm{d}x + \big\|(1-\Pi)\operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}})\big\|^2_{L^2(\Omega)}.
\end{aligned} \tag{47}$$

The combination of (43)–(47) concludes the proof. □

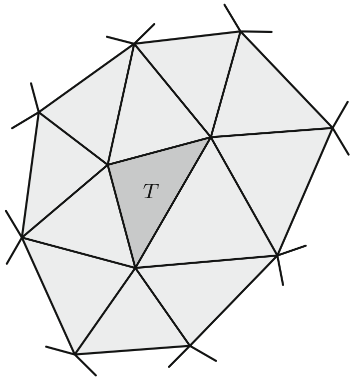


**Fig. 3** Enlarged triangle patch $\Omega_T$

*Proof* (Proof of Lemma 4.6) Let $v \in S_0^1(\mathcal{T};\mathbb{R}^2)$ be the Scott-Zhang quasi-interpolation of $\widehat{v} := \widehat{u}_{\mathrm{LS}} - u_{\mathrm{LS}} - \widehat{z} \in S_0^1(\widehat{\mathcal{T}};\mathbb{R}^2)$. For every $z \in \mathcal{N}$ in the construction of the quasi-interpolation [26, Section 2], choose $E \in \mathcal{E}(\omega_z)$ such that $E \in \mathcal{E} \cap \widehat{\mathcal{E}}$, whenever possible. This ensures that the error function $\widehat{w} := \widehat{v} - v \in S_0^1(\widehat{\mathcal{T}};\mathbb{R}^2)$ of the quasi-interpolation vanishes on any $T \in \mathcal{T} \cap \widehat{\mathcal{T}}$. The first-order approximation property [26, equation (4.3)] and the stability property [26, Theorem3.1] read

$$|T|^{-1/2} \|\widehat{w}\|_{L^2(T)} + \|\mathrm{D}\,\widehat{w}\|_{L^2(T)} \lesssim \|\mathrm{D}\,\widehat{v}\|_{L^2(\Omega_T)} \tag{48}$$

for the enlarged triangle patch $\Omega_T := \bigcup_{z \in \mathcal{N}(T)} \omega_z$ on the triangulation $\mathcal{T}$ of Fig. 3.

Since $v \in S_0^1(\mathcal{T};\mathbb{R}^2)$ is an admissible test function, (11) implies

$$\int_\Omega (\operatorname{dev} \boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\,u_{\mathrm{LS}}) : \mathrm{D}\,v \mathrm{d}x = 0.$$

This, the definition of $\widehat{w}$, and a piecewise integration by parts result in

$$\begin{aligned}\int_\Omega (\operatorname{dev} \boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\,u_{\mathrm{LS}}) : \mathrm{D}\,\widehat{v} \mathrm{d}x &= \int_\Omega (\operatorname{dev} \boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\,u_{\mathrm{LS}}) : \mathrm{D}\,\widehat{w} \mathrm{d}x \\ = - \sum_{T \in \mathcal{T}\setminus\widehat{\mathcal{T}}} \Bigg( \int_T & \widehat{w} \cdot \operatorname{div} \operatorname{dev} \boldsymbol{\sigma}_{\mathrm{LS}} \mathrm{d}x \\ + \sum_{E \in \mathcal{E}(T) \cap \mathcal{E}(\Omega)} & \int_E \widehat{w} \cdot \big( [\operatorname{dev} \boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\,u_{\mathrm{LS}}]_E\, \nu_E \big) \,\mathrm{d}s \Bigg).\end{aligned} \tag{49}$$

Given any $T \in \mathcal{T}\setminus\widehat{\mathcal{T}}$, a Cauchy-Schwarz inequality plus (48) prove

$$\begin{aligned}\int_T \widehat{w} \cdot \operatorname{div} \operatorname{dev} \boldsymbol{\sigma}_{\mathrm{LS}} \mathrm{d}x &\le |T|^{-1/2} \|\widehat{w}\|_{L^2(T)} |T|^{1/2} \|\operatorname{div} \operatorname{dev} \boldsymbol{\sigma}_{\mathrm{LS}}\|_{L^2(T)} \\ &\lesssim \|\mathrm{D}\,\widehat{v}\|_{L^2(\Omega_T)} |T|^{1/2} \|\operatorname{div} \operatorname{dev} \boldsymbol{\sigma}_{\mathrm{LS}}\|_{L^2(T)}.\end{aligned}$$

Given any $T \in \mathcal{T} \backslash \widehat{\mathcal{T}}$ with $E \in \mathcal{E}(T)$, a combination of a Cauchy-Schwarz inequality, a trace inequality, and (48) result in

$$\begin{aligned}
&\int_E \widehat{w} \cdot \big([\operatorname{dev} \boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\, u_{\mathrm{LS}}]_E\, \nu_E\big)\, \mathrm{d}s \\
&\le |T|^{-1/4} \|\widehat{w}\|_{L^2(E)} |T|^{1/4} \|[\operatorname{dev} \boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\, u_{\mathrm{LS}}]_E\, \nu_E\|_{L^2(E)} \\
&\lesssim \big(|T|^{-1/2} \|\widehat{w}\|_{L^2(T)} + \|\mathrm{D}\, \widehat{w}\|_{L^2(T)}\big) |T|^{1/4} \|[\operatorname{dev} \boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\, u_{\mathrm{LS}}]_E\, \nu_E\|_{L^2(E)} \\
&\lesssim \|\mathrm{D}\, \widehat{v}\|_{L^2(\Omega_T)} |T|^{1/4} \|[\operatorname{dev} \boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\, u_{\mathrm{LS}}]_E\, \nu_E\|_{L^2(E)}.
\end{aligned}$$

The combination of (49) with the last two preceding estimates and a finite overlap of the patches $\Omega_T$ from Fig. 3 conclude the proof. □

*Proof* (Proof of Lemma 4.7) Let $\beta \in S^1(\mathcal{T}; \mathbb{R}^2)$ be the Scott-Zhang quasi-interpolation of $\widehat{\beta} \in S^1(\widehat{\mathcal{T}}; \mathbb{R}^2)$. For every $z \in \mathcal{N}$ in the design of the quasi-interpolation [26, Section 2], choose $E \in \mathcal{E}(\omega_z)$ such that $E \in \mathcal{E} \cap \widehat{\mathcal{E}}$, whenever possible, so that the error function $\widehat{\gamma} := \widehat{\beta} - \beta \in S^1(\widehat{\mathcal{T}}; \mathbb{R}^2)$ of the quasi-interpolation vanishes on any $T \in \mathcal{T} \cap \widehat{\mathcal{T}}$. The first-order approximation property [26, equation (4.3)] and the stability property [26, Theorem 3.1] read, with the enlarged triangle patch $\Omega_T$ of Fig. 3, as

$$|T|^{-1/2} \|\widehat{\gamma}\|_{L^2(T)} + \|\mathrm{D}\, \widehat{\gamma}\|_{L^2(T)} \lesssim \|\mathrm{D}\, \widehat{\beta}\|_{L^2(\Omega_T)}. \tag{50}$$

For $x = (x_1, x_2) \in \overline{\Omega}$, define a modified quasi-interpolation $\widetilde{\beta} \in S^1(\mathcal{T}; \mathbb{R}^2)$ by

$$\widetilde{\beta}(x) := \beta(x) - c/2\, (-x_2, x_1)^\top \quad \text{with} \quad c := \int_\Omega \operatorname{tr} \operatorname{Curl} \beta \, \mathrm{d}x \Big/ |\Omega|.$$

This guarantees that $\operatorname{Curl} \widetilde{\beta} = \operatorname{Curl} \beta - c/2\, I_{2\times 2} \in \boldsymbol{\Sigma}(\mathcal{T})$ is an admissible and divergence-free test function. Therefore, the discrete equation (11) proves

$$\begin{aligned}
&\int_\Omega (\operatorname{dev} \boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\, u_{\mathrm{LS}}) : \operatorname{dev} \operatorname{Curl} \beta \, \mathrm{d}x \\
&\quad = \int_\Omega (\operatorname{dev} \boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\, u_{\mathrm{LS}}) : \operatorname{dev} \operatorname{Curl} \widetilde{\beta} \, \mathrm{d}x = 0.
\end{aligned}$$

This plus elementary algebra on the deviatoric part and a piecewise integration by parts imply

$$\begin{aligned}
&\int_\Omega (\operatorname{dev} \boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\, u_{\mathrm{LS}}) : \operatorname{dev} \operatorname{Curl} \widehat{\beta} \, \mathrm{d}x \qquad (51) \\
&= \int_\Omega (\operatorname{dev} \boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\, u_{\mathrm{LS}}) : \operatorname{dev} \operatorname{Curl} \widehat{\gamma} \, \mathrm{d}x = \int_\Omega \operatorname{dev}(\boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\, u_{\mathrm{LS}}) : \operatorname{Curl} \widehat{\gamma} \, \mathrm{d}x
\end{aligned}$$

$$= - \sum_{T\in\mathcal{T}\setminus\widehat{\mathcal{T}}} \Big( \int_T \widehat{\gamma}\cdot \operatorname{curl}\operatorname{dev}(\boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\, u_{\mathrm{LS}})\mathrm{d}x$$
$$+ \sum_{E\in\mathcal{E}(T)} \int_E \widehat{\gamma}\cdot \big([\operatorname{dev}(\boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\, u_{\mathrm{LS}})]_E\, \tau_E\big)\, \mathrm{d}s\Big).$$

Given any $T \in \mathcal{T}\setminus\widehat{\mathcal{T}}$, a Cauchy-Schwarz inequality plus (50) prove

$$\int_T \widehat{\gamma}\cdot \operatorname{curl}\operatorname{dev}\boldsymbol{\sigma}_{\mathrm{LS}}\mathrm{d}x \lesssim \|\mathrm{D}\,\widehat{\beta}\|_{L^2(\Omega_T)} |T|^{1/2} \|\operatorname{curl}\operatorname{dev}\boldsymbol{\sigma}_{\mathrm{LS}}\|_{L^2(T)}.$$

Given any $T \in \mathcal{T}\setminus\widehat{\mathcal{T}}$ with $E \in \mathcal{E}(T)$, a combination of a Cauchy-Schwarz inequality, a trace inequality, and (50) imply

$$\begin{aligned}
&\int_E \widehat{\gamma}\cdot ([\operatorname{dev}(\boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\, u_{\mathrm{LS}})]_E\, \tau_E)\, \mathrm{d}s\\
&\quad\le |T|^{-1/4} \|\widehat{\gamma}\|_{L^2(E)} |T|^{1/4} \|[\operatorname{dev}(\boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\, u_{\mathrm{LS}})]_E\, \tau_E\|_{L^2(E)}\\
&\quad\lesssim \|\mathrm{D}\,\widehat{\beta}\|_{L^2(\Omega_T)} |T|^{1/4} \|[\operatorname{dev}(\boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\, u_{\mathrm{LS}})]_E\, \tau_E\|_{L^2(E)}.
\end{aligned}$$

The combination of (51) with the last two preceding estimates and a finite overlap of the patches $\Omega_T$ from Fig. 3 prove

$$\begin{aligned}
\int_\Omega (\operatorname{dev}\boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\, u_{\mathrm{LS}}) : \operatorname{dev}\operatorname{Curl}\widehat{\beta}\mathrm{d}x \lesssim |||\widehat{\beta}||| \Big( \sum_{T\in\mathcal{T}\setminus\widehat{\mathcal{T}}} \Big( |T|\, \|\operatorname{curl}\operatorname{dev}\boldsymbol{\sigma}_{\mathrm{LS}}\|^2_{L^2(T)}\\
+ \sum_{E\in\mathcal{E}(T)} |T|^{1/2} \|[\operatorname{dev}(\boldsymbol{\sigma}_{\mathrm{LS}} - \mathrm{D}\, u_{\mathrm{LS}})]_E\, \tau_E\|^2_{L^2(E)}\Big)\Big)^{1/2}.
\end{aligned} \tag{52}$$

The subsequent stability property can be found in [13, Lemma 3.4] in different notation

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

The stability of PS-FEM applied to $\widehat{\boldsymbol{\tau}}_{\mathrm{PS}} - \widehat{\boldsymbol{\tau}}^*_{\mathrm{PS}}$ and $\boldsymbol{\tau}_{\mathrm{PS}}$ yields

$$\begin{aligned}
\|\widehat{\boldsymbol{\tau}}_{\mathrm{PS}} - \widehat{\boldsymbol{\tau}}^*_{\mathrm{PS}}\|_{L^2(\Omega)} &\lesssim \|(1-\Pi)\operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}})\|_{L^2(\Omega)} \quad \text{and}\\
\|\boldsymbol{\tau}_{\mathrm{PS}}\|_{L^2(\Omega)} &\lesssim \|\Pi\operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}})\|_{L^2(\Omega)}.
\end{aligned} \tag{53}$$

Since $\widehat{\beta} \in X(\widehat{\mathcal{T}})$, $\operatorname{Curl} \widehat{\beta} \in \boldsymbol{\Sigma}(\widehat{\mathcal{T}})$. Hence, the tr-dev-div lemma (15), Eq. (45), definition (44), a triangle inequality, and (53) imply

$$
\begin{aligned}
\|\!|\widehat{\beta}|\!\| &= \| \operatorname{Curl} \widehat{\beta} \|_{L^2(\Omega)} \lesssim \| \operatorname{dev} \operatorname{Curl} \widehat{\beta} \|_{L^2(\Omega)} = \| \operatorname{dev} \widehat{\boldsymbol{\rho}} \|_{L^2(\Omega)} \leq \| \widehat{\boldsymbol{\rho}} \|_{L^2(\Omega)} \\
&\leq \| \widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}} \|_{L^2(\Omega)} + \| \widehat{\boldsymbol{\tau}}_{\mathrm{PS}} - \widehat{\boldsymbol{\tau}}_{\mathrm{PS}}^{*} \|_{L^2(\Omega)} + \| \boldsymbol{\tau}_{\mathrm{PS}} \|_{L^2(\Omega)} \\
&\lesssim \| \widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}} \|_{L^2(\Omega)} + \| (1 - \Pi) \operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}) \|_{L^2(\Omega)} \\
&\qquad + \| \Pi \operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}) \|_{L^2(\Omega)} \\
&\lesssim \| \widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}} \|_{H(\operatorname{div},\Omega)}.
\end{aligned}
$$

This and (52) conclude the proof. □

## References


1. Adler, J.H., Manteuffel, T.A., McCormick, S.F., Nolting, J.W., Ruge, J.W., Tang, L.: Efficiency based adaptive local refinement for first-order system least-squares formulations. SIAM J. Sci. Comput. **33**(1), 1–24 (2011)
2. Aurada, M., Feischl, M., Kemetmüller, J., Page, M., Praetorius, D.: 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), 1207–1235 (2013)
3. Bartels, S., Carstensen, C., Dolzmann, G.: Inhomogeneous Dirichlet conditions in a priori and a posteriori finite element error analysis. Numer. Math. **99**(1), 1–24 (2004)
4. Becker, R., Mao, S.: Quasi-optimality of adaptive nonconforming finite element methods for the Stokes equations. SIAM J. Numer. Anal. **49**(3), 970–991 (2011)
5. Berndt, M., Manteuffel, T.A., McCormick, S.F.: Local error estimates and adaptive refinement for first-order system least squares (FOSLS). Electron. Trans. Numer. Anal. **6**, 35–43 (1997, (electronic) [Special issue on multilevel methods (Copper Mountain, CO, 1997)]
6. Binev, P., Dahmen, W., DeVore, R.: Adaptive finite element methods with convergence rates. Numer. Math. **97**(2), 219–268 (2004)
7. Binev, P., DeVore, R.: Fast computation in adaptive tree approximation. Numer. Math. **97**(2), 193–217 (2004)
8. Bochev, P.B., Gunzburger, M.D.: Least-squares finite element methods, Applied mathematical sciences, vol. 166. Springer, New York (2009)
9. Boffi, D., Brezzi, F., Fortin, M.: Mixed finite element methods and applications, Springer series in computational Mathematics, vol. 44. Springer, Heidelberg (2013)
10. Braess, D.: Finite elements, 3rd edn. Theory, fast solvers, and applications in elasticity theory, Translated from the German by Larry L. Schumaker. MR 2322235 (2008b:65142). Cambridge University Press, Cambridge (2007)
11. Brenner, S.C., Scott, L.R.: The mathematical theory of finite element methods, Texts in applied mathematics, vol. 15, 3rd edn. Springer, New York (2008)
12. Cai, Z., Lee, B., Wang, P.: Least-squares methods for incompressible Newtonian fluid flow: linear stationary problems. SIAM J. Numer. Anal. 4**2**(2), 843–859 (2004, electronic)
13. Cai, Z., Tong, C., Vassilevski, P.S., Wang, C.: Mixed finite element methods for incompressible flow: stationary Stokes equations. Numer. Methods Partial Differ. Equ. **26**(4), 957–978 (2010)
14. Cai, Z., Wang, Y.: A multigrid method for the pseudostress formulation of Stokes problems. SIAM J. Sci. Comput. **29**(5), 2078–2095 (2007, electronic)
15. Carstensen, C., Gallistl, D., Schedensack, M.: $L^2$ best-approximation of the elastic stress in the Arnold-Winther FEM. IMA J. Numer. Anal. (2015, published online). doi:10.1093/imanum/drv051
16. Carstensen, C.: An a posteriori error estimate for a first-kind integral equation. Math. Comput. **66**(217), 139–155 (1997)
17. Carstensen, C., Feischl, M., Page, M., Praetorius, D.: Axioms of adaptivity. Comput. Math. Appl. **67**(6), 1195–1253 (2014)

18. Carstensen, C., Gallistl, D., Schedensack, M.: Quasi-optimal adaptive pseudostress approximation of the stokes equations. SIAM J. Numer. Anal. **51**(3), 1715–1734 (2013)
19. Carstensen, C., Kim, D., Park, E.-J.: A priori and a posteriori pseudostress-velocity mixed finite element error analysis for the Stokes problem. SIAM J. Numer. Anal. **49**(6), 2501–2523 (2011)
20. Carstensen, C., Park, E.-J.: Convergence and optimality of adaptive least squares finite element methods. SIAM J. Numer. Anal. **53**(1), 43–62 (2015)
21. Carstensen, C., Peterseim, D., Rabus, H.: Optimal adaptive nonconforming FEM for the Stokes problem. Numer. Math. **123**(2), 291–308 (2013)
22. Carstensen, C., Rabus, H.: An optimal adaptive mixed finite element method. Math. Comput. **80**(274), 649–667 (2011)
23. Cascon, J.M., Kreuzer, C., 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)
24. Feischl, M., Page, M., Praetorius, D.: Convergence and quasi-optimality of adaptive FEM with inhomogeneous Dirichlet data. J. Comput. Appl. Math. **255**, 481–501 (2014)
25. Jun, H., Jinchao, X.: Convergence and optimality of the adaptive nonconforming linear element method for the Stokes problem. J. Sci. Comput. **55**(1), 125–148 (2013)
26. Scott, L.R., Zhang, S.: Finite element interpolation of nonsmooth functions satisfying boundary conditions. Math. Comput. **54**(190), 483–493 (1990)
27. Stevenson, R.: Optimality of a standard adaptive finite element method. Found. Comput. Math. **7**(2), 245–269 (2007)
28. Stevenson, R.: The completion of locally refined simplicial partitions created by bisection. Math. Comput. **77**(261), 227–241 (2008, electronic)
29. Verfürth, R.: A posteriori error estimation techniques for finite element methods, Numerical mathematics and scientific computation. Oxford University Press, Oxford (2013)
30. Zhong, L., Chen, L., Shu, S., Wittum, G., Jinchao, X.: Convergence and optimality of adaptive edge finite element methods for time-harmonic Maxwell equations. Math. Comput. **81**(278), 623–642 (2012)