SIAM J. NUMER. ANAL.
Vol. 56, No. 1, pp. 428–447

AN ADAPTIVE LEAST-SQUARES FEM FOR LINEAR ELASTICITY WITH OPTIMAL CONVERGENCE RATES*

P. BRINGMANN†, C. CARSTENSEN†, AND G. STARKE‡

Abstract. Adaptive mesh-refining is of particular importance in computational mechanics and established here for the lowest-order locking-free least-squares finite element scheme which solely employs conforming P_1 approximations for the displacement and lowest-order Raviart–Thomas approximations for the stress variables. This forms a competitive discretization in particular in three-dimensional linear elasticity with traction boundary conditions although the stress approximation does not satisfy the symmetry condition exactly. The paper introduces an adaptive mesh-refining algorithm based on separate marking and exact solve with some novel explicit a posteriori error estimator and proves optimal convergence rates. The point is robustness in the sense that the crucial input parameters Θ for the Dörfler marking and κ for the separate marking as well as the equivalence constants in the asymptotic convergence rates do *not* degenerate as the Lamé parameter λ tends to ∞.

Key words. least-squares finite element method, linear, elasticity, adaptive finite element method, optimal convergence rates, separate marking, supercloseness

AMS subject classifications. 65N12, 65N30, 65N50, 65Y20, 74B05

DOI. 10.1137/16M1083797

1. Introduction.

Quasi-optimality of an adaptive first-order system least-squares finite element method (LS-FEM) was invented for the two-dimensional (2D) Poisson model problem in [12] and exploited for the Stokes equations in [8]. This paper extends those results to the first-order system least-squares formulation of linear elasticity [9] in three dimensions.

Numerical experiments show optimal behavior of an adaptive algorithm with least-squares formulations driven by the local contributions of the least-squares functional, e.g., in [13] for the Poisson model problem. However, this approach does not fit into the known mathematical techniques to guarantee optimal convergence rates [10, 15]. The affirmative result in [13] requires the bulk parameter Θ to be close to one while the known optimality [10] follows exclusively for Θ sufficiently small. An alternative a posteriori error estimator is therefore derived in this paper for the framework of the axioms of adaptivity and separate marking in [15].

More information on the history of least-squares finite element schemes may be found in [3] and on the mathematical foundation of adaptive algorithms in [10].

The polyhedral boundary $\partial\Omega$ of the bounded Lipschitz domain $\Omega \subset \mathbb{R}^3$ is split into some compact part $\Gamma_{\mathrm{D}} \subset \partial\Omega$ with positive surface measure $|\Gamma_{\mathrm{D}}| > 0$ and the relatively open complement $\Gamma_{\mathrm{N}} \coloneqq \partial\Omega \backslash \Gamma_{\mathrm{D}} \neq \emptyset$. Throughout this paper, Γ_{D} is supposed to belong exclusively to one of the connectivity components of $\partial\Omega$ for an immediate application

*Received by the editors July 7, 2016; accepted for publication (in revised form) June 29, 2017; published electronically February 1, 2018.
http://www.siam.org/journals/sinum/56-1/M108379.html
Funding: The authors were supported by the Deutsche Forschungsgemeinschaft in the Priority Program 1748, "Reliable simulation techniques in solid mechanics. Development of non-standard discretization methods, mechanical and mathematical analysis," under the projects "First-order system least-squares finite elements for finite elasto-plasticity" (STA 402/12-1) and "Foundation and application of generalized mixed FEM towards nonlinear problems in solid mechanics" (CA 151/22-1).
†Institut für Mathematik, Humboldt-Universität zu Berlin, 10099 Berlin, Germany (bringman@math.hu-berlin.de, cc@math.hu-berlin.de).
‡Department of Mathematics, Universität Duisburg-Essen, Thea-Leymann-Straße 9, 45127 Essen, Germany (gerhard.starke@uni-due.de).

of the exact sequence property in Proposition 2.3 below. This is no restriction if $\partial\Omega$ is connected. The 2D case is simpler and more easy to derive as in [8] and, hence, shall not be discussed explicitly in this paper. The pure Dirichlet problem may be included with the additional condition $\int_\Omega \operatorname{tr}\boldsymbol{\sigma}\,\mathrm{d}x = 0$ on the stress variables and is omitted for ease of notation.

Given $\boldsymbol{f} \in L^2(\Omega;\mathbb{R}^3)$ and $\boldsymbol{g} \in L^2(\Gamma_\mathrm{N};\mathbb{R}^3)$, the first-order system formulation of linear elasticity for $\boldsymbol{\sigma} \in H(\operatorname{div},\Omega;\mathbb{R}^{3\times 3})$ and $\boldsymbol{u} \in H^1(\Omega;\mathbb{R}^3)$ reads

$$\begin{gathered}\mathbb{C}^{-1}\boldsymbol{\sigma} - \boldsymbol{\varepsilon}(\boldsymbol{u}) = \mathbf{0} \quad \text{and} \quad \boldsymbol{f} + \operatorname{div}\boldsymbol{\sigma} = \mathbf{0} \quad \text{in } \Omega,\\ \boldsymbol{u} = \mathbf{0} \quad \text{on } \Gamma_\mathrm{D}, \quad \text{and} \quad \boldsymbol{\sigma\nu} = \boldsymbol{g} \quad \text{on } \Gamma_\mathrm{N}.\end{gathered} \tag{1.1}$$

The isotropic material law with Lamé parameters $\lambda, \mu > 0$ is the linear operator $\mathbb{C} : \mathbb{R}^{3\times 3} \to \mathbb{R}^{3\times 3}$ (also viewed as a fourth-order tensor) with $\mathbb{C}\boldsymbol{E} = 2\mu\boldsymbol{E} + \lambda(\operatorname{tr}\boldsymbol{E})\boldsymbol{I}_{3\times 3}$ for all $\boldsymbol{E} \in \mathbb{R}^{3\times 3}$ and

$$\mathbb{C}^{-1}\boldsymbol{\tau} \coloneqq \frac{1}{2\mu}\left(\boldsymbol{\tau} - \frac{\lambda}{3\lambda + 2\mu}(\operatorname{tr}\boldsymbol{\tau})\,\boldsymbol{I}_{3\times 3}\right) \quad \text{for } \boldsymbol{\tau} \in \mathbb{R}^{3\times 3}. \tag{1.2}$$

Note that $\mathbb{C}$ and $\mathbb{C}^{-1}$ map $\mathbb{S} \coloneqq \mathbb{R}^{3\times 3}_{\mathrm{sym}}$ into itself. Recall that (1.2) remains meaningful in the incompressible limit $\lambda \to \infty$ as it tends to $1/(2\mu)$ $\mathbf{dev}$ with the $\mathbf{dev}\,\boldsymbol{\tau} = \boldsymbol{\tau} - (\operatorname{tr}\boldsymbol{\tau})/3\,\boldsymbol{I}_{3\times 3}$. The weak spaces for the stress $\boldsymbol{\sigma}$ and the displacement $\boldsymbol{u}$ read

$$\begin{aligned}\boldsymbol{\Sigma}_{\boldsymbol{g}} &\coloneqq \{\boldsymbol{\tau} \in H(\operatorname{div},\Omega;\mathbb{R}^{3\times 3}) : \boldsymbol{\tau}\cdot\boldsymbol{\nu} = \boldsymbol{g} \text{ on } \Gamma_\mathrm{N}\},\\ \boldsymbol{V} &\coloneqq \{\boldsymbol{v} \in H^1(\Omega;\mathbb{R}^3) : \boldsymbol{v} = \mathbf{0} \text{ on } \Gamma_\mathrm{D}\}.\end{aligned} \tag{1.3}$$

The unique solution $(\boldsymbol{\sigma}, \boldsymbol{u}) \in \boldsymbol{\Sigma}_{\boldsymbol{g}} \times \boldsymbol{V}$ to (1.1) minimizes the functional

$$LS(\boldsymbol{f};\boldsymbol{\tau},\boldsymbol{v}) \coloneqq \|\boldsymbol{f} + \operatorname{div}\boldsymbol{\tau}\|^2_{L^2(\Omega)} + \|\mathbb{C}^{-1}\boldsymbol{\tau} - \boldsymbol{\varepsilon}(\boldsymbol{v})\|^2_{L^2(\Omega)} \tag{1.4}$$

among all $(\boldsymbol{\tau}, \boldsymbol{v}) \in \boldsymbol{\Sigma}_{\boldsymbol{g}} \times \boldsymbol{V}$ [9, Thm. 3.1]. Given the piecewise constant approximation $\boldsymbol{g}_0$ of $\boldsymbol{g}$, the LS-FEM seeks minimizers $(\boldsymbol{\sigma}_\mathrm{LS}, \boldsymbol{u}_\mathrm{LS}) \in \boldsymbol{\Sigma}_{\boldsymbol{g}_0}(\mathcal{T}) \times \mathcal{A}(\mathcal{T})$ of this functional in the Raviart–Thomas and Courant finite element function spaces

$$\boldsymbol{\Sigma}_{\boldsymbol{g}_0}(\mathcal{T}) \coloneqq RT_0(\mathcal{T};\mathbb{R}^{3\times 3}) \cap \boldsymbol{\Sigma}_{\boldsymbol{g}_0} \quad \text{and} \quad \boldsymbol{V}(\mathcal{T}) \coloneqq P_1(\mathcal{T};\mathbb{R}^3) \cap \boldsymbol{V}. \tag{1.5}$$

The local contributions of the functional LS provide a reliable and efficient built-in a posteriori error estimator. This paper introduces a novel error estimator η with the volume contributions on each tetrahedron T

$$|T|^{2/3}\|\operatorname{div}\operatorname{sym}\mathbb{C}^{-1}\boldsymbol{\sigma}_\mathrm{LS}\|^2_{L^2(T)} + |T|^{2/3}\|\operatorname{curl}\mathbb{C}^{-2}\boldsymbol{\sigma}_\mathrm{LS}\|^2_{L^2(T)}$$

and the interior side contributions, for the face $F \in \mathcal{F}(T) \cap \mathcal{F}(\Omega)$ with jump $[\,\bullet\,]_F$ across F,

$$\begin{aligned}&|T|^{1/3}\|[\operatorname{sym}\mathbb{C}^{-1}\boldsymbol{\sigma}_\mathrm{LS} - \boldsymbol{\varepsilon}(\boldsymbol{u}_\mathrm{LS})]_F\boldsymbol{\nu}_F\|^2_{L^2(F)}\\ &+ |T|^{1/3}\|[\mathbb{C}^{-1}(\mathbb{C}^{-1}\boldsymbol{\sigma}_\mathrm{LS} - \boldsymbol{\varepsilon}(\boldsymbol{u}_\mathrm{LS}))]_F \times \boldsymbol{\nu}_F\|^2_{L^2(F)}\end{aligned}$$

plus additional contributions on the boundary faces which involve Neumann boundary data oscillations. In contrast to the built-in estimator, the novel explicit residual-based a posteriori error estimator η requires an exact solve of the discrete equations.

The least-squares functional LS contains the data error $\mu^2(\mathcal{T}) \coloneqq \|\boldsymbol{f} - \boldsymbol{\Pi}_0 \boldsymbol{f}\|^2_{L^2(\Omega)}$ between $\boldsymbol{f}$ and its piecewise constant best-approximation $\boldsymbol{\Pi}_0 \boldsymbol{f}$. This error contribution is not monitored by the novel estimator η and, therefore, a separate marking strategy is required with a quasi-optimal data approximation algorithm for the reduction of $\mu^2(\mathcal{T})$.

The resulting adaptive algorithm is quasi-optimal in terms of the nonlinear approximation class $\mathcal{A}_s$ of all $(\boldsymbol{u}, \boldsymbol{f}) \in \mathcal{A} \times L^2(\Omega; \mathbb{R}^3)$ with a finite seminorm

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

with the L^2 projection $\boldsymbol{g}_0$ onto $P_0(\mathcal{F}(\Gamma_{\mathrm{N}}); \mathbb{R}^3)$ of $\boldsymbol{g}$ with respect to $\mathcal{T} \in \mathbb{T}(N)$ in the best possible error

$$E(\boldsymbol{u}, \boldsymbol{f}, N) \coloneqq \min_{\mathcal{T} \in \mathbb{T}(N)} \min_{(\boldsymbol{\tau}_{\mathrm{LS}}, \boldsymbol{v}_{\mathrm{LS}}) \in \boldsymbol{\Sigma}_{\boldsymbol{g}_0}(\mathcal{T}) \times \boldsymbol{V}(\mathcal{T})} \big(LS(\boldsymbol{f}; \boldsymbol{\tau}_{\mathrm{LS}}, \boldsymbol{v}_{\mathrm{LS}}) + \operatorname{osc}^2(\boldsymbol{g}, \mathcal{F}(\Gamma_{\mathrm{N}})) \big).$$

The extension of the analysis from [12, 8] for 2D to 3D encounters divergence-free lowest-order Raviart–Thomas finite element functions as curls of Nédélec functions of the first kind. The related stability and quasi-interpolation of the latter are established with the commuting diagram property from [23]. This paper is restricted to the lowest-order discretization for ease of presentation. Nevertheless, all results in section 2 also hold for all polynomial degrees $k \geqslant 1$ and the higher-order discrete quasi-interpolation for Nédélec functions from [28] enables the generalization of the proof of discrete reliability Theorem 5.1. For further details on discrete quasi-interpolation of Nédélec functions, the authors refer to [18, 19] and the references therein.

The remaining parts of the paper are organized as follows. Section 2 provides some preliminaries and notation employed in this paper. Section 3 presents the first-order system least-squares formulation of linear elasticity as well as some super approximation of the equilibrium residual in the stress-based finite element discretization. The main contribution of this paper is a new adaptive algorithm with optimal convergence rates in section 4 and three crucial parameters, which are all independent of the value of the Lamé parameter λ and independent of the mesh-sizes and solely depend on the initial triangulation $\mathcal{T}_0$. The proof of quasi-optimal convergence is based on the axioms of adaptivity in the form of [15] and the proof of the quasi-orthogonality (A4) conclude the section. The main technical ingredient for discrete reliability (A3) in section 5 relies on the relation to auxiliary mixed formulations of intermediate problems.

Standard notation of Sobolev and Lebesgue spaces such as $H^k(\Omega)$, $H(\mathrm{curl}, \Omega)$, $H(\mathrm{div}, \Omega)$, and $L^2(\Omega)$ and the corresponding spaces of vector- or matrix-valued functions $H^k(\Omega; \mathbb{R}^3)$, $H^k(\Omega; \mathbb{R}^{3\times 3})$, $H(\mathrm{curl}, \Omega; \mathbb{R}^{3\times 3})$, $H(\mathrm{div}, \Omega; \mathbb{R}^{3\times 3})$, $L^2(\Omega; \mathbb{R}^3)$, and $L^2(\Omega; \mathbb{R}^{3\times 3})$ apply throughout the paper. Appropriate subscripts indicate the norms of the subspaces of $H^1(\Omega; \mathbb{R}^3)$, $H(\mathrm{div}, \Omega; \mathbb{R}^{3\times 3})$, $H(\mathrm{curl}, \Omega, \mathbb{R}^{3\times 3})$, and $L^2(\Omega; \mathbb{R}^3)$. Let $\langle \bullet, \bullet \rangle_{\partial\Omega}$ denote the duality pairing of $\widetilde{H}^{1/2}(\Gamma) \coloneqq \{v|_\Gamma : v \in H^1(\Omega) \text{ with } v = 0 \text{ on } \partial\Omega \backslash \Gamma\}$ and its dual $H^{-1/2}(\Gamma)$, which extends the L^2 scalar product on some measurable subset $\Gamma \subset \partial\Omega$. Let $|\!|\!| \bullet |\!|\!| \coloneqq |\bullet|_{H^1(\Omega)} = \|\mathbf{D} \bullet\|_{L^2(\Omega)}$ abbreviate the H^1-seminorm. For $a, b \in \mathbb{R}^3$ and $M, N \in \mathbb{R}^{3\times 3}$, $a \cdot b \coloneqq a^\top b$ and $M : N \coloneqq \operatorname{tr}(M^\top N)$ abbreviate the Euclidian scalar products in $\mathbb{R}^3$ and $\mathbb{R}^{3\times 3}$.

Throughout the paper, $A \lesssim B$ abbreviates the relation $A \leqslant CB$ with a generic constant $C > 0$, which solely depends on the material parameter μ and shape-regularity of the underlying triangulations and so only on $\mathcal{T}_0$ and the newest-vertex bisection (NVB) [26]. Especially, those constants are independent of the Lamé parameter $\lambda > 0$; $A \approx B$ abbreviates $A \lesssim B \lesssim A$.

2. Preliminaries. To reduce the technical descriptions, the exposition in this paper keeps to the most important case of three spatial dimensions and $|\Gamma_{\mathrm{D}}|$ and $|\Gamma_{\mathrm{N}}| > 0$. Let $\Omega \subseteq \mathbb{R}^3$ denote a polyhedral Lipschitz domain with boundary $\partial\Omega$ (and outer unit normal vector $\boldsymbol{\nu} \in L^\infty(\partial\Omega;\mathbb{R}^3)$) partitioned into the closed Dirichlet boundary $\Gamma_{\mathrm{D}} \subset \Gamma_j \subset \partial\Omega$ included in exactly one of the connectivity components $\Gamma_1,\dots,\Gamma_J$ of $\partial\Omega$ with positive surface measure $|\Gamma_{\mathrm{D}}| > 0$ and the remaining (relatively open) Neumann boundary $\Gamma_{\mathrm{N}} \coloneqq \partial\Omega\backslash\Gamma_{\mathrm{D}} \neq \varnothing$.

2.1. Triangulations. Let $\mathcal{T}$ denote a regular triangulation of Ω into tetrahedra and let $\mathcal{T}$ resolve the decomposition of the boundary into Γ_{D} and Γ_{N} (cf. [16, sect. 3.2]). Let $\mathcal{F}$ denote the set of faces subordinated to Γ_{D} and Γ_{N} in that $\mathcal{F}(\Gamma_{\mathrm{D}}) \coloneqq \{F \in \mathcal{F} : F \subseteq \Gamma_{\mathrm{D}}\}$ and $\mathcal{F}(\Gamma_{\mathrm{N}}) \coloneqq \{F \in \mathcal{F} : F \subseteq \overline{\Gamma}_{\mathrm{N}}\}$ partition the set $\mathcal{F}(\partial\Omega)$ of faces on the boundary $\partial\Omega$.

Given an initial shape-regular triangulation $\mathcal{T}_0$ into closed tetrahedra of the polyhedral Lipschitz domain Ω with some initial condition [26, sect. 4, (a)–(b)] on the refinement edges, the set of admissible triangulations

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

follows from the NVB rules [26, sect. 2] for three dimensions.

With the counting measure $|\bullet|$ and the cardinality $|\mathcal{T}|$ of $\mathcal{T} \in \mathbb{T}$, let

$$\mathbb{T}(N) \coloneqq \{\mathcal{T} \in \mathbb{T} : |\mathcal{T}| - |\mathcal{T}_0| \leqslant N\} \quad \text{for any } N \in \mathbb{N}_0.$$

All admissible triangulations are shape-regular in the sense that for each $T \in \bigcup \mathbb{T} \coloneqq \{K : \exists \mathcal{T} \in \mathbb{T}, K \in \mathcal{T}\}$, the radius $\rho(T)$ of the largest inscribed sphere and the diameter $\operatorname{diam}(T)$ of T are uniformly bounded by a constant $C_0 > 0$ which solely depends on the initial triangulation $\mathcal{T}_0$; cf. [1, 26] for details on mesh-refining.

2.2. Finite element function spaces. Recall the definition of the spaces $\boldsymbol{\Sigma_g}$ and $\boldsymbol{V}$ from (1.3) as well as $\boldsymbol{\Sigma}_{\boldsymbol{g}_0}(\mathcal{T})$ and $\boldsymbol{V}(\mathcal{T})$ from (1.5). Let $P_0(\mathcal{T})$ (resp., $P_0(\mathcal{T};\mathbb{R}^3)$ or $P_0(\mathcal{T};\mathbb{R}^{3\times3})$) denote the space of piecewise constants (resp., for vector- or matrix-valued functions). The piecewise constant averages $f_0 \coloneqq \Pi_0 f \in P_0(\mathcal{T})$ coincide with the orthogonal projection of an L^2 function f onto $P_0(\mathcal{T})$ and analogously for every component of vector- or matrix-valued functions. Let $\mathrm{id} : \Omega \to \mathbb{R}^3$ denote the identity mapping.

The discrete approximation of rowwise $H(\mathrm{div})$ functions in $\boldsymbol{\Sigma_0}$ employs the space of rowwise Raviart–Thomas functions [7, 5, 4]

$$\begin{aligned}RT_0(\mathcal{T}) &\coloneqq \big\{q_{\mathrm{RT}} \in H(\mathrm{div},\Omega) : \forall T \in \mathcal{T}\, \exists a_T \in P_0(T;\mathbb{R}^3)\, \exists b_T \in P_0(T),\\ &\qquad\qquad q_{\mathrm{RT}}^\top\big|_T = a_T + b_T\,\mathrm{id}\big\},\\ RT_0(\mathcal{T};\mathbb{R}^{3\times3}) &\coloneqq \big\{\boldsymbol{\tau}_{\mathrm{RT}} = (\boldsymbol{\tau}_{jk})_{j,k=1,\dots,3} \in H(\mathrm{div},\Omega;\mathbb{R}^{3\times3}) :\\ &\qquad\qquad \forall j = 1,2,3,\ (\boldsymbol{\tau}_{j1},\boldsymbol{\tau}_{j2},\boldsymbol{\tau}_{j3}) \in RT_0(\mathcal{T})\big\}.\end{aligned}$$

The rowwise Nédélec functions of the first kind [21, 22, 20] read

$$\begin{aligned}N_0(\mathcal{T}) &\coloneqq \big\{\boldsymbol{\beta}_{\mathrm{N}} \in H(\mathrm{curl},\Omega) : \forall T \in \mathcal{T}\, \exists a_T, b_T \in P_0(\mathcal{T};\mathbb{R}^3),\\ &\qquad\qquad \boldsymbol{\beta}_{\mathrm{N}}^\top\big|_T = a_T + b_T \times \mathrm{id}\big\},\\ N_0(\mathcal{T};\mathbb{R}^{3\times3}) &\coloneqq \big\{\boldsymbol{\beta}_{\mathrm{N}} = (\boldsymbol{\beta}_{jk})_{j,k=1,\dots,3} \in H(\mathrm{curl},\Omega;\mathbb{R}^{3\times3}) :\\ &\qquad\qquad \forall j = 1,2,3,\ (\boldsymbol{\beta}_{j1},\boldsymbol{\beta}_{j2},\boldsymbol{\beta}_{j3}) \in N_0(\mathcal{T})\big\}.\end{aligned}$$

2.3. Approximation of Neumann boundary data. Let $H^{-1/2}(\Gamma_\mathrm{N};\mathbb{R}^3)$ denote the dual space of the trace space $\widetilde{H}^{1/2}(\Gamma_\mathrm{N};\mathbb{R}^3) \coloneqq \boldsymbol{\gamma_\nu}(\boldsymbol{V})$ with the normal trace operator $\boldsymbol{\gamma_\nu}$ acting on Γ_N equipped with the norm

$$\|\boldsymbol{g}\|_{H^{-1/2}(\Gamma_\mathrm{N})} \coloneqq \sup_{\boldsymbol{v}\in\boldsymbol{V}\setminus\{\boldsymbol{0}\}} \int_{\Gamma_\mathrm{N}} \boldsymbol{g}\cdot\boldsymbol{v}\,\mathrm{d}s \Big/ \|\boldsymbol{v}\|_{H^1(\Omega)}.$$

Given a regular triangulation $\mathcal{T} \in \mathbb{T}$ with the set of Neumann boundary faces $\mathcal{F}(\Gamma_\mathrm{N})$, approximate inhomogeneous boundary values $\boldsymbol{g} \in L^2(\Gamma_\mathrm{N};\mathbb{R}^3)$ by the L^2 best-approximation $\boldsymbol{g}_0 \coloneqq \boldsymbol{\Pi}_0\boldsymbol{g}$ in $P_0(\mathcal{F}(\Gamma_\mathrm{N});\mathbb{R}^3)$ with the L^2 orthogonality

$$\boldsymbol{g} - \boldsymbol{g}_0 \perp P_0(\mathcal{F}(\Gamma_\mathrm{N});\mathbb{R}^3).$$

For any face $F \in \mathcal{F}(\Gamma_\mathrm{N})$ of area $|F|$ and diameter h_F, let $\bar{\omega}_F \in \mathcal{T}$ denote the unique tetrahedron with $F \in \mathcal{F}(\bar{\omega}_F)$ of volume $|\omega_F|$ and diameter $h_{\omega_F} \coloneqq \mathrm{diam}(\omega_F)$ and

$$\mathrm{osc}^2(\boldsymbol{g},\mathcal{F}(\Gamma_\mathrm{N})) \coloneqq \sum_{F\in\mathcal{F}(\Gamma_\mathrm{N})} |\omega_F|^{1/3} \|\boldsymbol{g}-\boldsymbol{g}_0\|^2_{L^2(F)}.$$

LEMMA 2.1. *It holds that* $\|\boldsymbol{g}-\boldsymbol{g}_0\|_{H^{-1/2}(\Gamma_\mathrm{N})} \lesssim \mathrm{osc}(\boldsymbol{g},\mathcal{F}(\Gamma_\mathrm{N}))$.

Proof. Since $H^{-1/2}(\Gamma_\mathrm{N})$ is the dual space to $\widetilde{H}^{1/2}(\Gamma_\mathrm{N})$ endowed with the minimal extension norm, it suffices for any $\boldsymbol{v} \in \boldsymbol{V}$ with norm $\|\boldsymbol{v}\|_{H^1(\Omega)} = 1$ to prove

$$\int_{\Gamma_\mathrm{N}} (\boldsymbol{g}-\boldsymbol{g}_0)\cdot\boldsymbol{v}\,\mathrm{d}s \leqslant C_1\,\mathrm{osc}(\boldsymbol{g},\mathcal{F}(\Gamma_\mathrm{N})). \tag{2.1}$$

Given such a $\boldsymbol{v}$, let $\boldsymbol{v}_F \coloneqq \fint_{\omega_F} \boldsymbol{v}(x)\,\mathrm{d}x$ be the average of v in the face-patch ω_F (i.e., the interior of the one tetrahedron with face F) of $F \in \mathcal{F}(\Gamma_\mathrm{N})$ with diameter h_{ω_F} so that a Poincaré inequality with Payne–Weinberger constant reads

$$\|\boldsymbol{v}-\boldsymbol{v}_F\|_{L^2(\omega_F)} \leqslant \frac{h_{\omega_F}}{\pi}\|\mathbf{D}\,\boldsymbol{v}\|_{L^2(\omega_F)}. \tag{2.2}$$

Since $\fint_{\Gamma_\mathrm{N}}(\boldsymbol{g}-\boldsymbol{g}_0)\,\mathrm{d}s = 0$, it follows for the left-hand side in (2.1) that

$$\begin{aligned}\int_{\Gamma_\mathrm{N}} (\boldsymbol{g}-\boldsymbol{g}_0)\cdot\boldsymbol{v}\,\mathrm{d}s &= \sum_{F\in\mathcal{F}(\Gamma_\mathrm{N})}\int_F (\boldsymbol{g}-\boldsymbol{g}_0)\cdot(\boldsymbol{v}-\boldsymbol{v}_F)\,\mathrm{d}s\\ &\leqslant \sum_{F\in\mathcal{F}(\Gamma_\mathrm{N})} |\omega_F|^{1/6}\|\boldsymbol{g}-\boldsymbol{g}_0\|_{L^2(F)}|\omega_F|^{-1/6}\|\boldsymbol{v}-\boldsymbol{v}_F\|_{L^2(F)}\\ &\leqslant \mathrm{osc}(\boldsymbol{g},\mathcal{F}(\Gamma_\mathrm{N}))\sqrt{\sum_{F\in\mathcal{F}(\Gamma_\mathrm{N})}|\omega_F|^{-1/3}\|\boldsymbol{v}-\boldsymbol{v}_F\|^2_{L^2(F)}}. \qquad (2.3)\end{aligned}$$

The trace identity for $f \coloneqq |\boldsymbol{v}-\boldsymbol{v}_F|^2 \in H^1(\omega_F)$ on $\bar{\omega}_F \eqqcolon \mathrm{conv}\{F, P_F\}$ reads [11, Lem. 2.1]

$$\begin{aligned}|F|^{-1}\|\boldsymbol{v}-\boldsymbol{v}_F\|^2_{L^2(F)} &= \fint_F f(x)\,\mathrm{d}s = \fint_{\omega_F} f(x)\,\mathrm{d}x + \frac{1}{3}\fint_{\omega_F}(x-P_F)\cdot\nabla f(x)\,\mathrm{d}x\\ &= |\omega_F|^{-1}\|\boldsymbol{v}-\boldsymbol{v}_F\|^2_{L^2(\omega_F)} + 2\frac{h_{\omega_F}}{3|\omega_F|}\|\boldsymbol{v}-\boldsymbol{v}_F\|_{L^2(\omega_F)}\|\mathbf{D}\,\boldsymbol{v}\|_{L^2(\omega_F)}.\end{aligned}$$

This and (2.2) prove

$$\|\boldsymbol{v}-\boldsymbol{v}_F\|^2_{L^2(F)} \leqslant \frac{|F|h^2_{\omega_F}}{\pi|\omega_F|}(1/\pi+2/3)\|\,\mathbf{D}\,\boldsymbol{v}\|^2_{L^2(\omega_F)}.$$

With the uniformly bounded constant

$$C_1^2 \coloneqq \frac{4}{\pi}\Big(\frac{2}{3}+\frac{1}{\pi}\Big)\max_{F\in\mathcal{F}(\Gamma_{\mathrm{N}})}|F|\,h^2_{\omega_F}|\omega_F|^{-4/3}\lesssim 1,$$

the weighted sum of all those contributions reads

$$\sum_{F\in\mathcal{F}(\Gamma_{\mathrm{N}})}|\omega_F|^{-1/3}\|\boldsymbol{v}-\boldsymbol{v}_F\|^2_{L^2(F)} \leqslant C_1^2/4\sum_{F\in\mathcal{F}(\Gamma_{\mathrm{N}})}\|\,\mathbf{D}\,\boldsymbol{v}\|^2_{L^2(\omega_F)}.$$

The finite overlap of the family $(\omega_F\ :\ F\in\mathcal{F}(\Gamma_{\mathrm{N}}))$ shows that the last term is $\leqslant C_1^2\|\,\mathbf{D}\,\boldsymbol{v}\|_{L^2(\Omega)}\leqslant C_1^2$. The combination with (2.3) concludes the proof of (2.1). ∎

LEMMA 2.2. *There exists some constant $C_{\mathrm{N}}\approx 1$, which solely depends on the geometry of Ω, Γ_{D}, and Γ_{N}, such that for any given $\boldsymbol{g}\in L^2(\Gamma_{\mathrm{N}};\mathbb{R}^3)\subseteq H^{-1/2}(\Gamma_{\mathrm{N}};\mathbb{R}^3)$ with L^2 best-approximation $\boldsymbol{g}_0\coloneqq\mathbf{\Pi}_0\boldsymbol{g}\in P_0(\mathcal{F}(\Gamma_{\mathrm{N}});\mathbb{R}^3)$, there exists some extension $\boldsymbol{\tau}\in H(\mathrm{div},\Omega;\mathbb{R}^{3\times 3})$ with*

$$\boldsymbol{\tau}\boldsymbol{\nu}=\boldsymbol{g}-\boldsymbol{g}_0 \text{ on } \Gamma_{\mathrm{N}} \quad\text{and}\quad \|\boldsymbol{\tau}\|_{H(\mathrm{div},\Omega)}\leqslant C_{\mathrm{N}}\,\mathrm{osc}(\boldsymbol{g},\mathcal{F}(\Gamma_{\mathrm{N}})). \tag{2.4}$$

If $\boldsymbol{g}\in P_0(\widehat{\mathcal{F}}(\Gamma_{\mathrm{N}});\mathbb{R}^3)$ for some admissible refinement $\widehat{\mathcal{T}}$ of $\mathcal{T}$, then $\boldsymbol{\tau}$ in (2.4) can be found in $RT_0(\widehat{\mathcal{T}};\mathbb{R}^{3\times 3})$.

Proof. Let $\boldsymbol{\tau}\in H(\mathrm{div},\Omega;\mathbb{R}^{3\times 3})$ and $\boldsymbol{v}\in H^1(\Omega;\mathbb{R}^3)$ solve

$$-\operatorname{div}\boldsymbol{\tau}=\mathbf{0} \text{ and } \mathbf{D}\,\boldsymbol{v}=\boldsymbol{\tau} \text{ in } \Omega,\quad \boldsymbol{v}=0 \text{ on } \Gamma_{\mathrm{D}}, \text{ and }\quad \boldsymbol{\tau}\boldsymbol{\nu}=\boldsymbol{g}-\boldsymbol{g}_0 \text{ on } \Gamma_{\mathrm{N}}. \tag{2.5}$$

The stability of the boundary value problem leads to

$$\|\boldsymbol{\tau}\|_{H(\mathrm{div},\Omega)}=\|\,\mathbf{D}\,\boldsymbol{v}\|_{L^2(\Omega)}\lesssim\|\boldsymbol{g}-\boldsymbol{g}_0\|_{H^{-1/2}(\Gamma_{\mathrm{N}})}.$$

This and Lemma 2.1 conclude the proof of (2.4).

For $\boldsymbol{g}\in P_0(\widehat{\mathcal{F}}(\Gamma_{\mathrm{N}});\mathbb{R}^3)$, let $\widehat{\boldsymbol{\tau}}_{\mathrm{RT}}\in RT_0(\widehat{\mathcal{T}};\mathbb{R}^{3\times 3})$ denote the mixed finite element solution to the boundary value problem (2.5). ∎

2.4. Auxiliary problem. Given $\boldsymbol{f}\in L^2(\Omega;\mathbb{R}^3)$ and $\boldsymbol{g}\in L^2(\Omega;\mathbb{R}^3)$, let $\boldsymbol{z}\in\boldsymbol{V}\equiv H^1_{\mathrm{D}}(\Omega;\mathbb{R}^3)$ denote the unique solution to

$$-\Delta\,\boldsymbol{z}=\boldsymbol{f}\quad\text{in }\Omega\quad\text{and}\quad\frac{\partial\boldsymbol{z}}{\partial\boldsymbol{\nu}}=\boldsymbol{g}\quad\text{on }\Gamma_{\mathrm{N}}.$$

Since $\boldsymbol{f}\in L^2(\Omega;\mathbb{R}^3)$ and $\boldsymbol{g}\in L^2(\Omega;\mathbb{R}^3)$, the reduced elliptic regularity [17] implies $\boldsymbol{\tau}\coloneqq\mathbf{D}\,\boldsymbol{z}\in H^s(\Omega;\mathbb{R}^{3\times 3})$ for some $s>0$ and

$$\|\boldsymbol{\tau}\|_{H(\mathrm{div},\Omega)}\leqslant C(\Omega,\Gamma_{\mathrm{N}})\big(\|\boldsymbol{f}\|_{L^2(\Omega)}+\|\boldsymbol{g}\|_{H^{-1/2}(\Gamma_{\mathrm{N}})}\big).$$

Let $\boldsymbol{\tau}_0\coloneqq\mathcal{I}_{\mathrm{F}}\boldsymbol{\tau}\in RT_0(\mathcal{T};\mathbb{R}^{3\times 3})$ denote the Fortin interpolation of $\boldsymbol{\tau}$ from [4, eq. (2.5.26)]. Then it holds that $-\operatorname{div}\boldsymbol{\tau}_0=-\mathbf{\Pi}_0\operatorname{div}\boldsymbol{\tau}=\Pi_0\boldsymbol{f}$ and

$$\|\boldsymbol{\tau}_0\|_{H(\mathrm{div},\Omega)}\lesssim\|\boldsymbol{\tau}\|_{H(\mathrm{div},\Omega)}\lesssim\big(\|\boldsymbol{f}\|_{L^2(\Omega)}+\|\boldsymbol{g}\|_{H^{-1/2}(\Gamma_{\mathrm{N}})}\big). \tag{2.6}$$

Moreover, for all $F\in\mathcal{F}(\Gamma_{\mathrm{N}})$ [4, eq. (2.5.10), p. 107],

$$\boldsymbol{\tau}_0\,\boldsymbol{\nu}_F=\mathbf{\Pi}_{0,F}(\boldsymbol{\tau}\,\boldsymbol{\nu}_F)=\boldsymbol{g}_0.$$

2.5. Divergence-free Raviart–Thomas functions. In the notation of [20, p. 37], let $\Gamma_0, \dots, \Gamma_J$ denote all connectivity components of $\partial\Omega$ and recall that $\Gamma_\mathrm{D} \subseteq \Gamma_j$ for exactly one Γ_j of the connectivity components.

PROPOSITION 2.3. *Given* $\boldsymbol{\rho}_\mathrm{RT} \in RT_0(\mathcal{T})$ *with* $\operatorname{div} \boldsymbol{\rho}_\mathrm{RT} = 0$ *and* $\boldsymbol{\rho}_\mathrm{RT} \cdot \boldsymbol{\nu} = 0$ *on* Γ_N*, there exists* $\boldsymbol{\beta}_\mathrm{N} \in N_0(\mathcal{T})$ *with* $\operatorname{curl} \boldsymbol{\beta}_\mathrm{N} = \boldsymbol{\rho}_\mathrm{RT}$ *and*

$$\|\boldsymbol{\beta}_\mathrm{N}\|_{H(\mathrm{curl},\Omega)} \lesssim \|\boldsymbol{\rho}_\mathrm{RT}\|_{L^2(\Omega)}. \tag{2.7}$$

The proof of the proposition employs the following function spaces:

$$\begin{aligned} H(\Omega, \operatorname{div} = 0) &\coloneqq \{\boldsymbol{q} \in H(\operatorname{div}, \Omega) \,:\, \operatorname{div} \boldsymbol{q} = 0 \text{ in } \Omega\}, \\ H_0(\Omega, \operatorname{div} = 0) &\coloneqq \{\boldsymbol{q} \in H(\Omega, \operatorname{div} = 0) \,:\, \boldsymbol{q} \cdot \boldsymbol{\nu} = 0 \text{ on } \partial\Omega\}, \\ H_0(\mathrm{curl}, \Omega) &\coloneqq \{\boldsymbol{v} \in H(\mathrm{curl}, \Omega) \,:\, \boldsymbol{\nu} \times \boldsymbol{v} = 0 \text{ on } \partial\Omega\}. \end{aligned}$$

Remark 2.4 ($\operatorname{curl} H_0(\mathrm{curl}, \Omega) \subseteq H_0(\Omega, \operatorname{div} = 0)$). For any $\boldsymbol{v} \in H_0(\mathrm{curl}, \Omega)$ and $w \in C^\infty(\Omega)$, $\operatorname{curl} \boldsymbol{v} \in H(\Omega, \operatorname{div} = 0)$ and Green's formulas for the gradient and the curl imply

$$\int_{\partial\Omega} w \, \operatorname{curl} \boldsymbol{v} \cdot \boldsymbol{\nu} \,\mathrm{d}s = \int_\Omega \nabla w \cdot \operatorname{curl} \boldsymbol{v} \,\mathrm{d}x = \int_{\partial\Omega} \nabla w \cdot (\boldsymbol{\nu} \times \boldsymbol{v}) \,\mathrm{d}x = 0.$$

This is the weak form of $(\operatorname{curl} \boldsymbol{v}) \cdot \boldsymbol{\nu} = 0$ on $\partial\Omega$; written $\operatorname{curl} \boldsymbol{v} \in H_0(\Omega, \operatorname{div} = 0)$.

Proof of Proposition 2.3. *Step* 1. The assumption $\Gamma_\mathrm{D} \subseteq \Gamma_j$ for some fixed index $j \in \{1, \dots, J\}$ implies $\Gamma_k \subseteq \Gamma_\mathrm{N}$ for every $k = 0, \dots, J$ with $k \neq j$. Consequently, $\int_{\Gamma_k} \boldsymbol{\rho}_\mathrm{RT} \cdot \boldsymbol{\nu} \,\mathrm{d}s = 0$. Since $\operatorname{div} \boldsymbol{\rho}_\mathrm{RT} = 0$ a.e. in Ω implies $\int_{\partial\Omega} \boldsymbol{\rho}_\mathrm{RT} \cdot \boldsymbol{\nu} \,\mathrm{d}s = 0$, it follows that

$$\int_{\Gamma_k} \boldsymbol{\rho}_\mathrm{RT} \cdot \boldsymbol{\nu} \,\mathrm{d}s = 0 \quad \forall\, k = 1, \dots, J.$$

Let $\widehat{\boldsymbol{\rho}} \in H(\operatorname{div}, \widehat{\Omega})$ denote the extension of $\boldsymbol{\rho}_\mathrm{RT}$ to some large ball $\widehat{\Omega}$, which includes Ω, with $\widehat{\boldsymbol{\rho}} = \boldsymbol{\rho}_\mathrm{RT}$ on Ω, $\operatorname{div} \widehat{\boldsymbol{\rho}} = 0$ on $\widehat{\Omega}$, and $\widehat{\boldsymbol{\rho}} \cdot \boldsymbol{\nu} = 0$ on $\partial\widehat{\Omega}$. Following [20, pp. 38, 46], the design of $\widehat{\boldsymbol{\rho}}$ is via some Laplace Neumann problem [20, (N), p. 38] with Neumann data $\boldsymbol{\rho}_\mathrm{RT} \cdot \boldsymbol{\nu}$ on Γ_D and $\boldsymbol{\rho}_\mathrm{RT} \cdot \boldsymbol{\nu} \equiv 0$ on Γ_N. This ensures standard estimates on a neighborhood of $\widehat{\Omega} \backslash \Omega$, in particular $\|\widehat{\boldsymbol{\rho}}\|_{L^2(\widehat{\Omega})} \lesssim \|\boldsymbol{\rho}_\mathrm{RT}\|_{L^2(\Omega)}$.

Step 2. Since $H_1 \coloneqq H(\mathrm{curl}, \widehat{\Omega}) \cap H_0(\widehat{\Omega}, \operatorname{div} = 0)$ is a closed subspace of $H(\mathrm{curl}, \widehat{\Omega})$ and since $H_2 \coloneqq H_0(\widehat{\Omega}, \operatorname{div} = 0)$ is a closed subspace of $H(\operatorname{div}, \widehat{\Omega})$, the linear map $\operatorname{curl} : H_1 \to H_2$ is bounded between the Hilbert spaces H_1 and H_2. Theorem I.3.5 from [20] asserts that $\operatorname{curl} : H_1 \to H_2$ is injective and surjective (as the ball $\widehat{\Omega}$ is simply connected). As a bounded bijection between Hilbert spaces, the inverse $\operatorname{curl}^{-1} : H_2 \to H_1$ is bounded as well. This leads to a generic constant $C > 0$ with

$$\|\operatorname{curl}^{-1} \boldsymbol{v}\|_{H(\mathrm{curl},\widehat{\Omega})} \leqslant C \|\boldsymbol{v}\|_{H(\operatorname{div},\widehat{\Omega})} \quad \text{for any } \boldsymbol{v} \in H_2 \equiv H(\widehat{\Omega}, \operatorname{div} = 0). \tag{2.8}$$

Hence, given $\widehat{\boldsymbol{\rho}}$ in the ball $\widehat{\Omega}$ with $\widehat{\boldsymbol{\rho}} \cdot \boldsymbol{\nu} = 0$ along $\partial\widehat{\Omega}$, i.e., $\widehat{\boldsymbol{\rho}} \in H_2$, there exists a $\widehat{\boldsymbol{\beta}} \in H_1$ with

$$\begin{aligned} &\operatorname{curl} \widehat{\boldsymbol{\beta}} = \widehat{\boldsymbol{\rho}}, \quad \operatorname{div} \widehat{\boldsymbol{\beta}} = 0, \quad \widehat{\boldsymbol{\beta}} \cdot \boldsymbol{\nu} = 0 \quad \text{on } \partial\widehat{\Omega}, \quad \text{and} \\ &\|\widehat{\boldsymbol{\beta}}\|_{H(\mathrm{curl},\widehat{\Omega})} \lesssim \|\widehat{\boldsymbol{\rho}}\|_{H(\operatorname{div},\widehat{\Omega})} = \|\widehat{\boldsymbol{\rho}}\|_{L^2(\widehat{\Omega})} \lesssim \|\boldsymbol{\rho}_\mathrm{RT}\|_{L^2(\Omega)}. \end{aligned} \tag{2.9}$$

Step 3. Let $\boldsymbol{\beta}_\mathrm{N} \coloneqq \boldsymbol{\Pi}_\mathrm{N}(\widehat{\boldsymbol{\beta}}|_\Omega) \in N_0(\Omega)$ denote the projection $\boldsymbol{\Pi}_\mathrm{N}$ of $\widehat{\boldsymbol{\beta}}$ from [23, Thm. 7] as part of the commuting diagram property of the quasi-interpolation operators from [23] of Figure 2.1. This and the projection property of $\boldsymbol{\Pi}_\mathrm{RT}$ read

$$
\begin{array}{ccc}
H(\mathrm{curl},\Omega) & \xrightarrow{\ \mathrm{curl}\ } & H(\mathrm{div},\Omega) \\
\Big\downarrow \boldsymbol{\Pi}_{\mathrm{N}} & & \Big\downarrow \boldsymbol{\Pi}_{\mathrm{RT}} \\
N_0(\mathcal{T}) & \xrightarrow{\ \mathrm{curl}\ } & RT_0(\mathcal{T})
\end{array}
$$

FIG. 2.1. *Commuting diagram property of Schöberl quasi-interpolation operators.*

$$\operatorname{curl} \boldsymbol{\beta}_{\mathrm{N}} = \operatorname{curl} \boldsymbol{\Pi}_{\mathrm{N}}(\widehat{\boldsymbol{\beta}}\big|_{\Omega}) = \boldsymbol{\Pi}_{\mathrm{RT}} \boldsymbol{\rho}_{\mathrm{RT}} = \boldsymbol{\rho}_{\mathrm{RT}}.$$

The L^2 stability of $\boldsymbol{\Pi}_{\mathrm{N}}$ [23, Thm. 8] plus (2.9) yield

$$\|\boldsymbol{\beta}_{\mathrm{N}}\|_{L^2(\Omega)} \lesssim \|\widehat{\boldsymbol{\beta}}\|_{L^2(\Omega)} \leqslant \|\widehat{\boldsymbol{\beta}}\|_{H(\mathrm{curl},\widehat{\Omega})} \lesssim \|\boldsymbol{\rho}_{\mathrm{RT}}\|_{L^2(\Omega)}.$$

This and $\|\operatorname{curl} \boldsymbol{\beta}_{\mathrm{N}}\|_{L^2(\Omega)} = \|\boldsymbol{\rho}_{\mathrm{RT}}\|_{L^2(\Omega)}$ conclude the proof. ∎

Remark 2.5. The authors conjecture that the result from Proposition 2.3 can be generalized to higher polynomial degrees. Given $\boldsymbol{\rho}_{\mathrm{RT}} \in RT_k(\mathcal{T})$ with $\operatorname{div} \boldsymbol{\rho}_{\mathrm{RT}} = 0$, the existence of a vector potential $\boldsymbol{\beta}_{\mathrm{N}} \in N_k(\mathcal{T})$ with $\operatorname{curl} \boldsymbol{\beta}_{\mathrm{N}} = \boldsymbol{\rho}_{\mathrm{RT}}$ is well-known in the context of exact sequences of finite element function spaces [4, subsect. 2.5.6, p. 116]. However, the L^2 stability (2.7) is not straightforward. The commuting and L^2 stable higher-order Nédélec quasi-interpolation operator from [18, Thm. 6.5] generalizes the Schöberl quasi-interpolation $\boldsymbol{\Pi}_{\mathrm{N}}$ in the proof at hand.

3. Stress-based finite element discretization. Recall the definition of the spaces $\boldsymbol{\Sigma}_{\boldsymbol{g}}$ and $\boldsymbol{V}$ from (1.3) as well as $\boldsymbol{\Sigma}_{\boldsymbol{g}_0}(\mathcal{T})$ and $\boldsymbol{V}(\mathcal{T})$ from (1.5). The minimization of the least-squares functional (1.4) is equivalent to the variational problem

$$\int_\Omega \operatorname{div} \boldsymbol{\sigma} \cdot \operatorname{div} \boldsymbol{\tau} \, \mathrm{d}x + \int_\Omega (\mathbb{C}^{-1}\boldsymbol{\sigma} - \boldsymbol{\varepsilon}(\boldsymbol{u})) : (\mathbb{C}^{-1}\boldsymbol{\tau} - \boldsymbol{\varepsilon}(\boldsymbol{v})) \, \mathrm{d}x = -\int_\Omega \boldsymbol{f} \cdot \operatorname{div} \boldsymbol{\tau} \, \mathrm{d}x \tag{3.1}$$

for all $(\boldsymbol{\tau}, \boldsymbol{v}) \in \boldsymbol{\Sigma}_{\mathbf{0}} \times \boldsymbol{V}$. The equivalence of the homogeneous least-squares functional with the associated norm on $\boldsymbol{\Sigma}_{\mathbf{0}} \times \boldsymbol{V}$ implies the uniqueness of the solution to (3.1).

PROPOSITION 3.1 (see [9, Thm. 3.1]). *Any $(\boldsymbol{\tau}, \boldsymbol{v}) \in \boldsymbol{\Sigma}_{\mathbf{0}} \times \boldsymbol{V}$ satisfies*

$$LS(\mathbf{0}; \boldsymbol{\tau}, \boldsymbol{v}) \approx \|\boldsymbol{\tau}\|^2_{H(\mathrm{div},\Omega)} + \|\boldsymbol{\varepsilon}(\boldsymbol{v})\|^2_{L^2(\Omega)}.$$

The first-order system least-squares finite element approximation reads as follows: Minimize (1.4) among all $(\boldsymbol{\tau}, \boldsymbol{v}) \in \boldsymbol{\Sigma}_{\boldsymbol{g}_0}(\mathcal{T}) \times \boldsymbol{V}(\mathcal{T})$. Since $\operatorname{div} \boldsymbol{\Sigma}(\mathcal{T}) = P_0(\mathcal{T}; \mathbb{R}^3)$, we may replace $\boldsymbol{f}$ by $\boldsymbol{f}_0 = \boldsymbol{\Pi}_0 \boldsymbol{f}$ in the discrete version of (3.1): Seek $(\boldsymbol{\sigma}_{\mathrm{LS}}, \boldsymbol{u}_{\mathrm{LS}}) \in \boldsymbol{\Sigma}_{\boldsymbol{g}_0}(\mathcal{T}) \times \boldsymbol{V}(\mathcal{T})$ with

$$\begin{aligned}
&\int_\Omega \operatorname{div} \boldsymbol{\sigma}_{\mathrm{LS}} \cdot \operatorname{div} \boldsymbol{\tau}_0 \, \mathrm{d}x + \int_\Omega (\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}})) : (\mathbb{C}^{-1}\boldsymbol{\tau}_0 - \boldsymbol{\varepsilon}(\boldsymbol{v}_0)) \, \mathrm{d}x \\
&\quad = -\int_\Omega \boldsymbol{f}_0 \cdot \operatorname{div} \boldsymbol{\tau}_0 \, \mathrm{d}x \ \forall \ (\boldsymbol{\tau}_0, \boldsymbol{v}_0) \in \boldsymbol{\Sigma}(\mathcal{T}) \times \boldsymbol{V}(\mathcal{T}).
\end{aligned} \tag{3.2}$$

The following proposition controls the compatibility of the traction boundary conditions with Lemma 2.2. Let $(\boldsymbol{\sigma}_{\mathrm{LS}}, \boldsymbol{u}_{\mathrm{LS}})$ (resp., $(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}, \widehat{\boldsymbol{u}}_{\mathrm{LS}})$) solve the discrete problem (3.2) with respect to the regular triangulation $\mathcal{T} \in \mathbb{T}$ (resp., some admissible refinement $\widehat{\mathcal{T}} \in \mathbb{T}$ of $\mathcal{T}$).

PROPOSITION 3.2. *The universal constant* $C_{\mathrm{osc}} \coloneqq \max\{1, 1/(4\mu^2)\} C_{\mathrm{N}}$ *with* C_{N} *from Lemma 2.2 and any positive* ε *satisfy*

$$\begin{aligned}(1-\varepsilon) LS(\boldsymbol{f}; \widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}, \widehat{\boldsymbol{u}}_{\mathrm{LS}}) &+ LS(\mathbf{0}; \widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}, \widehat{\boldsymbol{u}}_{\mathrm{LS}} - \boldsymbol{u}_{\mathrm{LS}}) \\ &\leqslant LS(\boldsymbol{f}; \boldsymbol{\sigma}_{\mathrm{LS}}, \boldsymbol{u}_{\mathrm{LS}}) + C_{\mathrm{osc}}/\varepsilon \ \mathrm{osc}^2(\widehat{\boldsymbol{g}}_0, \mathcal{F}(\Gamma_{\mathrm{N}})).\end{aligned}$$

Proof. Elementary algebra proves

$$\begin{aligned}& LS(\boldsymbol{f}; \widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}, \widehat{\boldsymbol{u}}_{\mathrm{LS}}) - LS(\boldsymbol{f}; \boldsymbol{\sigma}_{\mathrm{LS}}, \boldsymbol{u}_{\mathrm{LS}}) \\ &= -\| \operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}})\|_{L^2(\Omega)}^2 - \|\mathbb{C}^{-1}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}) - \boldsymbol{\varepsilon}(\widehat{\boldsymbol{u}}_{\mathrm{LS}} - \boldsymbol{u}_{\mathrm{LS}})\|_{L^2(\Omega)}^2 \\ &\quad + 2\int_\Omega (f + \operatorname{div} \widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}) \cdot \operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}) \,\mathrm{d}x \\ &\quad + 2\int_\Omega (\mathbb{C}^{-1}\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\widehat{\boldsymbol{u}}_{\mathrm{LS}})) : (\mathbb{C}^{-1}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}) - \boldsymbol{\varepsilon}(\widehat{\boldsymbol{u}}_{\mathrm{LS}} - \boldsymbol{u}_{\mathrm{LS}})) \,\mathrm{d}x.\end{aligned}$$

Given the extension $\widehat{\boldsymbol{\tau}}_{\mathrm{RT}} \in RT_0(\widehat{\mathcal{T}}; \mathbb{R}^{3\times 3})$ from Lemma 2.2 with $\widehat{\boldsymbol{\tau}}_{\mathrm{RT}}\boldsymbol{\nu} = \widehat{\boldsymbol{g}}_0 - \boldsymbol{g}_0$, the function $\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}} - \widehat{\boldsymbol{\tau}}_{\mathrm{RT}} \in \boldsymbol{\Sigma}_{\mathbf{0}}(\widehat{\mathcal{T}})$ is an admissible test function in the discrete equations (3.2) with respect to the refined triangulation $\widehat{\mathcal{T}}$. This, the Cauchy–Schwarz inequality, the Young inequality with respect to parameter $0 < \varepsilon < 1$, and $\|\mathbb{C}^{-1}\boldsymbol{\tau}\|_{L^2(\Omega)} \leqslant 1/(2\mu) \, \|\boldsymbol{\tau}\|_{L^2(\Omega)}$ imply

$$\begin{aligned}& LS(\boldsymbol{f}; \widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}, \widehat{\boldsymbol{u}}_{\mathrm{LS}}) - LS(\boldsymbol{f}; \boldsymbol{\sigma}_{\mathrm{LS}}, \boldsymbol{u}_{\mathrm{LS}}) \\ &+ \| \operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}})\|_{L^2(\Omega)}^2 + \|\mathbb{C}^{-1}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}) - \boldsymbol{\varepsilon}(\widehat{\boldsymbol{u}}_{\mathrm{LS}} - \boldsymbol{u}_{\mathrm{LS}})\|_{L^2(\Omega)}^2 \\ &= 2\int_\Omega (f + \operatorname{div} \widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}) \cdot \operatorname{div} \widehat{\boldsymbol{\tau}}_{\mathrm{RT}} \,\mathrm{d}x + 2\int_\Omega (\mathbb{C}^{-1}\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\widehat{\boldsymbol{u}}_{\mathrm{LS}})) : \mathbb{C}^{-1}\widehat{\boldsymbol{\tau}}_{\mathrm{RT}} \,\mathrm{d}x \\ &\leqslant 2\|\boldsymbol{f} + \operatorname{div} \widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}\|_{L^2(\Omega)} \|\operatorname{div} \widehat{\boldsymbol{\tau}}_{\mathrm{RT}}\|_{L^2(\Omega)} + 1/\mu \, \|\mathbb{C}^{-1}\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\widehat{\boldsymbol{u}}_{\mathrm{LS}})\|_{L^2(\Omega)} \|\widehat{\boldsymbol{\tau}}_{\mathrm{RT}}\|_{L^2(\Omega)} \\ &\leqslant \varepsilon LS(\boldsymbol{f}; \widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}, \widehat{\boldsymbol{u}}_{\mathrm{LS}}) + \max\{1, 1/(4\mu^2)\}/\varepsilon \, \|\widehat{\boldsymbol{\tau}}_{\mathrm{RT}}\|_{H(\operatorname{div},\Omega)}^2.\end{aligned}$$

This and Lemma 2.2 conclude the proof with $C_{\mathrm{osc}} \coloneqq \max\{1, 1/(4\mu^2)\} \widehat{C}_{\mathrm{N}}^2$. ∎

Let $(\widetilde{\boldsymbol{\sigma}}^{\boldsymbol{g}_0}, \widetilde{\boldsymbol{u}}^{\boldsymbol{g}_0}) \in \boldsymbol{\Sigma}_{\boldsymbol{g}_0} \times \boldsymbol{V}$ denote the corresponding solution to the continuous problem (1.1) with boundary values $\boldsymbol{g}_0$ instead of $\boldsymbol{g}$. Then, the ellipticity result from Proposition 3.1 implies, for any $(\boldsymbol{\tau}_0, \boldsymbol{v}_0) \in \boldsymbol{\Sigma}_{\boldsymbol{g}_0}(\mathcal{T}) \times \boldsymbol{V}(\mathcal{T})$, that

$$LS(\boldsymbol{f}; \boldsymbol{\tau}_0, \boldsymbol{v}_0) \approx \|\widetilde{\boldsymbol{\sigma}}^{\boldsymbol{g}_0} - \boldsymbol{\tau}_0\|_{H(\operatorname{div},\Omega)}^2 + \|\boldsymbol{\varepsilon}(\widetilde{\boldsymbol{u}}^{\boldsymbol{g}_0} - \boldsymbol{v}_0)\|_{L^2(\Omega)}^2. \tag{3.3}$$

In order to study of the approximation behavior of the first-order system LS-FEM (3.2), define the sets

$$\begin{aligned}\mathcal{S}(L^2(\Omega; \mathbb{R}^3)) &\coloneqq \{\boldsymbol{f} \in L^2(\Omega; \mathbb{R}^3) \,:\, \|\boldsymbol{f}\|_{L^2(\Omega)} = 1\}, \\ Q_0(\mathcal{T}, \boldsymbol{f}) &\coloneqq \{\boldsymbol{\tau}_0 \in \boldsymbol{\Sigma}_{\boldsymbol{g}_0}(\mathcal{T}) \,:\, \boldsymbol{\Pi}_0\boldsymbol{f} + \operatorname{div} \boldsymbol{\tau}_0 = \mathbf{0}\}.\end{aligned}$$

The quantities

$$\begin{aligned}\rho_0(\mathcal{T}) &\coloneqq \sup_{\boldsymbol{f} \in \mathcal{S}(L^2(\Omega;\mathbb{R}^3))} \inf_{\boldsymbol{\tau} \in Q_0(\mathcal{T},\boldsymbol{f})} \|\mathbb{C}^{-1}(\widetilde{\boldsymbol{\sigma}}^{\boldsymbol{g}_0} - \boldsymbol{\tau})\|_{L^2(\Omega)}, \\ \xi_0(\mathcal{T}) &\coloneqq \sup_{\boldsymbol{f} \in \mathcal{S}(L^2(\Omega;\mathbb{R}^3))} \inf_{\boldsymbol{v} \in \boldsymbol{V}(\mathcal{T})} \|\boldsymbol{\varepsilon}(\widetilde{\boldsymbol{u}}^{\boldsymbol{g}_0} - \boldsymbol{v})\|_{L^2(\Omega)}\end{aligned} \tag{3.4}$$

represent distances of certain finite element spaces to $(\widetilde{\boldsymbol{\sigma}}^{\boldsymbol{g}_0}, \widetilde{\boldsymbol{u}}^{\boldsymbol{g}_0})$ and depend on the regularity of the solution and on the way the triangulation is adapted.

The following result and its proof are motivated by the investigations in [6] about the supercloseness of first-order system least-squares approximations to those produced by mixed methods of saddle-point type.

THEOREM 3.3. *It holds that*

$$\|\operatorname{div}\boldsymbol{\sigma}_{\mathrm{LS}} + \boldsymbol{\Pi}_0\boldsymbol{f}\|^2_{L^2(\Omega)} \lesssim (\rho_0(\mathcal{T}) + \xi(\mathcal{T}))^2\, LS(\boldsymbol{f}; \boldsymbol{\sigma}_{\mathrm{LS}}, \boldsymbol{u}_{\mathrm{LS}}). \tag{3.5}$$

Proof. Step 1. Recall $\boldsymbol{f}_0 = \boldsymbol{\Pi}_0\boldsymbol{f}$ and $\operatorname{div}\boldsymbol{\sigma}_{\mathrm{LS}} \in P_0(\mathcal{T};\mathbb{R}^3)$. Let $(\widetilde{\boldsymbol{\sigma}}^{\boldsymbol{g}_0}, \widetilde{\boldsymbol{u}}^{\boldsymbol{g}_0}) \in \boldsymbol{\Sigma}_{\boldsymbol{g}_0} \times \boldsymbol{V}$ solve (1.1) with boundary values $\boldsymbol{g}_0$ replacing $\boldsymbol{g}$ and define the sphere

$$\mathcal{S}(P_0(\mathcal{T};\mathbb{R}^3)) := \{\boldsymbol{q}_0 \in P_0(\mathcal{T};\mathbb{R}^3) : \|\boldsymbol{q}_0\|_{L^2(\Omega)} = 1\}.$$

For any $\boldsymbol{z}_0 \in P_0(\mathcal{T};\mathbb{R}^3)$, determine $\boldsymbol{\Xi} \in \boldsymbol{\Sigma}_{\boldsymbol{0}}$ and $\boldsymbol{\eta} \in \boldsymbol{V}$ via the auxiliary boundary value problem

$$\operatorname{div}\boldsymbol{\Xi} = \boldsymbol{z}_0 \quad \text{and} \quad \mathbb{C}^{-1}\boldsymbol{\Xi} - \boldsymbol{\varepsilon}(\boldsymbol{\eta}) = \boldsymbol{0} \quad \text{in } \Omega. \tag{3.6}$$

Step 2. Let $\boldsymbol{\Xi}_0 \in \boldsymbol{\Sigma}_{\boldsymbol{0}}(\mathcal{T})$ with $\operatorname{div}\,\boldsymbol{\Xi}_0 = \boldsymbol{z}_0 = \operatorname{div}\,\boldsymbol{\Xi}$ and $\boldsymbol{\eta}_0 \in \boldsymbol{V}(\mathcal{T})$. Since $\operatorname{div}(\boldsymbol{\Xi}_0 - \boldsymbol{\Xi}) = \boldsymbol{0}$, the continuous equation (3.1) with the test functions $\boldsymbol{\tau} \equiv \boldsymbol{\Xi}_0 - \boldsymbol{\Xi}$ and $\boldsymbol{v} \equiv \boldsymbol{\eta}_0 - \boldsymbol{\eta}$ implies

$$\int_\Omega (\mathbb{C}^{-1}\widetilde{\boldsymbol{\sigma}}^{\boldsymbol{g}_0} - \boldsymbol{\varepsilon}(\widetilde{\boldsymbol{u}}^{\boldsymbol{g}_0})) : (\mathbb{C}^{-1}(\boldsymbol{\Xi}_0 - \boldsymbol{\Xi}) - \boldsymbol{\varepsilon}(\boldsymbol{\eta}_0 - \boldsymbol{\eta}))\,\mathrm{d}x = 0. \tag{3.7}$$

The continuous equation (3.1) with $\boldsymbol{\tau} \equiv \boldsymbol{\Xi}_0$ and $\boldsymbol{v} \equiv \boldsymbol{\eta}_0$ reads

$$\begin{aligned}
&\int_\Omega \operatorname{div}\widetilde{\boldsymbol{\sigma}}^{\boldsymbol{g}_0} \cdot \operatorname{div}\boldsymbol{\Xi}_0\,\mathrm{d}x + \int_\Omega (\mathbb{C}^{-1}\widetilde{\boldsymbol{\sigma}}^{\boldsymbol{g}_0} - \boldsymbol{\varepsilon}(\widetilde{\boldsymbol{u}}^{\boldsymbol{g}_0})) : (\mathbb{C}^{-1}\boldsymbol{\Xi}_0 - \boldsymbol{\varepsilon}(\boldsymbol{\eta}_0))\,\mathrm{d}x \\
&\quad = -\int_\Omega \boldsymbol{f} \cdot \operatorname{div}\boldsymbol{\Xi}_0\,\mathrm{d}x.
\end{aligned}$$

The discrete equation (3.2) with $\boldsymbol{\tau}_0 \equiv \boldsymbol{\Xi}_0$ and $\boldsymbol{v}_0 \equiv \boldsymbol{\eta}_0$ reads

$$\begin{aligned}
&-\int_\Omega \operatorname{div}\boldsymbol{\sigma}_{\mathrm{LS}} \cdot \operatorname{div}\boldsymbol{\Xi}_0\,\mathrm{d}x - \int_\Omega (\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}})) : (\mathbb{C}^{-1}\boldsymbol{\Xi}_0 - \boldsymbol{\varepsilon}(\boldsymbol{\eta}_0))\,\mathrm{d}x \\
&\quad = -\int_\Omega \boldsymbol{\Pi}_0\boldsymbol{f} \cdot \operatorname{div}\boldsymbol{\Xi}_0\,\mathrm{d}x.
\end{aligned}$$

The sum of the last and second to last displayed formulas plus the L^2 orthogonality, (3.7), the second equation in (3.6), and $\operatorname{div}\boldsymbol{\Xi}_0 = \boldsymbol{z}_0 \in P_0(\mathcal{T};\mathbb{R}^3)$ lead to

$$\begin{aligned}
&\int_\Omega \operatorname{div}(\widetilde{\boldsymbol{\sigma}}^{\boldsymbol{g}_0} - \boldsymbol{\sigma}_{\mathrm{LS}}) \cdot \boldsymbol{z}_0\,\mathrm{d}x \\
&\quad = -\int_\Omega (\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}})) : (\mathbb{C}^{-1}(\boldsymbol{\Xi} - \boldsymbol{\Xi}_0) - \boldsymbol{\varepsilon}(\boldsymbol{\eta} - \boldsymbol{\eta}_0))\,\mathrm{d}x.
\end{aligned} \tag{3.8}$$

Step 3. Since $\operatorname{div}\boldsymbol{\sigma}_{\mathrm{LS}} + \boldsymbol{\Pi}_0\boldsymbol{f} \in P_0(\mathcal{T};\mathbb{R}^3)$,

$$\begin{aligned}
\|\operatorname{div}\,\boldsymbol{\sigma}_{\mathrm{LS}} + \boldsymbol{\Pi}_0\boldsymbol{f}\|_{L^2(\Omega)} &= \sup_{\boldsymbol{z}_0 \in \mathcal{S}(P_0(\mathcal{T};\mathbb{R}^3))} \int_\Omega (\operatorname{div}\boldsymbol{\sigma}_{\mathrm{LS}} + \boldsymbol{f}) \cdot \boldsymbol{z}_0\,\mathrm{d}x \\
&= \sup_{\boldsymbol{z}_0 \in \mathcal{S}(P_0(\mathcal{T};\mathbb{R}^3))} \int_\Omega \operatorname{div}\,(\boldsymbol{\sigma}_{\mathrm{LS}} - \widetilde{\boldsymbol{\sigma}}^{\boldsymbol{g}_0}) \cdot \boldsymbol{z}_0\,\mathrm{d}x.
\end{aligned}$$

The combination with (3.8) leads to

$$\begin{aligned}
&\|\operatorname{div}\boldsymbol{\sigma}_{\mathrm{LS}}+\boldsymbol{\Pi}_0\boldsymbol{f}\|_{L^2(\Omega)}\\
&\leqslant \|\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}}-\boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}})\|_{L^2(\Omega)} \sup_{\substack{\boldsymbol{z}_0\in\mathcal{S}(P_0(\mathcal{T};\mathbb{R}^3))\\ (\boldsymbol{\Xi},\boldsymbol{\eta})\text{ with (3.6)}}} \|\mathbb{C}^{-1}(\boldsymbol{\Xi}-\boldsymbol{\Xi}_0)-\boldsymbol{\varepsilon}(\boldsymbol{\eta}-\boldsymbol{\eta}_0)\|_{L^2(\Omega)}\\
&\leqslant LS(\boldsymbol{f};\boldsymbol{\sigma}_{\mathrm{LS}},\boldsymbol{u}_{\mathrm{LS}})^{1/2} \sup_{\substack{\boldsymbol{z}_0\in\mathcal{S}(P_0(\mathcal{T};\mathbb{R}^3))\\ (\boldsymbol{\Xi},\boldsymbol{\eta})\text{ with (3.6)}}} \Big(\|\mathbb{C}^{-1}(\boldsymbol{\Xi}-\boldsymbol{\Xi}_0)\|_{L^2(\Omega)}+\|\boldsymbol{\varepsilon}(\boldsymbol{\eta}-\boldsymbol{\eta}_0)\|_{L^2(\Omega)}\Big).
\end{aligned}$$

Recall that $\boldsymbol{\Xi}_0\in\boldsymbol{\Sigma_0}(\mathcal{T})$ with $\operatorname{div}\boldsymbol{\Xi}_0=\boldsymbol{z}_0$ and $\boldsymbol{\eta}_0\in\boldsymbol{V}(\mathcal{T})$ are arbitrary, to conclude the proof of (3.5). ∎

4. Quasi-optimal adaptive algorithm.

4.1. Alternative a posteriori error estimator. Let $(\boldsymbol{\sigma}_{\mathrm{LS}},\boldsymbol{u}_{\mathrm{LS}})$ solve the discrete equations (3.2) and let $\eta^2(\mathcal{T}) \coloneqq \sum_{T\in\mathcal{T}}\eta^2(\mathcal{T},T)$ denote the alternative a posteriori error estimator with

$$\begin{aligned}
\eta^2(\mathcal{T},T) \coloneqq{}& |T|^{2/3}\|\operatorname{div}\operatorname{sym}\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}}\|^2_{L^2(T)}+|T|^{2/3}\|\operatorname{curl}\mathbb{C}^{-2}\boldsymbol{\sigma}_{\mathrm{LS}}\|^2_{L^2(T)}\\
&+|T|^{1/3}\sum_{F\in\mathcal{F}(T)\setminus\mathcal{F}(\Gamma_{\mathrm{D}})}\|[\operatorname{sym}\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}}-\boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}})]_F\boldsymbol{\nu}_F\|^2_{L^2(F)}\\
&+|T|^{1/3}\sum_{F\in\mathcal{F}(T)\setminus\mathcal{F}(\Gamma_{\mathrm{N}})}\|[\mathbb{C}^{-1}(\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}}-\boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}}))]_F\times\boldsymbol{\nu}_F\|^2_{L^2(F)}\\
&+|T|^{1/3}\sum_{F\in\mathcal{F}(T)\cap\mathcal{F}(\Gamma_{\mathrm{N}})}\|\boldsymbol{g}-\boldsymbol{g}_0\|^2_{L^2(F)}.
\end{aligned}\tag{4.1}$$

All the terms in the estimator (4.1) except the last one appear nonstandard in the scaling with the compliance tensor. They do, however, arise naturally from the treatment of the least-squares formulation in section 5.

4.2. Efficiency. The discrete test function technology due to Verfürth [27] leads to local efficiency of the estimator η from (4.1) in the following sense.

THEOREM 4.1 (efficiency). *It holds that*

$$\eta^2(\mathcal{T})+\|\boldsymbol{f}-\boldsymbol{\Pi}_0\boldsymbol{f}\|^2_{L^2(\Omega)} \lesssim LS(\boldsymbol{f};\boldsymbol{\sigma}_{\mathrm{LS}},\boldsymbol{u}_{\mathrm{LS}})+\operatorname{osc}^2(\boldsymbol{g},\mathcal{F}(\Gamma_{\mathrm{N}})).$$

Proof. The corresponding arguments from [12] and [8] can be adopted immediately and further details are omitted. ∎

4.3. Adaptive algorithm (ALS-FEM).

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

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

  Solve LS-FEM (3.2) with respect to regular triangulation $\mathcal{T}_\ell$ with solution $(\boldsymbol{\sigma}_\ell,\boldsymbol{u}_\ell)$ and the piecewise constant best-approximation $\boldsymbol{f}_\ell\coloneqq\boldsymbol{\Pi}_\ell\boldsymbol{f}$ on $\mathcal{T}_\ell$.

  Compute $\eta_\ell(T)\coloneqq\eta(\mathcal{T}_\ell,T)$ from (4.1) for all $T\in\mathcal{T}_\ell$ and set $\eta_\ell^2\coloneqq\eta^2(\mathcal{T}_\ell)$.

  if CASE A $\|\boldsymbol{f}-\boldsymbol{f}_\ell\|^2_{L^2(\Omega)}\leqslant\kappa\eta_\ell^2$ **then**

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

$$\theta\eta_\ell^2\leqslant\eta_\ell^2(\mathcal{M}_\ell)\coloneqq\sum_{T\in\mathcal{M}_\ell}\eta_\ell^2(T).$$

Compute smallest regular refinement $\mathcal{T}_{\ell+1}$ of $\mathcal{T}_\ell$ with $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell \backslash \mathcal{T}_{\ell+1}$ by NVB.
else (**CASE B** $\kappa\eta_\ell^2 < \|\boldsymbol{f} - \boldsymbol{f}_\ell\|_{L^2(\Omega)}^2$)
Compute an admissible refinement $\mathcal{T}_{\ell+1}$ of $\mathcal{T}_\ell$ with (almost) minimal cardinality $|\mathcal{T}_{\ell+1}|$ and $\|\boldsymbol{f} - \boldsymbol{f}_{\ell+1}\|_{L^2(\Omega)} \leqslant \rho\|\boldsymbol{f} - \boldsymbol{f}_\ell\|_{L^2(\Omega)}$. **fi od**
Output: Sequence of discrete solutions $(\boldsymbol{\sigma}_\ell, \boldsymbol{u}_\ell)_{\ell\in\mathbb{N}_0}$ and meshes $(\mathcal{T}_\ell)_{\ell\in\mathbb{N}_0}$.

Remark 4.2 (Case B). The thresholding second algorithm (TSA) of [2, sect. 5] is one possible realization 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.

4.4. Quasi-optimal convergence. The main result of this paper involves, for any given $0 < s < \infty$, the notion of approximation classes $\mathcal{A}_s$ which consist of all pairs $(\boldsymbol{u}, \boldsymbol{f}) \in \mathcal{A} \times L^2(\Omega; \mathbb{R}^3)$ such that

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

with the best possible error

$$E(\boldsymbol{u}, \boldsymbol{f}, N) := \min_{\mathcal{T}\in\mathbb{T}(N)} \min_{(\boldsymbol{\tau}_{\mathrm{LS}}, \boldsymbol{v}_{\mathrm{LS}})\in\boldsymbol{\Sigma}_{\boldsymbol{g}_0}(\mathcal{T})\times\boldsymbol{V}_{\boldsymbol{g}_0}(\mathcal{T})} \big(LS(\boldsymbol{f}; \boldsymbol{\tau}_{\mathrm{LS}}, \boldsymbol{v}_{\mathrm{LS}}) + \operatorname{osc}^2(\boldsymbol{g}, \mathcal{F}(\Gamma_{\mathrm{N}}))\big).$$

THEOREM 4.3. *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* $0 < \theta \leqslant \theta_0$*, for all* $0 < \kappa \leqslant \kappa_0$*, for all* $0 < \rho < 1$*, and for all* $0 < s < \infty$*, the output* $(\boldsymbol{\sigma}_\ell, \boldsymbol{u}_\ell)_\ell$ *of ALS-FEM with* $(\boldsymbol{u}, \boldsymbol{f}) \in \mathcal{A}_s$ *satisfies*

$$\sup_{\ell\in\mathbb{N}} \big(|\mathcal{T}_\ell| - |\mathcal{T}_0|\big)^s \big(LS(\boldsymbol{f}; \boldsymbol{\sigma}_\ell, \boldsymbol{u}_\ell) + \operatorname{osc}^2(\boldsymbol{g}, \mathcal{F}_\ell(\Gamma_{\mathrm{N}}))\big)^{1/2} \leqslant C_{\mathrm{qopt}} \big|(\boldsymbol{u}, \boldsymbol{f})\big|_{\mathcal{A}_s}.$$

The constant $C_{\mathrm{qopt}} < \infty$ *depends only on the initial mesh* $\mathcal{T}_0$ *the constant* s *and the parameters* $\rho, \theta,$ *and* κ*; all the parameters* κ_0*,* θ_0*, and* C_{qopt} *are* λ*-independent.*

The proof of the converse inequality "$\gtrsim$" is discussed in [15, Thm. 2.1.b] with arguments applicable to the situation at hand.

4.5. Axioms of adaptivity. This section summarizes the convergence analysis of [15] based on the axioms (A1)–(A4), (B1)–(B2), and (QM) for the proof of Theorem 4.3. The axioms (A1)–(A3) and (B2) concern an admissible refinement $\widehat{\mathcal{T}} \in \mathbb{T}$ of an arbitrary triangulation $\mathcal{T} \in \mathbb{T}$ and the associated discrete solutions $(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}, \widehat{\boldsymbol{u}}_{\mathrm{LS}})$ and $(\boldsymbol{\sigma}_{\mathrm{LS}}, \boldsymbol{u}_{\mathrm{LS}})$ to the discrete equations (3.2) in the definition of the distance

$$\begin{aligned} \delta^2(\widehat{\mathcal{T}}, \mathcal{T}) := {} & \|\operatorname{div}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}})\|_{L^2(\Omega)}^2 \\ & + \|\mathbb{C}^{-1}(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}}) - \boldsymbol{\varepsilon}(\widehat{\boldsymbol{u}}_{\mathrm{LS}} - \boldsymbol{u}_{\mathrm{LS}})\|_{L^2(\Omega)}^2 + \operatorname{osc}^2(\widehat{\boldsymbol{g}}_0, \mathcal{F}(\Gamma_{\mathrm{N}})). \end{aligned} \tag{4.2}$$

For any subset $\mathcal{M} \subseteq \mathcal{T}$, let $\eta^2(\mathcal{T}, \mathcal{M}) := \sum_{T\in\mathcal{M}} \eta^2(\mathcal{T}, T)$ abbreviate the sum over the corresponding error estimator contributions from (4.1). In particular, $\eta^2(\mathcal{T}) := \eta^2(\mathcal{T}, \mathcal{T})$. Let $\mu(\mathcal{T}) := \|\boldsymbol{f} - \boldsymbol{\Pi}_0\boldsymbol{f}\|_{L^2(\Omega)}$ with the L^2 projection $\boldsymbol{\Pi}_0$ on $P_0(\mathcal{T}; \mathbb{R}^3)$.

THEOREM 4.4 (stability and reduction). *It holds that*

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

$$\eta(\widehat{\mathcal{T}}, \widehat{\mathcal{T}} \backslash \mathcal{T}) \leqslant \rho_2 \eta(\mathcal{T}, \mathcal{T} \backslash \widehat{\mathcal{T}}) + \Lambda_2 \delta(\widehat{\mathcal{T}}, \mathcal{T}). \tag{A2}$$

Proof. The proofs of (A1)–(A2) are straightforward from [10, 8, 12]. ∎

The proof of the discrete reliability $\exists \mathcal{R} \subseteq \mathcal{T}$ with $\mathcal{T} \backslash \widehat{\mathcal{T}} \subseteq \mathcal{R}$, $|\mathcal{R}| \lesssim |\mathcal{T} \backslash \widehat{\mathcal{T}}|$, and

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

is postponed to section 5.

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

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

follows directly from (A1)–(A2), the following theorem, and [15, Thm. 3.1].

THEOREM 4.5 (quasi-orthogonality with $\varepsilon > 0$). *For any sequence of successive admissible refinements* $\mathcal{T}_0, \mathcal{T}_1, \dots \in \mathbb{T}$ *and all positive* ε*, there exists a generic constant* $\Lambda_4(\varepsilon) \approx 1$ *with*

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

Proof. For all $k = \ell, \dots, \ell + m$ and positive ε, Proposition 3.2 proves

$$\begin{aligned}(1-\varepsilon) LS(\boldsymbol{f}; \boldsymbol{\sigma}_{k+1}, \boldsymbol{u}_{k+1}) + LS(\boldsymbol{0}; \boldsymbol{\sigma}_{k+1} - \boldsymbol{\sigma}_k, \boldsymbol{u}_{k+1} - \boldsymbol{u}_k)\\ \leqslant LS(\boldsymbol{f}; \boldsymbol{\sigma}_k, \boldsymbol{u}_k) + C_{\mathrm{osc}}/\varepsilon \ \operatorname{osc}^2(\boldsymbol{g}_{k+1}, \mathcal{F}_k(\Gamma_{\mathrm{N}})).\end{aligned}$$

The orthogonality of the boundary data oscillations leads to

$$\operatorname{osc}^2(\boldsymbol{g}_{k+1}, \mathcal{F}_k(\Gamma_{\mathrm{N}})) + \operatorname{osc}^2(\boldsymbol{g}, \mathcal{F}_{k+1}(\Gamma_{\mathrm{N}})) \leqslant \operatorname{osc}^2(\boldsymbol{g}, \mathcal{F}_k(\Gamma_{\mathrm{N}})).$$

Consequently,

$$\begin{aligned}&\delta^2(\mathcal{T}_{k+1}, \mathcal{T}_k) - \varepsilon LS(\boldsymbol{f}; \boldsymbol{\sigma}_{k+1}, \boldsymbol{u}_{k+1})\\ &\quad\leqslant LS(\boldsymbol{f}; \boldsymbol{\sigma}_k, \boldsymbol{u}_k) - LS(\boldsymbol{f}; \boldsymbol{\sigma}_{k+1}, \boldsymbol{u}_{k+1})\\ &\qquad + (1 + C_{\mathrm{osc}}/\varepsilon)\big(\operatorname{osc}^2(\boldsymbol{g}, \mathcal{F}_k(\Gamma_{\mathrm{N}})) - \operatorname{osc}^2(\boldsymbol{g}, \mathcal{F}_{k+1}(\Gamma_{\mathrm{N}}))\big).\end{aligned}$$

The telescoping sum over all $k = \ell, \dots, \ell + m$ and the reliability from Corollary 5.2 conclude the proof with $\Lambda_4(\varepsilon) := 1 + C_{\mathrm{rel}} + C_{\mathrm{osc}}/\varepsilon$. ∎

The subsequent assumptions (B1)–(B2) transfer directly from [15] to the situation at hand in three components for the TSA plus completion (called APPROX in [14]).

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

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

Since $\widehat{\Lambda}_3$ may be large, the following result is required and proven explicitly.

THEOREM 4.6 (quasi-monotonicity of $\eta + \mu$). *It holds that*

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

Proof. The efficiency from subsection 4.2 plus Proposition 3.2 and the reliability from Corollary 5.2 prove (QM). ∎

5. Discrete reliability. Let $\widehat{\mathcal{T}} \in \mathbb{T}$ denote an admissible refinement of $\mathcal{T} \in \mathbb{T}$ with respective discrete solutions $(\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}, \widehat{\boldsymbol{u}}_{\mathrm{LS}})$ and $(\boldsymbol{\sigma}_{\mathrm{LS}}, \boldsymbol{u}_{\mathrm{LS}})$ to (3.2). Recall the definition of $\delta(\widehat{\mathcal{T}}, \mathcal{T})$ from (4.2). Let $\mathcal{R} \coloneqq \{K \in \mathcal{T} : \exists K' \in \mathcal{T} \backslash \widehat{\mathcal{T}},\ K \cap K' \neq \emptyset\}$ be one layer of simplices around and including $\mathcal{T} \backslash \widehat{\mathcal{T}}$. This set satisfies $|\mathcal{R}| \lesssim |\mathcal{T} \backslash \widehat{\mathcal{T}}|$.

THEOREM 5.1 (discrete reliability). *It holds that*

$$\delta^2(\widehat{\mathcal{T}}, \mathcal{T}) \lesssim \eta^2(\mathcal{T}, \mathcal{R}) + \|\boldsymbol{f} - \boldsymbol{f}_0\|^2_{L^2(\Omega)} + LS(\boldsymbol{f}; \widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}, \widehat{\boldsymbol{u}}_{\mathrm{LS}}). \tag{5.1}$$

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

COROLLARY 5.2 (reliability). *For any admissible triangulation $\mathcal{T} \in \mathbb{T}$ with discrete solutions $(\boldsymbol{\sigma}_{\mathrm{LS}}, \boldsymbol{u}_{\mathrm{LS}}) \in \boldsymbol{\Sigma}_{\boldsymbol{g}_0}(\mathcal{T}) \times \boldsymbol{V}(\mathcal{T})$ to* (3.2)*, it holds that*

$$LS(\boldsymbol{f}; \boldsymbol{\sigma}_{\mathrm{LS}}, \boldsymbol{u}_{\mathrm{LS}}) \leqslant C_{\mathrm{rel}}\big(\eta^2(\mathcal{T}) + \|\boldsymbol{f} - \boldsymbol{f}_0\|^2_{L^2(\Omega)}\big). \tag{5.2}$$

Proof. The proof of [8, Cor. 4.4] relies on the discrete reliability (5.1) with $\widehat{\mathcal{T}}$ replaced by successive uniform refinements of $\mathcal{T}$ and applies literally to the situation at hand. The convergence of the LS-FEM in the limit as the maximal mesh-sizes tend to zero proves (5.2). ∎

Proof of (A3). The combination of the estimate (5.1) with (5.2) and $|\mathcal{R}| \lesssim |\mathcal{T} \backslash \widehat{\mathcal{T}}|$ proves (A3) from subsection 4.5. ∎

The remaining part of this section is devoted to the proof of Theorem 5.1 and utilizes the abbreviations

$$\boldsymbol{\delta} \coloneqq \widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\sigma}_{\mathrm{LS}} \quad \text{and} \quad \boldsymbol{e} \coloneqq \widehat{\boldsymbol{u}}_{\mathrm{LS}} - \boldsymbol{u}_{\mathrm{LS}}.$$

Three intermediate solutions to the auxiliary problem from subsection 2.4 are important. Let $\boldsymbol{\tau}_{\mathrm{RT}} \in \boldsymbol{\Sigma}_{\mathbf{0}}(\mathcal{T})$ and $\widehat{\boldsymbol{\tau}}^*_{\mathrm{RT}} \in \boldsymbol{\Sigma}_{\mathbf{0}}(\mathcal{T})$ satisfy

$$\operatorname{div} \boldsymbol{\tau}_{\mathrm{RT}} = \operatorname{div} \widehat{\boldsymbol{\tau}}^*_{\mathrm{RT}} = \boldsymbol{\Pi}_0 \operatorname{div} \boldsymbol{\delta}.$$

Let $\widehat{\boldsymbol{\tau}}_{\mathrm{RT}} \in RT_0(\mathcal{T}; \mathbb{R}^{3\times 3})$ satisfy

$$\operatorname{div} \widehat{\boldsymbol{\tau}}_{\mathrm{RT}} = \operatorname{div} \boldsymbol{\delta} \quad \text{and} \quad \widehat{\boldsymbol{\tau}}_{\mathrm{RT}}\, \boldsymbol{\nu} = \widehat{\boldsymbol{g}}_0 - \boldsymbol{g}_0 \quad \text{on } \Gamma_{\mathrm{N}}.$$

The stability estimate (2.6) and Lemma 2.1 lead to

$$\begin{aligned} \|\boldsymbol{\tau}_{\mathrm{RT}}\|_{H(\operatorname{div},\Omega)} + \|\widehat{\boldsymbol{\tau}}^*_{\mathrm{RT}}\|_{H(\operatorname{div},\Omega)} &\lesssim \|\boldsymbol{\Pi}_0 \operatorname{div} \boldsymbol{\delta}\|_{L^2(\Omega)} \quad \text{and} \\ \|\widehat{\boldsymbol{\tau}}_{\mathrm{RT}}\|_{H(\operatorname{div},\Omega)} &\lesssim \|\operatorname{div} \boldsymbol{\delta}\|_{L^2(\Omega)} + \operatorname{osc}(\widehat{\boldsymbol{g}}_0, \mathcal{F}(\Gamma_{\mathrm{N}})). \end{aligned} \tag{5.3}$$

The analysis of $\delta^2(\widehat{\mathcal{T}}, \mathcal{T})$ departs with elementary algebra.

LEMMA 5.3. *There exists some $\widehat{\boldsymbol{\beta}}_{\mathrm{N}} \in N_0(\widehat{\mathcal{T}}; \mathbb{R}^{3\times 3})$ with*

$$\begin{aligned} &\|\widehat{\beta}_{\mathrm{N}}\|_{L^2(\Omega)} \lesssim \|\boldsymbol{\delta}\|_{H(\operatorname{div},\Omega)} + \operatorname{osc}(\widehat{\boldsymbol{g}}_0, \mathcal{F}(\Gamma_{\mathrm{N}})) \quad \textit{and} \\ &LS(\mathbf{0}; \boldsymbol{\delta}, \boldsymbol{e}) = \|(1 - \boldsymbol{\Pi}_0) \operatorname{div} \widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}\|^2_{L^2(\Omega)} \\ &\qquad + \int_\Omega (\mathbb{C}^{-1}\boldsymbol{\delta} - \boldsymbol{\varepsilon}(\boldsymbol{e})) : (\mathbb{C}^{-1}(\widehat{\boldsymbol{\tau}}_{\mathrm{RT}} - \widehat{\boldsymbol{\tau}}^*_{\mathrm{RT}}))\,\mathrm{d}x \\ &\qquad + \int_\Omega (\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}})) : (\boldsymbol{\varepsilon}(\boldsymbol{e}) - \mathbb{C}^{-1} \operatorname{curl} \widehat{\boldsymbol{\beta}}_{\mathrm{N}})\,\mathrm{d}x. \end{aligned} \tag{5.4}$$

Proof. *Step* 1. Since $\operatorname{div} \boldsymbol{\tau}_{\mathrm{RT}} = \operatorname{div} \boldsymbol{\delta}$, the discrete equations (3.2) with respect to the triangulation $\mathcal{T}$ and the test functions $\boldsymbol{\tau}_{\mathrm{LS}} = \boldsymbol{\tau}_{\mathrm{RT}}$ and $\boldsymbol{v}_{\mathrm{LS}} = \mathbf{0}$ read

$$\int_\Omega (\boldsymbol{f}_0 + \operatorname{div} \boldsymbol{\sigma}_{\mathrm{LS}}) \cdot \operatorname{div} \boldsymbol{\delta} \,\mathrm{d}x + \int_\Omega (\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}})) : \mathbb{C}^{-1}\boldsymbol{\tau}_{\mathrm{RT}} \,\mathrm{d}x = 0. \tag{5.5}$$

The same arguments with respect to $\widehat{\mathcal{T}}$, $\boldsymbol{\tau}_{\mathrm{LS}} = \boldsymbol{\delta} + \widehat{\boldsymbol{\tau}}_{\mathrm{RT}}^* - \widehat{\boldsymbol{\tau}}_{\mathrm{RT}}$, and $\boldsymbol{v}_{\mathrm{LS}} = \boldsymbol{e}$ show

$$\begin{aligned} &\int_\Omega (\widehat{\boldsymbol{f}}_0 + \operatorname{div} \widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}) \cdot \boldsymbol{\Pi}_0 \operatorname{div} \boldsymbol{\delta} \,\mathrm{d}x \\ &\quad + \int_\Omega (\mathbb{C}^{-1}\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\widehat{\boldsymbol{u}}_{\mathrm{LS}})) : (\mathbb{C}^{-1}(\boldsymbol{\delta} + \widehat{\boldsymbol{\tau}}_{\mathrm{RT}}^* - \widehat{\boldsymbol{\tau}}_{\mathrm{RT}}) - \boldsymbol{\varepsilon}(\boldsymbol{e})) \,\mathrm{d}x = 0. \end{aligned} \tag{5.6}$$

The summation of (5.5) and $LS(\mathbf{0}; \boldsymbol{\delta}, \boldsymbol{e})$ leads to

$$\begin{aligned} LS(\mathbf{0}; \boldsymbol{\delta}, \boldsymbol{e}) = \|\operatorname{div} \boldsymbol{\delta}\|_{L^2(\Omega)}^2 &+ \int_\Omega (\mathbb{C}^{-1}\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\widehat{\boldsymbol{u}}_{\mathrm{LS}})) : (\mathbb{C}^{-1}\boldsymbol{\delta} - \boldsymbol{\varepsilon}(\boldsymbol{e})) \,\mathrm{d}x \\ &+ \int_\Omega (\boldsymbol{f}_0 + \operatorname{div} \boldsymbol{\sigma}_{\mathrm{LS}}) \cdot \operatorname{div} \boldsymbol{\delta} \,\mathrm{d}x \\ &- \int_\Omega (\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}})) : (\mathbb{C}^{-1}(\boldsymbol{\delta} - \boldsymbol{\tau}_{\mathrm{RT}}) - \boldsymbol{\varepsilon}(\boldsymbol{e})) \,\mathrm{d}x. \end{aligned} \tag{5.7}$$

Step 2. Since $\boldsymbol{\rho}_{\mathrm{RT}} \coloneqq \boldsymbol{\delta} - \widehat{\boldsymbol{\tau}}_{\mathrm{RT}} + \widehat{\boldsymbol{\tau}}_{\mathrm{RT}}^* - \boldsymbol{\tau}_{\mathrm{RT}} \in \boldsymbol{\Sigma}_{\mathbf{0}}(\widehat{\mathcal{T}})$ is divergence-free with $\boldsymbol{\rho}_{\mathrm{RT}}\boldsymbol{\nu} = \mathbf{0}$ on Γ_{N}, Proposition 2.3 yields existence of some $\widehat{\boldsymbol{\beta}}_{\mathrm{N}} \in N_0(\widehat{\mathcal{T}}; \mathbb{R}^{n\times n})$ with

$$\boldsymbol{\rho}_{\mathrm{RT}} = \operatorname{curl} \widehat{\boldsymbol{\beta}}_{\mathrm{N}} \quad \text{in } \Omega, \quad \operatorname{curl} \widehat{\boldsymbol{\beta}}_{\mathrm{N}} \cdot \boldsymbol{\nu} = \mathbf{0} \quad \text{on } \Gamma_{\mathrm{N}}, \quad \text{and} \quad \|\boldsymbol{\beta}_{\mathrm{N}}\|_{L^2(\Omega)} \lesssim \|\boldsymbol{\rho}_{\mathrm{RT}}\|_{L^2(\Omega)}.$$

This, the triangle inequality, and the stability estimates (5.3) imply

$$\begin{aligned} \|\boldsymbol{\beta}_{\mathrm{N}}\|_{L^2(\Omega)} \lesssim \|\boldsymbol{\rho}_{\mathrm{RT}}\|_{L^2(\Omega)} &\leq \|\boldsymbol{\delta}\|_{L^2(\Omega)} + \|\widehat{\boldsymbol{\tau}}_{\mathrm{RT}} - \widehat{\boldsymbol{\tau}}_{\mathrm{RT}}^*\|_{L^2(\Omega)} + \|\boldsymbol{\tau}_{\mathrm{RT}}\|_{L^2(\Omega)} \\ &\lesssim \|\boldsymbol{\delta}\|_{L^2(\Omega)} + \|(1 - \boldsymbol{\Pi}_0) \operatorname{div} \boldsymbol{\delta}\|_{L^2(\Omega)} + \|\boldsymbol{\Pi}_0 \operatorname{div} \boldsymbol{\delta}\|_{L^2(\Omega)} + \operatorname{osc}(\widehat{\boldsymbol{g}}_0, \mathcal{F}(\Gamma_{\mathrm{N}})) \\ &\lesssim \|\boldsymbol{\delta}\|_{H(\operatorname{div},\Omega)} + \operatorname{osc}(\widehat{\boldsymbol{g}}_0, \mathcal{F}(\Gamma_{\mathrm{N}})). \end{aligned}$$

Step 3. The split $\mathbb{C}^{-1}(\boldsymbol{\delta} - \boldsymbol{\tau}_{\mathrm{RT}}) = \mathbb{C}^{-1}\boldsymbol{\rho}_{\mathrm{RT}} - \mathbb{C}^{-1}(\widehat{\boldsymbol{\tau}}_{\mathrm{RT}}^* - \widehat{\boldsymbol{\tau}}_{\mathrm{RT}})$, elementary algebra, and (5.7) prove

$$\begin{aligned} LS(\mathbf{0}; \boldsymbol{\delta}, \boldsymbol{e}) = \|\operatorname{div} \boldsymbol{\delta}\|_{L^2(\Omega)}^2 &+ \int_\Omega (\mathbb{C}^{-1}\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\widehat{\boldsymbol{u}}_{\mathrm{LS}})) : (\mathbb{C}^{-1}(\boldsymbol{\delta} + \widehat{\boldsymbol{\tau}}_{\mathrm{RT}}^* - \widehat{\boldsymbol{\tau}}_{\mathrm{RT}}) - \boldsymbol{\varepsilon}(\boldsymbol{e})) \,\mathrm{d}x \\ + \int_\Omega (\boldsymbol{f}_0 + \operatorname{div} \boldsymbol{\sigma}_{\mathrm{LS}}) \cdot \operatorname{div} \boldsymbol{\delta} \,\mathrm{d}x &+ \int_\Omega (\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}})) : (\boldsymbol{\varepsilon}(\boldsymbol{e}) - \mathbb{C}^{-1}\boldsymbol{\rho}_{\mathrm{RT}}) \,\mathrm{d}x \\ + \int_\Omega (\mathbb{C}^{-1}\boldsymbol{\delta} - \boldsymbol{\varepsilon}(\boldsymbol{e})) &: \mathbb{C}^{-1}(\widehat{\boldsymbol{\tau}}_{\mathrm{RT}} - \widehat{\boldsymbol{\tau}}_{\mathrm{RT}}^*) \,\mathrm{d}x. \end{aligned}$$

The equation (5.6) plus elementary algebra with the L^2 orthogonal projection $\boldsymbol{\Pi}_0$ conclude the proof. ∎

LEMMA 5.4. *It holds that*

$$\int_\Omega (\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}})) : \mathbb{C}^{-1} \operatorname{curl} \boldsymbol{\beta}_{\mathrm{N}} \,\mathrm{d}x$$
$$\leqslant C_1\big(\|\boldsymbol{\delta}\|_{H(\mathrm{div},\Omega)} + \operatorname{osc}(\widehat{\boldsymbol{g}}_0, \mathcal{F}(\Gamma_{\mathrm{N}}))\big)\Bigg(\sum_{T\in\mathcal{R}}\Bigg(|T|^{2/3}\|\operatorname{curl}\mathbb{C}^{-2}\boldsymbol{\sigma}_{\mathrm{LS}}\|^2_{L^2(T)}$$
$$+ \sum_{F\in\mathcal{F}(T)} |T|^{1/3}\|[\mathbb{C}^{-1}(\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}}))]_F \times \boldsymbol{\nu}\|^2_{L^2(F)}\Bigg)\Bigg)^{1/2}.$$

Proof. The operator $\mathcal{I}_{\mathrm{N}} : N_0(\widehat{\mathcal{T}}; \mathbb{R}^{3\times 3}) \to N_0(\mathcal{T}; \mathbb{R}^{3\times 3})$ satisfies [28, Thm. 4.1]

$$(1 - \mathcal{I}_{\mathrm{N}})\widehat{\boldsymbol{\beta}}_{\mathrm{N}} = 0 \quad \text{on any } K \in \mathcal{R} \quad \text{and} \quad \|\mathcal{I}_{\mathrm{N}}\widehat{\boldsymbol{\beta}}_{\mathrm{N}}\|_{H(\mathrm{curl},\Omega)} \lesssim \|\widehat{\boldsymbol{\beta}}_{\mathrm{N}}\|_{H(\mathrm{curl},\Omega)}. \tag{5.8}$$

The quasi-interpolation operator $\mathcal{S}_{\mathrm{N}} : H(\mathrm{curl}, \Omega; \mathbb{R}^{3\times 3}) \to N_0(\mathcal{T}; \mathbb{R}^{3\times 3})$ from [24, Thm. 1] allows for a split (5.9) and a local approximation error estimate (5.10)

$$(1 - \mathcal{S}_{\mathrm{N}})(1 - \mathcal{I}_{\mathrm{N}})\widehat{\boldsymbol{\beta}}_{\mathrm{N}} = \nabla\phi + \boldsymbol{z} \tag{5.9}$$

for some $\phi \in H^1_{\mathrm{D}}(\Omega)$ and $\boldsymbol{z} \in H^1_{\mathrm{D}}(\Omega; \mathbb{R}^3)$ and for every $K \in \mathcal{T}$ with $\Omega_K := \bigcup\{T \in \mathcal{T} : T \cap K \neq \varnothing\}$ that

$$\|h_0^{-1}\boldsymbol{z}\|_{L^2(K)} + \|\mathbf{D}\,\boldsymbol{z}\|_{L^2(K)} \lesssim \|\operatorname{curl}(1 - \mathcal{I}_{\mathrm{N}})\widehat{\boldsymbol{\beta}}_{\mathrm{N}}\|_{L^2(\Omega_K)}. \tag{5.10}$$

Since $\operatorname{curl}\widehat{\boldsymbol{\beta}}_{\mathrm{N}} \in RT_0(\mathcal{T}; \mathbb{R}^{3\times 3})$ is divergence-free, the discrete equations (3.2) imply

$$\int_\Omega (\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(u_{\mathrm{LS}})) : \mathbb{C}^{-1}\operatorname{curl}\widehat{\boldsymbol{\beta}}_{\mathrm{N}}\,\mathrm{d}x$$
$$= \int_\Omega \mathbb{C}^{-1}(\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(u_{\mathrm{LS}})) : \operatorname{curl}(1 - \mathcal{S}_{\mathrm{N}})(1 - \mathcal{I}_{\mathrm{N}})\widehat{\boldsymbol{\beta}}_{\mathrm{N}}\,\mathrm{d}x.$$

The combination of the locality (5.8) of $\mathcal{I}_{\mathrm{N}}$ [28, Thm. 4.1(1)] and the local estimate (5.10) proves, for any $K \in \mathcal{R}$ and any $F \in \mathcal{F}(K)$, $\boldsymbol{z} \equiv \mathbf{0}$ in K and $\boldsymbol{z} \equiv 0$ on F in the sense of traces. This, (5.9) and an integration by parts result in

$$\int_\Omega \mathbb{C}^{-1}(\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}})) : \operatorname{curl}(1 - \mathcal{S}_{\mathrm{N}})(1 - \mathcal{I}_{\mathrm{N}})\widehat{\boldsymbol{\beta}}_{\mathrm{N}}\,\mathrm{d}x$$
$$= \sum_{T\in\mathcal{T}} \int_T \mathbb{C}^{-1}(\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}})) : \operatorname{curl}\boldsymbol{z}\,\mathrm{d}x$$
$$= \sum_{T\in\mathcal{T}} \int_T \operatorname{curl}\mathbb{C}^{-2}\boldsymbol{\sigma}_{\mathrm{LS}} : \boldsymbol{z}\,\mathrm{d}x$$
$$+ \sum_{F\in\mathcal{F}\setminus\mathcal{F}(\Gamma_{\mathrm{D}})} \int_F [\mathbb{C}^{-1}(\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}})) \times \boldsymbol{\nu}]_F : \boldsymbol{z}\,\mathrm{d}s$$
$$= \sum_{T\in\mathcal{R}} \int_T \operatorname{curl}\mathbb{C}^{-2}\boldsymbol{\sigma}_{\mathrm{LS}} : \boldsymbol{z}\,\mathrm{d}x$$
$$+ \sum_{F\in\mathcal{F}(\mathcal{R})\setminus\mathcal{F}(\Gamma_{\mathrm{D}})} \int_F [\mathbb{C}^{-1}(\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}})) \times \boldsymbol{\nu}]_F : \boldsymbol{z}\,\mathrm{d}s.$$

A Cauchy–Schwarz inequality and (5.10) prove, for every $T \in \mathcal{R}$,

$$\begin{aligned}\int_T &\operatorname{curl} \mathbb{C}^{-2}\boldsymbol{\sigma}_{\mathrm{LS}} : \boldsymbol{z}\,\mathrm{d}x\\ &\leqslant \left\| h_T \operatorname{curl} \mathbb{C}^{-2}\boldsymbol{\sigma}_{\mathrm{LS}} \right\|_{L^2(T)} \left\| h_T^{-1}\boldsymbol{z} \right\|_{L^2(T)}\\ &\lesssim \left\| h_T \operatorname{curl} \mathbb{C}^{-2}\boldsymbol{\sigma}_{\mathrm{LS}} \right\|_{L^2(T)} \left\| \operatorname{curl}(1-\mathcal{I}_{\mathrm{N}})\widehat{\boldsymbol{\beta}}_{\mathrm{N}} \right\|_{L^2(\Omega_T)}.\end{aligned}$$

A Cauchy–Schwarz inequality, a trace inequality, and (5.10) show, for every $F \in \mathcal{F}(\mathcal{R})\backslash\mathcal{F}(\Gamma_{\mathrm{D}})$, that

$$\begin{aligned}\int_F &[\mathbb{C}^{-1}(\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}})) \times \boldsymbol{\nu}]_F : \boldsymbol{z}\,\mathrm{d}s\\ &\lesssim \left\| h_T^{1/2}[\mathbb{C}^{-1}(\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}})) \times \boldsymbol{\nu}]_F \right\|_{L^2(F)} \left\| \operatorname{curl}(1-\mathcal{I}_{\mathrm{N}})\widehat{\boldsymbol{\beta}}_{\mathrm{N}} \right\|_{L^2(\Omega_T)}.\end{aligned}$$

The bounded overlap of the patches $(\Omega_T : T \in \mathcal{R})$, the stability estimate (5.8), and (5.4) conclude the proof. ∎

Remark 5.5. The quasi-interpolation operator $\mathcal{I}_{\mathrm{N}}$ has been established for arbitrary polynomial degrees in [28, Thm. 4.1] and allows for the generalization of Lemma 5.4 to higher-order discretizations.

Proof of Theorem 5.1. *Step* 1. Since $\boldsymbol{g}_0|_K = \widehat{\boldsymbol{g}}_0|_K$ for all $K \in \widehat{\mathcal{T}} \cap \mathcal{T}$,

$$\operatorname{osc}^2(\widehat{\boldsymbol{g}}_0, \mathcal{F}(\Gamma_{\mathrm{N}})) = \sum_{F \in \mathcal{F}(\mathcal{T}\backslash\widehat{\mathcal{T}})} |\bar{\omega}_F| \left\| \widehat{\boldsymbol{g}}_0 - \boldsymbol{g}_0 \right\|^2_{L^2(F)} \leqslant \eta^2(\mathcal{T}, \mathcal{R}). \tag{5.11}$$

Step 2. Step 1 and the stability estimate (2.6) yield a generic constant $C_2 \approx 1$ with

$$\left\| \mathbb{C}^{-1}(\widehat{\boldsymbol{\tau}}_{\mathrm{RT}} - \widehat{\boldsymbol{\tau}}^*_{\mathrm{RT}}) \right\|_{L^2(\Omega)} \leqslant C_2 \left(\left\| (1-\boldsymbol{\Pi}_0) \operatorname{div} \boldsymbol{\delta} \right\|_{L^2(\Omega)} + \eta(\mathcal{T}, \mathcal{R}) \right).$$

This, Lemma 5.3, the Cauchy–Schwarz inequality, and the upper bound $\|\mathbb{C}^{-1}\boldsymbol{\delta} - \boldsymbol{\varepsilon}(\boldsymbol{e})\|_{L^2(\Omega)} \leqslant LS(\boldsymbol{0}; \boldsymbol{\delta}, \boldsymbol{e})^{1/2}$ imply

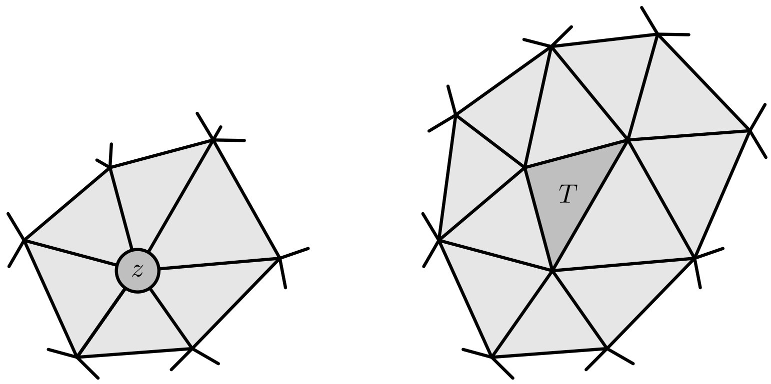
$$\begin{aligned}LS(\boldsymbol{0}; \boldsymbol{\delta}, \boldsymbol{e}) \leqslant{}& \left\| (1-\boldsymbol{\Pi}_0) \operatorname{div} \boldsymbol{\delta} \right\|^2_{L^2(\Omega)}\\ &+ C_2 LS(\boldsymbol{0}; \boldsymbol{\delta}, \boldsymbol{e})^{1/2} \left(\left\| (1-\boldsymbol{\Pi}_0) \operatorname{div} \boldsymbol{\delta} \right\|_{L^2(\Omega)} + \eta(\mathcal{T}, \mathcal{R}) \right)\\ &+ \int_\Omega (\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}})) : (\boldsymbol{\varepsilon}(\boldsymbol{e}) - \mathbb{C}^{-1} \operatorname{curl} \widehat{\boldsymbol{\beta}}_{\mathrm{N}})\,\mathrm{d}x.\end{aligned} \tag{5.12}$$

Step 3. Let $\mathcal{J}\boldsymbol{e} \in \boldsymbol{V}(\mathcal{T}; \mathbb{R}^3)$ denote the Scott–Zhang quasi-interpolation of $\boldsymbol{e}$. For every node $z \in \mathcal{N}$ in the construction of the quasi-interpolation [25, sect. 2], choose $E \in \mathcal{E}(\omega_z)$ such that $E \in \mathcal{E} \cap \widehat{\mathcal{E}}$, whenever possible. This ensures for the error function

$$(1-\mathcal{J})\boldsymbol{e} \equiv 0 \quad \text{on } \mathcal{T}\backslash\widehat{\mathcal{T}}, \tag{5.13}$$

i.e., $(1-\mathcal{J})\boldsymbol{e}$ vanishes on any triangle $T \in \mathcal{T} \cap \widehat{\mathcal{T}}$ and on any of its faces $F \in \mathcal{F}(T)$ in the sense of traces. The first-order approximation property [25, eq. (4.3)] and the stability property [25, Thm. 3.1] read

$$|T|^{-1/2} \left\| (1-\mathcal{J})\boldsymbol{e} \right\|_{L^2(T)} + \left\| \mathbf{D}(1-\mathcal{J})\boldsymbol{e} \right\|_{L^2(T)} \lesssim \left\| \mathbf{D}\,\boldsymbol{e} \right\|_{L^2(\Omega_T)} \tag{5.14}$$



(a) Nodal patch ω_z (b) Enlarged triangle patch Ω_T

FIG. 5.1. *Patches.*

for the enlarged triangle patch $\Omega_T := \bigcup_{z\in\mathcal{N}(T)} \omega_z$ of Figure 5.1. The discrete equations (3.2) imply

$$\int_\Omega (\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}})) : \boldsymbol{\varepsilon}(\mathcal{J}\boldsymbol{e})\,\mathrm{d}x = 0.$$

This, an integration by parts, and (5.13) prove

$$\begin{aligned}
&\int_\Omega (\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}})) : \boldsymbol{\varepsilon}(\boldsymbol{e})\,\mathrm{d}x\\
&= \int_\Omega (\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}})) : \boldsymbol{\varepsilon}((1-\mathcal{J})\boldsymbol{e})\,\mathrm{d}x\\
&= \int_\Omega (\operatorname{sym}\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}})) : \mathbf{D}(1-\mathcal{J})\boldsymbol{e}\,\mathrm{d}x\\
&= \sum_{T\in\mathcal{T}} \int_T (\operatorname{div}\operatorname{sym}\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}}) \cdot (1-\mathcal{J})\boldsymbol{e}\,\mathrm{d}x\\
&\qquad + \sum_{F\in\mathcal{F}\setminus\mathcal{F}(\Gamma_{\mathrm{D}})} \int_F [(\operatorname{sym}\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}}))\cdot\boldsymbol{\nu}_F]_F\,(1-\mathcal{J})\boldsymbol{e}\,\mathrm{d}s\\
&= \sum_{T\in\mathcal{T}\setminus\widehat{\mathcal{T}}} \int_T (\operatorname{div}\operatorname{sym}\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}}) \cdot (1-\mathcal{J})\boldsymbol{e}\,\mathrm{d}x\\
&\qquad + \sum_{F\in\mathcal{F}(\mathcal{T}\setminus\widehat{\mathcal{T}})\setminus\mathcal{F}(\Gamma_{\mathrm{D}})} \int_F [(\operatorname{sym}\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}}))\boldsymbol{\nu}_F]_F \cdot (1-\mathcal{J})\boldsymbol{e}\,\mathrm{d}s.
\end{aligned}$$

A Cauchy–Schwarz inequality followed by approximation and stability properties of the quasi-interpolation from (5.14) prove, for every $T\in\mathcal{T}\setminus\widehat{\mathcal{T}}$, that

$$\begin{aligned}
&\int_T (\operatorname{div}\operatorname{sym}\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}}) \cdot (1-\mathcal{J})\boldsymbol{e}\,\mathrm{d}x\\
&\qquad \leqslant \|h_T\operatorname{div}\operatorname{sym}\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}}\|_{L^2(T)} \|h_T^{-1}(1-\mathcal{J})\boldsymbol{e}\|_{L^2(T)}\\
&\qquad \lesssim \|h_T\operatorname{div}\operatorname{sym}\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}}\|_{L^2(T)} \|\mathbf{D}\,\boldsymbol{e}\|_{L^2(\Omega_T)}.
\end{aligned}$$

A Cauchy–Schwarz inequality, a trace inequality, and (5.14) show, for every $F \in \mathcal{F}(\mathcal{T}\backslash\widehat{\mathcal{T}})\backslash\mathcal{F}(\Gamma_{\mathrm{D}})$, that

$$\begin{aligned}&\int_F [(\operatorname{sym}\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}}))\boldsymbol{\nu}]_F \cdot (1-\mathcal{J})\boldsymbol{e}\,\mathrm{d}s \\ &\qquad \lesssim \big\| h_T^{1/2}[(\operatorname{sym}\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}}))\boldsymbol{\nu}]_F\big\|_{L^2(F)} \big\|\mathbf{D}\,\boldsymbol{e}\big\|_{L^2(\Omega_T)}.\end{aligned}$$

The bounded overlap of the patches $(\Omega_T : T \in \mathcal{T}\backslash\widehat{\mathcal{T}})$ in the two previously displayed formulas and the estimate $|\!|\!|\boldsymbol{e}|\!|\!|^2 \lesssim LS(\mathbf{0};\boldsymbol{\delta},\boldsymbol{e}) + \operatorname{osc}^2(\widehat{\boldsymbol{g}}_0,\mathcal{F}(\Gamma_{\mathrm{N}})) = \delta^2(\widehat{\mathcal{T}},\mathcal{T})$ show

$$\begin{aligned}&\int_\Omega (\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}})) : \boldsymbol{\varepsilon}(\boldsymbol{e})\,\mathrm{d}x \\ &\leqslant C_3\delta(\widehat{\mathcal{T}},\mathcal{T})\Bigg(\sum_{T\in\mathcal{T}\backslash\widehat{\mathcal{T}}}\Bigg(|T|^{2/3}\big\|\operatorname{div}\operatorname{sym}\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}}\big\|^2_{L^2(T)} \\ &\qquad + \sum_{F\in\mathcal{F}(T)\cap\mathcal{F}(\Omega)}|T|^{1/3}\big\|[(\operatorname{sym}\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}}))\boldsymbol{\nu}]_F\big\|^2_{L^2(F)}\Bigg)\Bigg)^{1/2}.\end{aligned} \tag{5.15}$$

Step 4. The triangle and Young inequalities and the L^2 orthogonality of $\boldsymbol{\Pi}_0$ prove

$$\begin{aligned}\frac{1}{2}\big\|(1-\boldsymbol{\Pi}_0)\operatorname{div}\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}}\big\|^2_{L^2(\Omega)} &\leqslant \big\|\boldsymbol{f}-\boldsymbol{\Pi}_0\boldsymbol{f}\big\|^2_{L^2(\Omega)} + \big\|(1-\boldsymbol{\Pi}_0)(\boldsymbol{f}+\operatorname{div}\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}})\big\|^2_{L^2(\Omega)} \\ &\leqslant \big\|\boldsymbol{f}-\boldsymbol{\Pi}_0\boldsymbol{f}\big\|^2_{L^2(\Omega)} + LS(\boldsymbol{f};\widehat{\boldsymbol{\sigma}}_{\mathrm{LS}},\widehat{\boldsymbol{u}}_{\mathrm{LS}}).\end{aligned}$$

Step 5. The combination of (5.12) with Lemma 5.4 and Step 3 reads

$$\begin{aligned}LS(\mathbf{0};\boldsymbol{\delta},\boldsymbol{e}) \leqslant{}& \big\|(1-\boldsymbol{\Pi}_0)\operatorname{div}\boldsymbol{\delta}\big\|^2_{L^2(\Omega)} + C_1\Big(\|\boldsymbol{\delta}\|_{H(\operatorname{div},\Omega)} + \operatorname{osc}(\widehat{\boldsymbol{g}}_0,\mathcal{F}(\Gamma_{\mathrm{N}}))\Big) \\ &\times\Bigg(\sum_{T\in\mathcal{R}}\Bigg(|T|^{2/3}\big\|\operatorname{curl}\mathbb{C}^{-2}\boldsymbol{\sigma}_{\mathrm{LS}}\big\|^2_{L^2(T)} \\ &\qquad + \sum_{F\in\mathcal{F}(T)}|T|^{1/3}\big\|[\mathbb{C}^{-1}(\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}}))]_F\times\boldsymbol{\nu}\big\|^2_{L^2(F)}\Bigg)\Bigg)^{1/2} \\ &+ C_2 LS(\mathbf{0};\boldsymbol{\delta},\boldsymbol{e})^{1/2}\big(\|(1-\boldsymbol{\Pi}_0)\operatorname{div}\boldsymbol{\delta}\|_{L^2(\Omega)} + \eta(\mathcal{T},\mathcal{R})\big) \\ &+ C_3\delta(\widehat{\mathcal{T}},\mathcal{T})\Bigg(\sum_{T\in\mathcal{T}\backslash\widehat{\mathcal{T}}}\Bigg(|T|^{2/3}\big\|\operatorname{div}\operatorname{sym}\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}}\big\|^2_{L^2(T)} \\ &\qquad + \sum_{F\in\mathcal{F}(T)\cap\mathcal{F}(\Omega)}|T|^{1/3}\big\|[(\operatorname{sym}\mathbb{C}^{-1}\boldsymbol{\sigma}_{\mathrm{LS}} - \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{LS}}))\boldsymbol{\nu}]_F\big\|^2_{L^2(F)}\Bigg)\Bigg)^{1/2}.\end{aligned}$$

Steps 1 and 4 plus some standard rearrangements with the Young inequality conclude the proof. ∎

REFERENCES

[1] P. Binev, W. Dahmen, and R. DeVore, *Adaptive finite element methods with convergence rates*, Numer. Math., 97 (2004), pp. 219–268.
[2] P. Binev and R. DeVore, *Fast computation in adaptive tree approximation*, Numer. Math., 97 (2004), pp. 193–217.
[3] P. B. Bochev and M. D. Gunzburger, *Least-Squares Finite Element Methods*, Appl. Math. Sci. 166, Springer, New York, 2009.
[4] D. Boffi, F. Brezzi, and M. Fortin, *Mixed Finite Element Methods and Applications*, Springer Ser. Comput. Math., Springer, New York, 2013.
[5] D. Braess, *Finite Elements: Theory, Fast Solvers, and Applications in Elasticity Theory*, 3rd ed., Cambridge University Press, Cambridge, UK, 2007.
[6] J. Brandts, Y. Chen, and J. Yang, *A note on least-squares mixed finite elements in relation to standard and mixed finite elements*, IMA J. Numer. Anal., 26 (2006), pp. 779–789.
[7] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, 3rd ed., Texts in Appl. Math. 15, Springer, New York, 2008.
[8] P. Bringmann and C. Carstensen, *An adaptive least-squares FEM for the Stokes equations with optimal convergence rates*, Numer. Math., 135 (2017), pp. 459–492.
[9] Z. Cai and G. Starke, *Least-squares methods for linear elasticity*, SIAM J. Numer. Anal., 42 (2004), pp. 826–842.
[10] C. Carstensen, M. Feischl, M. Page, and D. Praetorius, *Axioms of adaptivity*, Comput. Math. Appl., 67 (2014), pp. 1195–1253.
[11] C. Carstensen, J. Gedicke, and D. Rim, *Explicit error estimates for Courant, Crouzeix-Raviart and Raviart-Thomas finite element methods*, J. Comput. Math., 30 (2012), pp. 337–353.
[12] C. Carstensen and E.-J. Park, *Convergence and optimality of adaptive least squares finite element methods*, SIAM J. Numer. Anal., 53 (2015), pp. 43–62.
[13] C. Carstensen, E. J. Park, and P. Bringmann, *Convergence of natural adaptive least squares fems*, Numer. Math., 136 (2017), pp. 1097–1115, https://doi.org/10.1007/s00211-017-0866-x.
[14] C. Carstensen and H. Rabus, *An optimal adaptive mixed finite element method*, Math. Comp., 80 (2011), pp. 649–667.
[15] C. Carstensen and H. Rabus, *Axioms of adaptivity with separate marking for data resolution*, SIAM J. Numer. Anal., 55 (2017), pp. 2644–2665.
[16] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, Stud. Math. Appl. 4, North-Holland, Amsterdam, 1978.
[17] M. Dauge, *Elliptic Boundary Value Problems on Corner Domains: Smoothness and Asymptotics of Solutions*, Lecture Notes in Math. 1341, Springer, New York, 1988.
[18] A. Ern and J.-L. Guermond, *Mollification in strongly Lipschitz domains with application to continuous and discrete de Rham complexes*, Comput. Methods Appl. Math., 16 (2016), pp. 51–75.
[19] A. Ern and J.-L. Guermond, *Analysis of the Edge Finite Element Approximation of the Maxwell Equations with Low Regularity Solutions*, arXiv:1706.00600, 2017.
[20] V. Girault and P.-A. Raviart, *Finite Element Methods for Navier-Stokes Equations: Theory and Algorithms*, Springer Ser. Comput. Math. 5, Springer, New York, 1986.
[21] J.-C. Nédélec, *Mixed finite elements in* $\mathbf{R}^3$, Numer. Math., 35 (1980), pp. 315–341.
[22] J.-C. Nédélec, *Éléments finis mixtes incompressibles pour l'équation de Stokes dans* $\mathbf{R}^3$, Numer. Math., 39 (1982), pp. 97–112.
[23] J. Schöberl, *A multilevel decomposition result in H(curl)*, in Multigrid, Multilevel and Multiscale Methods, P. Wesseling, C. Oosterlee, and P. Hemker, eds., TU Delft, 2005.
[24] J. Schöberl, *A posteriori error estimates for Maxwell equations*, Math. Comp., 77 (2008), pp. 633–649.
[25] L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493.
[26] R. Stevenson, *The completion of locally refined simplicial partitions created by bisection*, Math. Comp., 77 (2008), pp. 227–241 (electronic).
[27] R. Verfürth, *A Posteriori Error Estimation Techniques for Finite Element Methods*, Numer. Math. Sci. Comput., Oxford University Press, Oxford, UK, 2013.
[28] L. Zhong, L. Chen, S. Shu, G. Wittum, and J. Xu, *Convergence and optimality of adaptive edge FEMs for time-harmonic Maxwell equations*, Math. Comp., 81 (2012), pp. 623–642.