A UNIFYING THEORY OF A POSTERIORI FINITE ELEMENT ERROR CONTROL

C. CARSTENSEN

Abstract. Residual-based a posteriori error estimates are derived within a unified setting for lowest-order conforming, nonconforming, and mixed finite element schemes. The various residuals are identified for all techniques and problems as the operator norm $\|\ell\|$ of a linear functional of the form

$$\ell(v) := \int_\Omega p_h : Dv\,dx + \int_\Omega g_\Omega \cdot v\,dx$$

in the variable v of a Sobolev space V. The main assumption is that the first-order finite element space $\mathcal{S}^1_0(\Omega) \subset \ker \ell \subset V$ is included in the kernel $\ker \ell$ of ℓ. As a consequence, *any residual estimator* that is a computable bound of $\|\ell\|$ can be used within the proposed frame *without* further analysis for nonconforming or mixed FE schemes. Applications are given for the Laplace, Stokes, and Navier-Lamè equations.

1. Unifying Theory of A Posteriori Error Control

This section sets up an abstract framework for a posteriori estimation which is filled with details for low-order finite element methods for the Laplace, Stokes, and Navier-Lamè equations in Section 2, 3, and 4, respectively. This unifying approach generalizes known techniques based on a Helmholtz decomposition [A, C, CD, CF1, CBK, BC, DDVP] as well as comparison schemes [HW]. The final result of the presented theory is that *one unified* type of residuals has to be analyzed *once* and the resulting estimator can be simultaneously used for a posteriori error control of conforming, nonconforming, mixed, or other finite element schemes.

1.1. **Residual-Based A Posteriori Error Control.** Let $A : X \to X^*$ be a linear and bounded operator between the (real) Banach spaces X and its dual X^*. Suppose that A is surjective and injective such

---

Date: March 2, 2004.

Key words and phrases. A posteriori, error analysis, finite element method, nonconforming finite element method, mixed finite element method, adaptive algorithm.

that A^{-1} is bounded as well. In particular, given $y \in X^*$, there exists a unique $x \in X$ such that

$$Ax = y. \tag{1.1}$$

Suppose we are given some finite element approximation $x_h \in X$ (possibly with some additional properties generated by the computational scheme that provided x_h). Then we address the issue of approximating the error $\|x - x_h\|$. Notice that the linear functional

$$\mathcal{R}es := y - Ax_h = A(x - x_h) \in X^* \tag{1.2}$$

is known or, at least computable. Throughout this paper, an inequality $a \lesssim b$ replaces $a \leq c\,b$ with a multiplicative mesh-size independent constant c that depends only on the domain Ω and the shape (e.g. through the aspect ratio) of finite elements. Finally, $a \approx b$ abbreviates $a \lesssim b \lesssim a$. Since the operator norms $\|A\| \approx 1 \approx \|A^{-1}\|$ of A and A^{-1} are uniformly bounded, there holds

$$\|x - x_h\|_X \approx \|A(x - x_h)\|_{X^*} \approx \|\mathcal{R}es\|_{X^*}. \tag{1.3}$$

Hence, any residual-based a posteriori error control means the approximation of lower and upper bounds of the dual norm $\|\mathcal{R}es\|_{X^*}$ of $\mathcal{R}es$. Throughout the paper, $X = H \times L$ will (essentially) be fixed and the discrete subspaces vary.

1.2. **Goal-Oriented Error Control.** The analysis of this paper focuses on the estimation of the *norm* $\|x - x_h\|_X$. In some applications, there is a given (hence known) linear and bounded *functional* $\rho : X \to \mathbb{R}$ that monitors the error $|\rho(x - x_h)|$ (e.g. the error of an averaged strain or traction over a small but fixed region). To assess the latter quantity, let $A^* : X \to X^*$ be the dual operator of A (for reflexive spaces $X = X^{**}$) and let $z \in X$ be the solution to

$$A^*z = \rho.$$

Then it remains to estimate

$$\rho(x - x_h) = (A^*z)(x - x_h) = (A(x - x_h))z = \mathcal{R}es(z).$$

An immediate consequence of this reads

$$|\rho(x - x_h)| \leq \|\mathcal{R}es\|_{X^*}\|z\|_X \approx \|\rho\|_{X^*}\|\mathcal{R}es\|_{X^*}. \tag{1.4}$$

This global estimate (1.4) is (a) presumably too coarse and (b) does not convey local information of ρ via z. But it indicates that the evaluation of $\mathcal{R}es(z)$ may follow localized arguments from the assessment of $\|\mathcal{R}es\|_{X^*}$ addressed in this paper (cf. [BR1, AO, BaS]).

1.3. **Mixed Approach to Flux or Stress Error Control.** The primal variable $u \in H$ (e.g. the displacement field) is accompanied by a dual variable $p \in L$ (e.g. the flux or stress). The pair $(p, u) =: x \in L \times H =: X$ plays the role of the variable x in Subsection 1.1 above. Below, L will be a Lebesgue space (e.g. $L = L^2(\Omega)^n$) and H will be a Sobolev space (e.g. $H = H_0^1(\Omega)$), defined on a bounded domain Ω in $\mathbb{R}^n$. At the moment, it suffices to consider L and H as reflexive Banach spaces. The linear operator $A : X \to X^*$ is defined via a mixed framework, namely,

$$(A(p, u))(q, v) := a(p, q) + b(p, v) + b(q, u) \tag{1.5}$$

for bounded bilinear forms $a : L \times L \to \mathbb{R}$ and $b : L \times H \to \mathbb{R}$. Under well-analyzed conditions on a and b [B, BF] the operator A is bijective. Hence, given right-hand sides $f \in L^*$ and $g \in H^*$ with $y \in X^*$ defined by $y(q, v) = f(q) + g(v)$, there exists a unique $x = (p, u) \in L \times H$ that solves (1.1). Let $x_h = (p_h, \tilde{u}_h) \in L \times H$ be an approximation to x and define $\mathcal{R}es$ as $\mathcal{R}es_L + \mathcal{R}es_H$ by (1.2), namely

$$\begin{aligned} \mathcal{R}es_L(q) &:= f(q) - a(p_h, q) - b(q, \tilde{u}_h) \quad \text{for } q \in L, \\ \mathcal{R}es_H(v) &:= g(v) - b(p_h, v) \quad \text{for } v \in H. \end{aligned} \tag{1.6}$$

The notation $\tilde{u}_h \in H$ here and below asserts that $\tilde{u}_h$ is a continuous and not necessarily a discrete function; the subindex in $\tilde{u}_h$ refers to the fact that $\tilde{u}_h$ is closely related to u_h and is on our disposal. With (1.5)-(1.6), Equivalence (1.3) becomes

$$\|p - p_h\|_L + \|u - \tilde{u}_h\|_H \approx \|\mathcal{R}es_L\|_{L^*} + \|\mathcal{R}es_H\|_{H^*}. \tag{1.7}$$

This is the starting point of the unifying theory. The fact that $\tilde{u}_h$ has to belong to H is a crucial point in the sequel.

Remark 1.1. For non-conforming or mixed finite element schemes we obtain an approximation u_h to u which, below, is *not* in H. Consequently, $\tilde{u}_h \in H$ is, in general, different from u_h. To achieve an error estimation of the dual variable $p - p_h$, we will choose $\tilde{u}_h$ properly. The choice $\tilde{u}_h = u$ might be possible and minimizes $\|u - \tilde{u}_h\|$ but, in general, leads to difficulties in the evaluation of $b(q, \tilde{u}_h)$ in $\mathcal{R}es_L$ in $(1.6)_a$.

Remark 1.2. It should be notified clearly that (1.5) is a *primal mixed* formulation, also called hybrid in [BF], where L is *not* a subspace of $H(\mathrm{div}; \Omega)$. This is because the derivatives act on $u \in H = H_0^1(\Omega)^m$. For instance, the bilinear form b looks like

$$b(q, v) = -\int_\Omega q \cdot \nabla v \, dx \quad \text{for the Laplace equation}$$

(without an integration by parts as for the dual mixed formulation) and similar expressions hold for Stokes and Navier-Lamè equations.

1.4. **Residuals.** It is an aim of this paper to emphasize that there is essentially only one type of residual that arise in a posteriori error control: Given some $g \in L^2(\Omega)^m$ and some $g_\mathcal{E} \in L^2(\cup\mathcal{E})^m$ one encounters the linear functional $\mathrm{Res} : H_0^1(\Omega)^m \to \mathbb{R}$,

$$\mathrm{Res}(v) := \int_\Omega g \cdot v \, dx + \int_{\cup\mathcal{E}} g_\mathcal{E} \cdot v \, ds \tag{1.8}$$

with the property that

$$\mathcal{S}_0^1(\mathcal{T})^m \subseteq \ker \mathrm{Res} \tag{1.9}$$

for globally continuous and piecewise polynomials $\mathcal{S}^1(\mathcal{T})^m$ of (partial or total) degree ≤ 1 and $\mathcal{S}_0^1(\mathcal{T}) := \mathcal{S}^1(\mathcal{T}) \cap H_0^1(\Omega)$. (Details on the notation follow in Section 2). The norm of Res enters as an explicit bound in the form of

$$\|\mathrm{Res}\| := \sup_{v \in H_0^1(\Omega)^m \setminus \{0\}} \mathrm{Res}(v)/\|\nabla v\|_{L^2(\Omega)}$$

or, for $m = 3 = n$ and nonconforming terms, in the form

$$\|\mathrm{Res}\| := \sup_{v \in H^1(\Omega)^3,\, \mathrm{Curl}\, v \not\equiv 0} \mathrm{Res}(v)/\|\,\mathrm{Curl}\, v\|_{L^2(\Omega)}.$$

It will be seen in Section 2, 3, and 4 that all arising residuals can be written and hence estimated in this unified form (1.8)-(1.9).

2. Application to Laplace Equation

This section is devoted to the Poisson problem as the simplest elliptic PDE and its residual-based a posteriori finite element error control. Subsection 2.1 introduces the model problem and some required notation while Subsection 2.2, 2.3, and 2.4 concern technical details in increasing difficulty for the conforming, non-conforming, and mixed low-order finite element methods. An application to discontinuous Galerkin schemes is in preparation.

2.1. **Model Problem.** Throughout this paper, Ω denotes a bounded Lipschitz domain in $\mathbb{R}^n$ with piecewise flat boundary $\partial\Omega$ such that $\overline{\Omega}$ is the union of a regular triangulation $\mathcal{T}$, $\overline{\Omega} = \cup\mathcal{T}$ (no hanging nodes). The Lebesgue and Sobolev spaces $L^2(\Omega)$ and $H^1(\Omega)$ are defined as usual and we define

$$L := L^2(\Omega)^n \text{ and } H := H_0^1(\Omega) := \{w \in H^1(\Omega) :\, w = 0 \text{ on } \partial\Omega\}. \tag{2.1}$$

Then, the gradient operator, $\nabla : H \to L$, maps H into L. Given $g \in L^2(\Omega)$ let $u \in H$ denote the solution to the *Poisson Problem*

$$\Delta u + g = 0 \text{ in } \Omega \quad \text{and} \quad u = 0 \text{ on } \partial\Omega. \tag{2.2}$$

Then, the flux $p := \nabla u \in L$ and $u \in H$ solve the problem

$$\begin{aligned}(A(p,u))(q,v) &:= a(p,q) + b(p,v) + b(q,u)\\ &\overset{!}{=} -\int_\Omega gv\,dx \quad \text{for all } (q,v) \in X = L \times H\end{aligned} \tag{2.3}$$

of the form considered in Subsection 1.3 with

$$a(p,q) := \int_\Omega p \cdot q\,dx \quad \text{and} \quad b(p,v) := -\int_\Omega p \cdot \nabla v\,dx. \tag{2.4}$$

Theorem 2.1. *The operator $A : X \to X^*$ defined in (2.3)$_a$ is bounded, linear, and bijective. For any $p_h \in L$ and $\tilde{u}_h \in H$ there holds*

$$\|p - p_h\|_L + \|u - \tilde{u}_h\|_H \approx \|\mathcal{R}es_L\|_{L^*} + \|\mathcal{R}es_H\|_{H^*} \tag{2.5}$$

for $\mathcal{R}es_L \in L^*$ and $\mathcal{R}es_H \in H^*$ defined for all $q \in L$ and $v \in H$ by

$$\begin{aligned}\mathcal{R}es_L(q) &:= \int_\Omega q \cdot (\nabla\tilde{u}_h - p_h)\,dx \quad \textit{and}\\ \mathcal{R}es_H(v) &:= -\int_\Omega gv\,dx + \int_\Omega p_h \cdot \nabla v\,dx.\end{aligned} \tag{2.6}$$

Proof. The assertions on A are well known; a direct proof of an inf-sup condition follows for any $(p,u) \in L \times H$ from $(q,v) := (p - \nabla u; -2u) \in L \times H$ and

$$\begin{aligned}1/6\,\|(p,u)\|_X\|(q,v)\|_X &\le 1/6\,(\|p\|_L + \|u\|_H)(\|p\|_L + 3\|u\|_H)\\ &\le \|p\|_L^2 + \|u\|_H^2 = (A(p,u))(q,v).\end{aligned}$$

The (generalised) Lax-Milgram lemma then yields bijectivity of A. The remaining assertions follow with the arguments of Subsection 1.3 which lead to (1.7) which, here, reads (2.5)-(2.6). □

2.2. **Conforming Finite Element Methods.** The aforementioned triangulation $\mathcal{T}$ into triangles or parallelograms for 2D and into tetrahedra or parallelepipeds for 3D is the basis of the conforming low-order finite element space $H_h = \mathcal{S}_0^1(\mathcal{T})$. Let $\mathcal{P}_k(T) = P_k(T)$ and $\mathcal{P}_k(T) = Q_k(T)$ for a triangle (or tetrahedron) and parallelogram (or parallelepiped), respectively, and the space $P_k(T)$ and $Q_k(T)$ algebraic polynomials of total and partial degree $\le k$, respectively, and define

$$\begin{aligned}&\mathcal{L}^k(\mathcal{T}) := \{v \in L^2(\Omega) : \forall T \in \mathcal{T},\ v|_T \in \mathcal{P}_k(T)\} \quad \text{for } k = 0,1;\\ &\mathcal{S}^1(\mathcal{T}) := \mathcal{L}^1(\mathcal{T}) \cap C(\overline{\Omega}) \quad \text{and} \quad \mathcal{S}_0^1(\mathcal{T}) := \mathcal{S}^1(\mathcal{T}) \cap H_0^1(\Omega).\end{aligned} \tag{2.7}$$

Let $\mathcal{N}$ denote the set of nodes (i.e. the vertices of elements in $\mathcal{T}$) and let $\mathcal{E}$ denote the edges in 2D (or faces in 3D) in $\mathcal{T}$. Let $h_\mathcal{T}$ and $h_\mathcal{E}$ be $\mathcal{T}$- and $\mathcal{E}$-piecewise constant on Ω and $\cup\mathcal{E} = \cup_{E\in\mathcal{E}}E$ defined by $h_\mathcal{T}|_T := h_T := \operatorname{diam}(T)$ and $h_\mathcal{E}|_E := h_E := \operatorname{diam}(E)$ for $T \in \mathcal{T}$ and $E \in \mathcal{E}$, respectively.

Given $u_h \in H_h$ with

$$\int_\Omega \nabla u_h \cdot \nabla v_h \, dx = \int_\Omega g v_h \, dx \quad \text{for all } v_h \in H_h \tag{2.8}$$

and $p_h := \nabla u_h$ as an approximation to $p := \nabla u$, we aim to estimate $\|p - p_h\|_L$. With Theorem 2.1 and $\tilde{u}_h = u_h \in H_h \subset H$ there holds

$$\mathcal{R}es_L \equiv 0 \quad \text{and} \quad \mathcal{R}es_H(v) = -\int_\Omega gv\,dx + \int_\Omega p_h \cdot \nabla v\,dx. \tag{2.9}$$

Notice that an elementwise integration by parts shows that $\mathcal{R}es_H$ is of the form (1.8)-(1.9). The evaluation of the residual $\mathcal{R}es_H$, namely the estimation of lower and upper bounds of

$$\|\mathcal{R}es_H\|_{H^*} := \sup_{v\in H\setminus\{0\}} (\int_\Omega gv\,dx - \int_\Omega p_h \cdot \nabla v\,dx)/\|v\|_H, \tag{2.10}$$

is subject of a vast literature. Although possibly sometimes not stated explicitly in this form, it is in fact the content of the books [V2, EJ, AO, BaS]. The point in this paper is that any of the (energy error) estimators thereof can be used. The standard explicit estimator reads

$$\eta_R^{(1)} := \|h_\mathcal{T}(g + \operatorname{div}_\mathcal{T} p_h)\|_{L^2(\Omega)} + \|h_\mathcal{E}^{1/2}[p_h \cdot \nu_\mathcal{E}]\|_{L^2(\cup\mathcal{E}_\Omega)} \tag{2.11}$$

and can be refined [CV] for $g \in H^1(\Omega)$ to

$$\eta_R^{(2)} := \|h_\mathcal{T}^2 \nabla g\|_{L^2(\Omega)} + \|h_\mathcal{E}^{1/2}[p_h \cdot \nu_\mathcal{E}]\|_{L^2(\cup\mathcal{E}_\Omega)}. \tag{2.12}$$

Another simple and easy-to-evaluate estimate is based on gradient-recovery: For any node $z \in \mathcal{N}$ with patch $\omega_z := \operatorname{int}(\cup\{T \in \mathcal{T} : z \in T\})$ let $A_z p_h := \int_{\omega_z} p_h\,dx/|\omega_z| \in \mathbb{R}^n$ be the average of p_h on ω_z. With the nodal basis function φ_z (defined by $\varphi_z \in \mathcal{S}^1(\mathcal{T})$ and $\varphi_z(z) = 1$ and $\varphi_z(x) = 0$ for all $x \in \mathcal{N}\setminus\{z\}$) let

$$\eta_A := \|h_\mathcal{T}^2 \nabla g\|_{L^2(\Omega)} + \|p_h - \sum_{z\in\mathcal{N}} (A_z p_h)\varphi_z\|_{L^2(\Omega)}. \tag{2.13}$$

There holds (assuming $g \in H^1(\Omega)$ and $\|h_\mathcal{T}^2 \nabla g\|_{L^2(\Omega)} \lesssim \|\mathcal{R}es_H\|_{H^*}$)

$$\|\mathcal{R}es_H\|_{H^*} \approx \eta_R^{(1)} \approx \eta_R^{(2)} \approx \eta_A. \tag{2.14}$$

Remark 2.1. Notice that $\|h_\mathcal{T}^2 \nabla g\|_{L^2(\Omega)}$ is of higher order for the low-order finite element scheme analyzed in this paper.

Remark 2.2. The proof of $\|p-p_h\|_L \lesssim \eta_R^{(1)}$ goes back to [BaR, BaM, EJ], the proof of $\eta_R^{(1)} \lesssim \|p-p_h\|_L + \|h_{\mathcal{T}}^2 \nabla g\|_{L^2(\Omega)}$ to [BaM, V2].

Remark 2.3. There are more expensive implicit error estimates, cf. [AO, V2, BaS, CBK].

2.3. **Nonconforming Finite Element Methods.** Based on the regular triangulation $\mathcal{T}$ into simplices (no parallelograms), the non-conforming finite element schemes due to Crouzeix-Raviart reads

$$\mathcal{S}_0^{1,NC}(\mathcal{T}) := \{v \in \mathcal{L}^1(\mathcal{T}) : \ v \text{ continuous at } \mathcal{M} \cap \Omega \\ \text{and } v = 0 \text{ at } \mathcal{M} \cap \partial\Omega\} \tag{2.15}$$

where $\mathcal{M}$ is the set of midpoints of edges (of faces) $E \in \mathcal{E}$. Notice that

$$\mathcal{S}_0^1(\mathcal{T}) \subset \mathcal{S}_0^{1,NC}(\mathcal{T}) \subset H^1(\mathcal{T}) := \{v \in L^2(\Omega) : \forall T \in \mathcal{T}, \ v|_T \in H^1(T)\}$$

and this is, in general, $\not\subset H^1(\Omega)$. Let $\nabla_{\mathcal{T}}$ (resp. $\nabla_{\mathcal{T}}^k$) denote the $\mathcal{T}$-piecewise action of the gradient operator (resp. the matrix of all partial derivatives of order k). Then, for any $v_h \in \mathcal{S}_0^{1,NC}(\mathcal{T})$, $\nabla_{\mathcal{T}} v_h \in L$. The finite element solution $u_h \in \mathcal{S}_0^{1,NC}(\mathcal{T})$ is the unique solution to

$$\int_\Omega \nabla_{\mathcal{T}} u_h \cdot \nabla_{\mathcal{T}} v_h \, dx = \int_\Omega g v_h \, dx \quad \text{for all } v_h \in \mathcal{S}_0^{1,NC}(\mathcal{T}). \tag{2.16}$$

The aim is to estimate the flux error $p - p_h$ for the discrete flux $p_h := \nabla_{\mathcal{T}} u_h \in L$. One difficulty is that, in general, $u_h \notin H$ and so $\tilde{u}_h$ cannot be chosen as in Subsection 2.2. However, for any $\tilde{u}_h \in H$, Theorem 2.1 yields

$$\|p - p_h\|_L + \|u - \tilde{u}_h\|_H \approx \|\mathcal{R}es_L\|_{L^*} + \|\mathcal{R}es_H\|_{H^*} \tag{2.17}$$

with $\mathcal{R}es_H$ treated as in (2.9)-(2.14) (notice that (2.16) guarantees (1.9) because of $\mathcal{S}_0^1(\mathcal{T}) \subset \mathcal{S}_0^{1,NC}(\mathcal{T})$) and

$$\mathcal{R}es_L(q) := \int_\Omega (\nabla \tilde{u}_h - p_h) \cdot q \, dx \quad \text{for all } q \in L. \tag{2.18}$$

The focus of the remaining part of this subsection is therefore on

$$\|\mathcal{R}es_L\|_{L^*} = \sup_{q \in L \setminus \{0\}} \mathcal{R}es_L(q)/\|q\|_L = \|p_h - \nabla \tilde{u}_h\|_{L^2(\Omega)} \tag{2.19}$$

and so on a proper choice of $\tilde{u}_h$.

Remark 2.4. A direct approach towards an upper bound of (2.19) is to compute some $\tilde{u}_h := I u_h \in \mathcal{S}_0^1(\mathcal{T})$ from the nonconforming finite element solution u_h; cf. [HW] for an example of I.

Remark 2.5. The minimizing $\tilde{u}_h$ in (2.19) within the class of conforming finite element approximations is equal to the conforming finite element approximation u_h^c of Subsection 2.2. [Proof: Since $\int_\Omega p_h \cdot \nabla v_h^c\, dx = \int_\Omega g v_h^c\, dx = \int_\Omega \nabla u_h^c \cdot \nabla v_h^c\, dx$ for all $v_h^c \mathcal{S}_0^1(\mathcal{T}) \subset \mathcal{S}_0^{1,NC}(\mathcal{T})$ there holds $\int_\Omega (p_h - \nabla u_h^c) \cdot \nabla v_h^c\, dx = 0$.] The split $\nabla_\mathcal{T} u_h = \nabla u_h^c + \mathrm{Curl}_\mathcal{T}\, b_h$, due to [AF], is known as the discrete Helmholtz decomposition. [In comparison with this, the version of [AF] interchanges the role of nonconforming and conforming terms; a change of the two components plus one change of signs proves the two versions equivalent.]

Definition 2.1. Given $p_h := \nabla_\mathcal{T} u_h$ define the linear functional

$$\mathcal{R}es_{NC} : H^1(\Omega)^m \to \mathbb{R} \quad \text{by} \quad \mathcal{R}es_{NC}(v) := \int_\Omega p_h \cdot \mathrm{Curl}\, v\, dx$$

for $v \in H^1(\Omega)^m$ and for $m := 1$ if $n = 2$ and $m := 3$ if $n = 3$. Set

$$\|\mathcal{R}es_{NC}\| := \sup_{v \in H^1(\Omega)^m,\, \mathrm{Curl}\, v \neq 0} \int_\Omega p_h \cdot \mathrm{Curl}\, v\, dx / \|\mathrm{Curl}\, v\|_L.$$

The following result relates $\|\mathcal{R}es_{NC}\|$ to (1.8)-(1.9). Notice carefully that $\tilde{u}_h$ is an arbitrary element in H (and not necessarily some discrete function).

Theorem 2.2. *There holds*

$$\min_{\tilde{u}_h \in H} \|p_h - \nabla \tilde{u}_h\|_L = \|\mathcal{R}es_{NC}\| \quad \text{and} \quad \mathcal{S}^1(\mathcal{T})^m \subset \ker \mathcal{R}es_{NC}.$$

Proof. The Helmholtz decomposition

$$p_h = \nabla a + \mathrm{Curl}\, b \tag{2.20}$$

holds for a unique $a \in H_0^1(\Omega) = H$ and some $b \in H^1(\Omega)^m$ with the L^2 orthogonality of $\mathrm{Curl}(H^1(\Omega)^m)$ and $\nabla(H)$. Then

$$\begin{aligned} \min_{\tilde{u}_h \in H} \|p_h - \nabla \tilde{u}_h\|_L^2 &= \|\mathrm{Curl}\, b\|_L^2 \\ &= \int_\Omega p_h \cdot \mathrm{Curl}\, b\, dx = \mathcal{R}es_{NC}(b) \\ &\leq \|\mathcal{R}es_{NC}\|\, \|\mathrm{Curl}\, b\|_L. \end{aligned} \tag{2.21}$$

For any $v \in H^1(\Omega)^m$ there holds, with (2.20) and $\int_\Omega \nabla a \cdot \mathrm{Curl}\, v\, dx = 0$, that

$$\begin{aligned} \mathcal{R}es_{NC}(v) &= \int_\Omega (\nabla a + \mathrm{Curl}\, b) \cdot \mathrm{Curl}\, v\, dx \\ &= \int_\Omega \mathrm{Curl}\, b \cdot \mathrm{Curl}\, v\, dx \\ &\leq \|\mathrm{Curl}\, v\|_L \|\mathrm{Curl}\, b\|_L. \end{aligned} \tag{2.22}$$

The combination of (2.21)-(2.22) shows

$$\|\mathcal{R}es_{NC}\| = \|\operatorname{Curl} b\|_L = \min_{\tilde{u}_h \in H} \|p_h - \nabla \tilde{u}_h\|_L.$$

This proves the first assertion of the theorem. The second is simpler in 2D and so solely shown for $n = m = 3$ and $v_h \in \mathcal{S}^1(\mathcal{T})^3$. An elementwise integration by parts yields

$$\begin{aligned}\mathcal{R}es_{NC}(v_h) &= \sum_{T\in\mathcal{T}} \int_T \nabla u_h \cdot \operatorname{Curl} v_h \, dx \\ &= \sum_{T\in\mathcal{T}} \int_{\partial T} u_h (\operatorname{Curl} v_h) \cdot \nu \, ds \\ &= \sum_{E\in\mathcal{E}} \int_E [u_h \operatorname{Curl} v_h] \cdot \nu_E \, ds\end{aligned}$$

for the unit normal ν_E on the element face E and the jump $[\cdot]$ across E. Recall that $\int_E [u_h]\, ds = 0$ by construction of the nonconforming finite element space $\mathcal{S}_0^{1,NC}(\mathcal{T})$. We claim that $[\operatorname{Curl} v_h] \cdot \nu_E = 0$ on an interior face $E \in \mathcal{E}_\Omega := \{E \in \mathcal{E} : E \not\subset \partial\Omega\}$. Since v_h is a polynomial on T_+ and $T_- \in \mathcal{T}$, $E = T_+ \cap T_-$, and continuous along E, there holds ($\otimes$ denotes the dyadic product)

$$Dv_h|_{T_+} - Dv_h|_{T_-} = a \otimes \nu_E \quad \text{on } E$$

for some polynomial a in three components on E. A direct calculation shows that, therefore, the jump of $(\operatorname{Curl} v_h) \cdot \nu_E$ along E vanishes. This proves our claim. We conclude continuity of $(\operatorname{Curl} v_h) \cdot \nu_E$ along $E \in \mathcal{E}$ and so

$$\mathcal{R}es_{NC}(v_h) = \sum_{E\in\mathcal{E}} \int_E [u_h] (\operatorname{Curl} v_h) \cdot \nu_E \, ds.$$

Since $\operatorname{Curl} v_h$ is constant along E and $\int_E [u_h]\, ds = 0$ we conclude

$$\mathcal{R}es_{NC}(v_h) = 0. \quad \square$$

Remark 2.6. For $n = 2$ dimensions, $|\operatorname{Curl} v| = |\nabla v|$ and $p_h \cdot \operatorname{Curl} v = p_h^\perp \cdot \nabla v$ for $p_h^\perp := (-p_{h2}, p_{h1})$. Hence

$$\begin{aligned}\mathcal{R}es_{NC}(v) &= \int_\Omega p_h^\perp \cdot \nabla v \, dx \quad \text{and} \\ \|\mathcal{R}es_{NC}\| &= \sup_{v\in H^1(\Omega)\setminus\mathbb{R}} \int_\Omega p_h^\perp \cdot \nabla v \, dx / \|\nabla v\|_L\end{aligned} \tag{2.23}$$

Element	$M_k(T)$	$D_k(T)$
RT	$\mathcal{P}_k^2 + x \cdot \mathcal{P}_k$	$\mathcal{P}_k$
BDM	$\mathcal{P}_{k+1}^2$	$\mathcal{P}_k$
BDFM	$\{q \in \mathcal{P}_{k+1}^2 : (q \cdot n)\|_{\partial T} \in \mathcal{R}_k(\partial T)\}$	$\mathcal{P}_k$

TABLE 1. Standard 2D Mixed FEMs allowed in Theorem 2.4. Here, $\mathcal{P}_k$ denotes polynomials of total degree at most $k = 0, 1, 2, ...$ and $\mathcal{R}_k(\partial T)$ denotes (not necessarily continuous) functions on ∂T which equal a polynomial of degree at most k on each edge.

is the usual operator norm as in (2.10). Notice the differences in the boundary conditions in (2.10) (where $v = 0$ on $\partial\Omega$) and (2.23) (where v has integral mean zero on Ω). Since $\operatorname{div}_{\mathcal{T}} p_h^{\perp} = 0$, we obtain

$$\eta_R^{(NC)} := \|h_{\mathcal{E}}^{1/2}[p_h \cdot \tau_{\mathcal{E}}]\|_{L^2(\cup\mathcal{E})} \tag{2.24}$$

(with the piecewise tangential unit vector $\tau_{\mathcal{E}}$) instead of (2.11)-(2.12) and, as in (2.13),

$$\eta_A := \|p_h - \sum_{z\in\mathcal{N}} (A_h p_h)\varphi_z\|_{L^2(\Omega)}. \tag{2.25}$$

Remark 2.7. We stress that all other estimators, for instance the localized or equilibrated implicit estimators of [AO, BS], are available for the assessment of

$$\|\mathcal{R}es_{NC}\| \approx \eta_R^{(NC)} \lesssim \eta_A \tag{2.26}$$

as well. The averaging estimator (2.25) concerns discontinuities in normal *and* tangential components and so

$$\|\mathcal{R}es_{NC}\| + \|\mathcal{R}es_H\|_H + \|h_{\mathcal{T}}^2 \nabla g\|_{L^2(\Omega)} \approx \eta_A + \|h_{\mathcal{T}}^2 \nabla_{\mathcal{T}} g\|_{L^2(\Omega)}. \tag{2.27}$$

Remark 2.8. The situation for $n = 3$ dimensions is more delicate and we refer to [CBJ] for reliable and efficient explicit error estimators.

2.4. **Mixed Finite Element Methods.** The Laplace equation is split into $\operatorname{div} p + g = 0$ and the weak form of $p = \nabla u$. The resulting mixed formulation involves a bilinear form as in (1.5). Its discrete version involves

$$p_h \in L_h \subseteq \mathcal{L}^{k+1}(\mathcal{T})^n \cap H(\operatorname{div};\Omega) \quad \text{and} \quad u_h \in H_h \subseteq \mathcal{L}^k(\mathcal{T})$$

for $k = 0, 1$. The couple (p_h, u_h) is supposed to satisfy

$$\begin{aligned} &\int_\Omega p_h \cdot q_h \, dx + \int_\Omega u_h \operatorname{div} q_h \, dx = 0 \quad \text{for all } q_h \in L_h, \\ &\int_\Omega v_h \operatorname{div} p_h \, dx = - \int_\Omega g \, v_h \, dx \quad \text{for all } v_h \in \mathcal{L}^k(\mathcal{T}). \end{aligned} \tag{2.28}$$

With the above sets $D_k(T)$ and $M_k(T)$ from Table 1 for $n = 2$ we define

$$\begin{aligned} L_h &:= M_k(\mathcal{T}) := \{q_h \in H(\operatorname{div}; \Omega) : \forall T \in \mathcal{T}, \, q_h|_T \in M_k(T)\}, \\ H_h &:= D_k(\mathcal{T}) := \{v_h \in L^\infty(\Omega) : \forall T \in \mathcal{T}, \, v_h|_T \in D_k(T)\}. \end{aligned}$$

Theorem 2.1 is applied to estimate $\|p - p_h\|_L + \|u - \tilde{u}_h\|_H$ for some $\tilde{u}_h \in H$ different from u_h. The evaluation of $\mathcal{R}es_H$ follows the arguments of the conforming finite element situation in Subsection 2.2.

Theorem 2.3. *Given $g \in H^{1+k}(\mathcal{T}) := \{g \in L^2(\Omega) : \forall T \in \mathcal{T}, \, g|_T \in H^{1+k}(T)\}$ and $k = 0, 1$,*

$$\|\mathcal{R}es_H\|_{H^*} \lesssim \|h_\mathcal{T}^{2+k} \nabla_\mathcal{T}^{1+k} g\|_{L^2(\Omega)} \quad \textit{is of higher order.}$$

Proof. An integration by parts shows, for all $v \in H$,

$$\mathcal{R}es_H(v) := - \int_\Omega g v \, dx + \int_\Omega p_h \cdot \nabla v \, dx = - \int_\Omega v (g + \operatorname{div} p_h) \, dx.$$

Note that there are no jump terms across interior element boundaries since $p_h \in H(\operatorname{div}; \Omega)$ (and, equivalently, $[p_h] \cdot \nu_\mathcal{E} = 0$). In the lowest-order cases, $H_h = \mathcal{L}^0(\mathcal{T})$ and $(2.28)_b$ lead to

$$- \operatorname{div} p_h|_T = \int_T g \, dx / |T| =: g_h|_T \quad \text{for all } T \in \mathcal{T}.$$

Consequently, if v_h and g_h denote the $\mathcal{T}$-piecewise constant averages of v and g, respectively, Poincaré inequalities show

$$\begin{aligned} \mathcal{R}es_H(v) &= - \int_\Omega (v - v_h)(g - g_h) \, dx \\ &\le \|h_\mathcal{T}(g - g_h)\|_{L^2(\Omega)} \|(v - v_h)/h_\mathcal{T}\|_{L^2(\Omega)} \\ &\lesssim \|h_\mathcal{T}^2 \nabla g\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)}. \end{aligned} \tag{2.29}$$

The proof is finished for $k = 0$. For $k = 1$, the second equation in (2.28) implies on each element domain $T \in \mathcal{T}$ that $- \operatorname{div} p_h = g_h := \Pi_1 g$ is the L^2 projection of g onto $P_1(T)$. Hence (2.29) can be moified to yield the upper bound $\|h_\mathcal{T}^3 \nabla_\mathcal{T}^2 g\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)}$. $\square$

For the evaluation of $\|\mathcal{R}es_L\|_{L^*}$ we have the following analogue of Theorem 2.2 which shows (1.8)-(1.9) in the notation from Definition 2.1;

recall that $\tilde{u}_h$ denotes an arbitrary (not necessarily discrete) element in H.

Theorem 2.4. *For each of the mixed FEM of Table 1 there holds*

$$\min_{\tilde{u}_h \in H} \|p_h - \nabla \tilde{u}_h\|_L = \|\mathcal{R}es_{NC}\| \quad \textit{and} \quad \mathcal{S}^1(\mathcal{T})^m \subset \ker \mathcal{R}es_{NC}.$$

Proof. The assertion follows as in Theorem 2.2: Given $v_h \in \mathcal{S}^1(\mathcal{T})^m$ the second part of its proof showed $\operatorname{Curl} v_h \in H(\operatorname{div};\Omega)$ by $[\operatorname{Curl} v_h]\cdot\nu_E = 0$. Hence,

$$q_h := \operatorname{Curl} v_h \in \mathcal{L}^1(\mathcal{T})^n \cap H(\operatorname{div},\Omega) \subset L_h.$$

The last inclusion holds for the finite element spaces of Table 1. Thus, $q_h := \operatorname{Curl} v_h$ may be considered in $(2.28)_a$ and shows

$$\mathcal{R}es_{NC}(v_h) = \int_\Omega p_h \cdot \operatorname{Curl} v_h \, dx = -\int_\Omega u_h \operatorname{div} \operatorname{Curl} v_h \, dx = 0. \quad \square$$

Remark 2.9. Based on Theorem 2.4, the evaluation of $\|\mathcal{R}es_{NC}\|$ follows the arguments of Subsections 2.2 and 2.3; e.g. for $n = 2$, $\eta_R^{(NC)}$, and η_A from (2.24)-(2.25) there holds

$$\|\mathcal{R}es_{NC}\| \approx \eta_R^{(NC)} \approx \eta_A.$$

3. Applications to the Stokes Problem

The stationary incompressible fluid flow can be modelled by the Stokes equations: Given $g \in L^2(\Omega)$ seek $(u,p) \in H \times L := H_0^1(\Omega)^n \times L_0^2(\Omega)$ with

$$\begin{aligned} &\int_\Omega 2\mu\, \varepsilon(u) : \varepsilon(v)\, dx - \int_\Omega p \operatorname{div} v \, dx = \int_\Omega g\, v\, dx, \\ &-\int_\Omega q \operatorname{div} u \, dx = 0 \quad \text{for all } (v,q) \in H_0^1(\Omega)^n \times L_0^2(\Omega). \end{aligned} \tag{3.1}$$

Here, $L_0^2(\Omega) := \{q \in L^2(\Omega) : \int_\Omega q\, dx = 0\} \equiv L^2(\Omega)/\mathbb{R}$ fixes a global additive constant in the pressure (because of lacking Neumann boundary conditions). In this case, (3.1) is equivalent to a formulation with the non-symmetric gradient ∇u instead of its symmetric part

$$\varepsilon(u) := (\nabla u + \nabla u^T)/2 \in L := L^2(\Omega; \mathbb{R}^{n\times n}_{\text{sym}}) \tag{3.2}$$

here; $\mathbb{R}^{n\times n}_{\text{sym}} = \{A \in \mathbb{R}^{n\times n} : A = A^T\}$. (Colon denotes the Euclidean scalar product, $A : B = \sum_{j,k=1}^n A_{jk}B_{jk}$ for $A, B \in \mathbb{R}^{n\times n}$.)

It is well-known that (3.1) has a unique solution (u,p). We discuss conforming and nonconforming finite element approximations of the

Stokes equations. Given a regular triangulation $\mathcal{T}$ and a $\mathcal{T}$-piecewise H^1 function (written $v \in H^1(\mathcal{T})$) set

$$\varepsilon_{\mathcal{T}}(v) := (\nabla_{\mathcal{T}} v + \nabla_{\mathcal{T}} v^T))/2 \in L := L^2(\Omega; \mathbb{R}^{n\times n}_{\mathrm{sym}})$$

for the $\mathcal{T}$-piecewise gradient $\nabla_{\mathcal{T}}$; $\nabla_{\mathcal{T}} v$ equals $\nabla(v|_T)$ on each $T \in \mathcal{T}$; let $\operatorname{div}_{\mathcal{T}}$ denote the $\mathcal{T}$-piecewise divergence operator. To describe conforming and nonconforming finite element methods simultaneously, suppose $u_h \in \mathcal{L}^k(\mathcal{T})^n$ and $p_h \in \mathcal{L}^k(\mathcal{T}) \cap L^2_0(\Omega)$ satisfy

$$\int_\Omega 2\mu\, \varepsilon_{\mathcal{T}}(u_h) : \varepsilon_{\mathcal{T}}(v_h)\, dx - \int_\Omega p_h \operatorname{div}_{\mathcal{T}} v_h\, dx = \int_\Omega g\, v_h\, dx \tag{3.3}$$

for all $v_h \in \mathcal{S}^1_0(\mathcal{T})^n$.

Remark 3.1. Even for nonconforming schemes we suppose that (3.3) holds for a continuous test function v_h. For the lowest-order finite element schemes, this implies the restriction to triangular finite elements.

Remark 3.2. Throughout the discussion of this paper, the discrete u_h and p_h are supposed to be piecewise polynomials of some degree $\leq k$. This does not mean that we propose some $P^n_k \times P_k$ finite element method — they may be instable. However, our a posteriori analysis partly includes error control even for unstable methods.

Remark 3.3. The condition $(3.1)_b$ has no discrete analog in (3.3) because that is not needed in our a posteriori error analysis. However, since $\| \operatorname{div}_{\mathcal{T}} u_h\|_{L^2(\Omega)}$ arises in estimates of $\mathcal{R}es_L$ below, it is understood below that $\| \operatorname{div}_{\mathcal{T}} u_h\|_{L^2(\Omega)}$ is small.

Remark 3.4. The list of examples for $n = 2$ includes conforming finite elements such as the MINI element, the P_2-P_0 finite element, and the Taylor-Hood element [BF] and the nonconforming finite element due to Kouhia and Stenberg [KS].

Remark 3.5. There are also finite element methods for the unsymmetric formulation [BF] such as the popular Crouzeix-Raviart finite element. Since the arguments of this section work verbatim (if not simpler), we omit details and refer to [DDP, V1].

Definition 3.1. Given $u_h \in H^1(\mathcal{T})^n$ and $p_h \in L^2_0(\Omega)$ set

$$\sigma_h := 2\mu\, \varepsilon_{\mathcal{T}}(u_h) - p_h \mathbf{1} \in L := L^2(\Omega; \mathbb{R}^{n\times n}_{\mathrm{sym}}) \tag{3.4}$$

and define the *linear functional* $\mathcal{R}es_H : H \to \mathbb{R}$ by

$$\mathcal{R}es_H(v) = \int_\Omega (g \cdot v - \sigma_h : \varepsilon(v))\, dx \quad \text{for } v \in H := H^1_0(\Omega)^n. \tag{3.5}$$

The linear function space H is endowed with the norm $\|v\|_H := |v|_{1,2} := \|\nabla v\|_2$ for $v \in H$ such that

$$\|\mathcal{R}es_H\|_{H^*} = \sup_{v \in H \setminus \{0\}} \mathcal{R}es_H(v)/|v|_{1,2}.$$

The residual $\|\mathcal{R}es_L\|_{L^*}$ involves the deviatoric-part operator

$$\operatorname{dev} A = A - (\operatorname{tr}(A)/n)\, \mathbf{1} \quad \text{for any } A \in \mathbb{R}^{n \times n} \tag{3.6}$$

(where $\operatorname{tr}(A) = A_{11} + \cdots + A_{nn}$ is the trace of A).

Theorem 3.1. *Let $(u, p) \in H \times L^2_0(\Omega)$ solve (3.1) and set $\sigma := 2\mu\, \varepsilon(u) - p\, \mathbf{1} \in L$. Let (u_h, p_h) satisfy (3.3) and define σ_h as in (3.4). Then, for any $\tilde{u}_h \in H$, there holds*

$$\|\sigma - \sigma_h\|_L + \|u - \tilde{u}_h\|_H \approx \|\varepsilon(\tilde{u}_h) - \operatorname{dev} \varepsilon_{\mathcal{T}}(u_h)\|_L + \|\mathcal{R}es_H(v)\|_{H^*}. \tag{3.7}$$

Before we focus on its proof, we briefly comment on applications of the theorem. The residual $\mathcal{R}es_H$ satisfies $\mathcal{R}es_H(v_h) = 0$ for all $v_h \in \mathcal{S}^1_0(\mathcal{T})^n$ and, since σ_h is symmetric, can be recast into

$$\mathcal{R}es_H(v) = \int_\Omega (g \cdot v - \sigma_h : \nabla v)\, dx. \tag{3.8}$$

That is, $\mathcal{R}es_H$ in (3.5) is the sum of $j = 1, 2, \ldots, n$ residuals $\mathcal{R}es(v\, e_j)$ of the form in $(2.6)_b$ where e_j is the j-th canonical unit vector in $\mathbb{R}^n$ and, here, v in $H^1_0(\Omega)$ is a scalar. As a consequence, the residual evaluation can follow the same lines as in Subsection 2.1.

The discussion of $\|\mathcal{R}es_L\|_{L^*} = \|\varepsilon(\tilde{u}_h) - \operatorname{dev} \varepsilon_{\mathcal{T}}(u_h)\|_L$ follows two cases. In case I, for any conforming approximation $u_h \in H$, the choice $\tilde{u}_h = u_h$ yields

$$\|\mathcal{R}es_L\|_{L^*} = \|\varepsilon(u_h) - \operatorname{dev} \varepsilon_{\mathcal{T}}(u_h)\|_L = n^{-1} \|\operatorname{div} u_h\|_{L^2(\Omega)}. \tag{3.9}$$

This is an appropriate error contribution and, at the same time, an error estimator.

In case II, $u_h \notin H$ and the estimation of the nonconformity terms is analogous to that of Subsection 2.3 but slightly more involved because of the interaction of the divergence residual and the $u_h - \tilde{u}_h$ approximation. If one accepts $\|\operatorname{div}_{\mathcal{T}} u_h\|_{L^2(\Omega)}$ as a proper error term (cf. Remark 3.3), the upper bound

$$\|\mathcal{R}es_L\|_{L^*} \le \|\varepsilon_{\mathcal{T}}(u_h - \tilde{u}_h)\|_L + n^{-1/2} \|\operatorname{div}_{\mathcal{T}} u_h\|_{L^2(\Omega)} \tag{3.10}$$

can be minimised according to a symmetric form of a Helmholtz decomposition. To quote results from the literature let $n = 2$ for a moment.

Lemma 3.2. ([CD, Lemma 3.2]). Given any $\tau \in L^2(\Omega;\mathbb{R}^{2\times 2}_{\rm sym})$ there exist $a \in H^1_0(\Omega)^2$ and $b \in H^2(\Omega)$ such that

$$\tau = \varepsilon(a) + \operatorname{Curl}\operatorname{Curl} b \tag{3.11}$$

(where $\operatorname{Curl}\operatorname{Curl} b$ has the four entries $b_{,22}, -b_{,12}, -b_{,12}, b_{,11}$ where $b_{,\alpha\beta} = \partial^2 b/\partial x_\alpha \partial x_\beta$). □

Suppose there holds $\tau = \varepsilon_{\mathcal{T}}(u_h) = \varepsilon(a) + \operatorname{Curl}\operatorname{Curl} b$ in (3.11). The, the lemma suggests the choice $\tilde{u}_h = a$ and, as for Theorem 2.2, one proves

$$\min_{\tilde{u}_h \in H} \|\varepsilon(u_h - \tilde{u}_h)\|_L = \|\operatorname{Curl}\operatorname{Curl} b\|_L = \|\mathcal{R}es_{NC}\|_{H^*} \tag{3.12}$$

for

$$\mathcal{R}es_{NC}(v) := \int_\Omega \varepsilon_{\mathcal{T}}(u_h) : \operatorname{Curl} v\, dx \quad \text{for all } v \in H.$$

[The proof follows closely (2.20)-(2.22) and is hence omitted.] More details may be found in [CF1] where it is in particular shown (for the Kouhia-Stenberg FEM and $n = 2$) that

$$\|\mathcal{R}es_{NC}\|_{H^*} \lesssim \|h_{\mathcal{E}}^{1/2}[\nabla u_h \cdot \tau_{\mathcal{E}}]\|_{L^2(\cup\mathcal{E})}. \tag{3.13}$$

The remaining part of this section is devoted to the proof of Theorem 3.1 for $n = 2$ or $n = 3$. To employ the mixed approach of Subsection 1.3 set

$$\begin{aligned} a(\sigma,\tau) &:= \int_\Omega 1/(2\mu)\, \operatorname{dev}\sigma : \operatorname{dev}\tau\, dx \quad \text{for } \sigma,\tau \in L, \\ b(\sigma; v) &:= -\int_\Omega \sigma : \varepsilon(v)\, dx \quad \text{for } (\sigma, v) \in L \times H. \end{aligned} \tag{3.14}$$

Let $L/\mathbb{R} := \{\sigma \in L : \int_\Omega \operatorname{tr}\sigma\, dx = 0\}$.

Lemma 3.3. The operator $A : X \to X^*$, defined for $(\sigma, u) \in X := (L/\mathbb{R}) \times H$ by

$$(A(\sigma,u))(\tau,v) := a(\sigma,\tau) + b(\sigma,v) + b(\tau,u) \tag{3.15}$$

is linear, bounded and bijective. [This result holds for $n = 2, 3$.]

Proof. The bijectivity of A is the only not so immediate part of the lemma. Proposition 3.1 in [BF, Chapter IV] states

$$\|\sigma\|_L \lesssim \|\operatorname{dev}\sigma\|_L + \|\operatorname{div}\sigma\|_{L^2(\Omega)} \quad \text{for all } \sigma \in L/\mathbb{R}.$$

Since any $\sigma \in L/\mathbb{R}$ with $b(\sigma;\cdot) = 0$ (written $\sigma \in \ker B$) satisfies $\operatorname{div}\sigma = 0$ this implies

$$\|\sigma\|_L^2 \lesssim a(\sigma,\sigma) \quad \text{for all } \sigma \in \ker B.$$

Hence a is elliptic on $\ker B$. This is one of the main ingredients of the general theory on mixed finite element mehods [BF]. The remaining details are omitted. □

Proof of Theorem 3.1. The inf-sup condition for A follows from Lemma 3.3. The resulting equivalence (1.7) reads

$$\|\sigma - \sigma_h\|_L + \|u - \tilde{u}_h\|_H \approx \|\mathcal{R}es_L\|_{L^*} + \|\mathcal{R}es_H\|_{H^*}.$$

The residuals on the right-hand side result from (1.6). In particular, if we employ $\operatorname{dev}\varepsilon(u) = \varepsilon(u)$ (from $\operatorname{div} u = 0$), there holds

$$\begin{aligned}\mathcal{R}es_L(\tau) &= \int_\Omega \Big((1/(2\mu)\ \operatorname{dev}(\sigma - \sigma_h) - \varepsilon(u - \tilde{u}_h) \Big) : \tau\, dx \\ &= \int_\Omega \Big(\varepsilon(\tilde{u}_h) - \operatorname{dev}_{\mathcal{T}} \varepsilon(u_h) \Big) : \tau\, dx.\end{aligned}$$

The remaining details are omitted. □

4. Applications to linear elasticity

This section is devoted to the Navier-Lamè equations and its conforming, nonconforming, and mixed finite element approximation. An analysis of enhanced finite elements in the same framework is given in [BCR]. It is an important feature of the presented unifying theory that the resulting a posteriori error estimates are robust with respect to the Lamè parameter $\lambda \to \infty$.

4.1. **Model Problem.** We adopt the notation of the previous two sections and continue with a linear stress-strain relation of the form

$$\mathbb{C}A := \lambda\ \operatorname{tr}(A)\, \mathbf{1} + 2\mu\, A \quad \text{for } A \in \mathbb{R}^{n\times n}$$

with inverse relation

$$\mathbb{C}^{-1}A = 1/(2\mu)\, A - \lambda/(2\mu(n\lambda - 2\mu))\ \operatorname{tr}(A)\, \mathbf{1} \quad \text{for } A \in \mathbb{R}^{n\times n}.$$

In the continuous model, $\sigma = \mathbb{C}\varepsilon(u) \in L := L^2(\Omega; \mathbb{R}^{n\times n}_{\mathrm{sym}})$ and the resulting model problem reads: Given $g \in L^2(\Omega)^n$ find $u \in H := H^1_0(\Omega)^n$ with

$$g + \operatorname{div} \mathbb{C}\, \varepsilon(u) = 0 \quad \text{in } \Omega. \tag{4.1}$$

The material parameters λ and μ are positive and hence (4.1) is an elliptic PDE with a unique solution u.

To employ the unified theory of Subsection 1.3 let

$$a(\sigma, \tau) := \int_\Omega (\mathbb{C}^{-1}\sigma) : \tau\, dx \quad \text{for } \sigma, \tau \in L := L^2(\Omega; \mathbb{R}^{n\times n}_{\mathrm{sym}}) \tag{4.2}$$

replace $(3.14)_a$ and adopt $b(\sigma; v)$ from $(3.14)_b$.

Lemma 4.1 ([BCR])**.** The operator $A : X \to X^*$ from (3.15) is linear, bounded, and bijective and the operator norms of A and A^{-1} are λ-independent. □

Remark 4.1. The operator A represents the weak form of the Hellinger-Reissner principle in mechanics. Lemma 4.1 is the analogue of Lemma 3.3.

4.2. **Conforming Finite Element Methods.** Although the (lower order) conforming finite element methods are not robust in $\lambda \to \infty$ we introduce a robust error estimation. Given a finite element approximation $u_h \in H = H_0^1(\Omega)^n$ with

$$\int_\Omega \varepsilon(u_h) : \mathbb{C}\,\varepsilon(v_h)\,dx = \int_\Omega g \cdot v_h\,dx \quad \text{for all } v_h \in \mathcal{S}_0^1(\mathcal{T})^n, \tag{4.3}$$

let $e := u - u_h \in H_0^1(\Omega)$ (where u solves (4.1)) denote the error.

Theorem 4.2. *With λ-independent constants in $\approx$, there holds*

$$\|\mathbb{C}\,\varepsilon(e)\|_{L^2(\Omega)} + \|e\|_{H^1(\Omega)} \approx \sup_{v \in H\setminus\{0\}} \int_\Omega (g\cdot v - \varepsilon(u_h) : \mathbb{C}\,\varepsilon(v))\,dx / \|v\|_{H^1(\Omega)}.$$

Proof. With $\tilde{u}_h = u_h \in H$, $\sigma := \mathbb{C}\,\varepsilon(u)$, and $\sigma_h := \mathbb{C}\,\varepsilon(u_h)$, Equivalence (1.7) results in

$$\|\sigma - \sigma_h\|_L + \|e\|_{H^1(\Omega)} \approx \|\mathcal{R}es_H\|_{H^*}$$

where $\mathcal{R}es_H(v)$ is defined in (3.8). This implies the assertion. □

As in the previous section (cf. the discussion about (3.8)), the estimation of $\|\mathcal{R}es_H\|_{H^*}$ follows the lines of Subsection 2.1. Given any estimator η (of the various choices (2.11)-(2.14)) with

$$\|\mathcal{R}es_H\|_{H^*} \approx \eta \quad \text{(up to h.o.t.)} \tag{4.4}$$

the estimate [the proof follows from (4.4) and $\|e\|_{H^1(\Omega)} \lesssim \|\sigma - \sigma_h\|_{L^2(\Omega)}$]

$$\|\sigma - \sigma_h\|_{L^2(\Omega)} \approx \eta \quad \text{(up to h.o.t.)} \tag{4.5}$$

appears to be new (where h.o.t. refers to $\|h_\mathcal{T}^2 \nabla g\|_{L^2(\Omega)}$ for the first-order finite element schemes). The point is that the constants behind $\approx$ in (4.5) are λ-independent. This is different for the standard estimate

$$\|\mathbb{C}^{1/2}\varepsilon(e)\|_{L^2(\Omega)} \lesssim \eta \lesssim \|\sigma - \sigma_h\|_{L^2(\Omega)} \quad \text{(up to h.o.t.).} \tag{4.6}$$

In fact, for the proof of (4.6) one argues

$$\int_\Omega \varepsilon(e) : \mathbb{C}\,\varepsilon(e)\,dx = \mathcal{R}es_H(e) \le \|\mathcal{R}es_H\|_{H^*}\|e\|_H$$
$$\lesssim \|\mathbb{C}^{-1/2}\|_\infty \|\mathbb{C}^{1/2}\varepsilon(e)\|_{L^2(\Omega)}\,\eta \lesssim \|\mathbb{C}^{1/2}\varepsilon(e)\|_{L^2(\Omega)}\eta.$$

Finally, inverse estimates verify $(4.6)_b$ up to higher order terms (h.o.t.). Notice that (4.5) is balanced in λ while (4.6) is not. Thus (4.4) establishes λ-robust a posteriori error control of the L^2-stress error $\|\sigma - \sigma_h\|_{L^2(\Omega)}$. A corresponding result for the energy norm $\|\mathbb{C}^{1/2}\varepsilon(e)\|_{L^2(\Omega)}$ remains as an open problem.

4.3. **Nonconforming Finite Element Methods.** Robust a priori convergence estimates are known for nonconforming finite element methods such as

$$\mathcal{S}_0^1(\mathcal{T})^2 \subset \mathcal{S}_0^1(\mathcal{T}) \times \mathcal{S}_0^{1,NC}(\mathcal{T}) \subset H_0^1(\Omega) \times H^1(\mathcal{T})$$

due to Kouhia-Stenberg [KS]. Here and throughout this subsection, let $n = 2$ and consider merely triangles. Suppose that the discrete solution $u_h \in H^1(\mathcal{T})^2$ satisfies

$$\int_\Omega \varepsilon_\mathcal{T}(u_h) : \mathbb{C}\,\varepsilon(v_h)\,dx = \int_\Omega g \cdot v_h\,dx \quad \text{for all } v_h \in \mathcal{S}_0^1(\mathcal{T})^2 \tag{4.7}$$

and set $\sigma_h := \mathbb{C}\,\varepsilon_\mathcal{T}(u_h)$, $\sigma := \mathbb{C}\,\varepsilon(u) \in L := L^2(\Omega; \mathbb{R}^{2\times 2}_{\text{sym}})$.

Theorem 4.3. *For any $\tilde{u}_h \in H$ there holds*

$$\|\sigma - \sigma_h\|_L + \|u - \tilde{u}_h\|_H \approx \|\varepsilon_\mathcal{T}(u_h - \tilde{u}_h)\|_L + \|\mathcal{R}es_H\|_{H^*}. \tag{4.8}$$

Proof. This is a result of Equivalence (1.7) and Lemma 4.1. The details are similar to those of the previous sections and hence omitted. □

The discussion of $\|\varepsilon_\mathcal{T}(u_h - \tilde{u}_h)\|_L$ follows the lines of (3.10)-(3.13) in the previous section and hence are omitted; cf. [CF2].

4.4. **Mixed Finite Element Methods.** Another λ-robust approximation is feasible with a mixed finite element method with reduced symmetry [ABD, S] for

$$\sigma_h \in H(\mathrm{div}; \Omega),$$

i.e. $\mathrm{div}\,\sigma_h \in L^2(\Omega; \mathbb{R}^n)$ and $\sigma_h \in \mathcal{L}^k(\mathcal{T}; \mathbb{R}^{n\times n})$ with, in general, $As(\sigma_h) := \sigma_h - \mathrm{sym}\,\sigma_h \not\equiv 0$, $\mathrm{sym}\,\sigma_h := (\sigma_h^T + \sigma_h)/2$. Suppose that σ_h satisfies

$$-\int_\Omega v_h \cdot \mathrm{div}\,\sigma_h\,dx = \int_\Omega v_h \cdot g\,dx \quad \text{for all } v_h \in \mathcal{L}^0(\mathcal{T})^n. \tag{4.9}$$

Theorem 4.4. *For any $\tilde{u}_h \in H$ there holds*

$$\begin{aligned} &\|\sigma - \operatorname{sym}\sigma_h\|_L + \|u - \tilde{u}_h\|_H \\ &\quad \approx \|\varepsilon(\tilde{u}_h) - \mathbb{C}^{-1}\operatorname{sym}\sigma_h\|_L + \|\mathcal{R}es_H\|_{H^*}. \end{aligned} \tag{4.10}$$

Proof. Lemma 4.1 and Equivalence (1.7) apply to $\sigma - \operatorname{sym}\sigma_h$ and $u - \tilde{u}_h$. We omit the details. □

Notice that the residual satisfies

$$\begin{aligned} \mathcal{R}es_H(v) &= \int_\Omega (v \cdot g - \varepsilon(v) : \sigma_h)\, dx \\ &= \int_\Omega v \cdot (g + \operatorname{div}\sigma_h)\, dx - \int_\Omega As(\sigma_h) : \nabla(v)\, dx \quad \text{for } v \in H. \end{aligned}$$

(For a proof observe that $\varepsilon(v) : \operatorname{sym}\sigma_h = \nabla(v) : (\sigma_h - As(\sigma_h))$ and employ an integration by parts.) The estimation of $\|\mathcal{R}es_H\|_{H^*}$ may hence follow as in (2.29) and yields

$$\|\mathcal{R}es_H\|_{H^*} \lesssim \|h_\mathcal{T}^2 \nabla_\mathcal{T} g\|_{L^2(\Omega)} + \|As(\sigma_h)\|_{L^2(\Omega)}. \tag{4.11}$$

The first term on the right-hand side of (4.11) is of higher order for $g \in H^1(\mathcal{T})^n$ and first order schemes such as PEERS. The estimation of

$$\min_{\tilde{u}_h \in H} \|\varepsilon(\tilde{u}_h) - \mathbb{C}^{-1}\operatorname{sym}\sigma_h\|_L = \|\mathcal{R}es_{NC}\| \tag{4.12}$$

follows closely the discussion of (3.10)-(3.13). We refer to [CD, CDFH] for the remaining details on the approximation of

$$\mathcal{R}es_{NC}(v) := \int_\Omega \operatorname{Curl} v : \operatorname{sym}\mathbb{C}^{-1}\sigma_h\, dx.$$

Notice that (4.10) concerns the symmetric part of the error $\|\sigma - \sigma_h\|_L$. Furthermore,

$$\|\sigma - \sigma_h\|_L^2 = \|\sigma - \operatorname{sym}\sigma_h\|_L^2 + \|As(\sigma_h)\|_L^2$$

and $\|As(\sigma_h)\|_L$ may simultaneously be regarded as an error contribution and as a (computable) contribution to an a posteriori error estimate. Hence the estimate (4.10) results in

$$\|\sigma - \sigma_h\|_L \approx \|\mathcal{R}es_{NC}\|_{H^*} + \|\mathcal{R}es_H\|_{H^*} + \|As(\sigma_h)\|_L. \tag{4.13}$$

Acknowledgments

The results of this paper were presented entitled *A Posteriori Mixed Finite Element Error Control* during the ENUMATH 2001 conference in July within the workshop on mixed finite element methods. The invitation to this is greatfully acknowledged.

References

[AO] Ainsworth, M .and Oden, J.T. (2000). *A posteriori error estimation in finite element analysis.* Wiley-Interscience [John Wiley & Sons], New York.

[A] Alonso, A. (1996). *Error estimators for a mixed method.* Numer. Math., **74**, 4, 385–395.

[AF] Arnold, D.N. and Falk, R.S. (1989). *A uniformly accurate finite element method for the Reissner-Mindlin plate model*, SIAM J. Numer. Anal., **26**, 1276–1290.

[ABD] Arnold, D.N., Brezzi, F. and Douglas, J. (1984). *PEERS: A new finite element for plane elasticity.* Japan J. Appl. Math, **1**, 2, 347–367.

[BaM] Babuška, I. and Miller, A. (1987). *A feedback finite element method with a posteriori error estimation: Part I, The finite element method and some properties of the a posteriori estimator.* Comp. Methods Appl. Mech. Engrg., **61**, 1, 1–40.

[BaR] Babuška, I. and Rheinboldt, W.C. (1981). *A posteriori error analysis of finite element solutions for one–dimensional problems.* SIAM Journal on Numerical Analysis, **18**, 565–589.

[BaS] Babuška, I. and Strouboulis, T. (2001). *The Finite Element Method and its Reliability.* The Clarendon Press Oxford University Press, xii+802.

[BC] Bartels, S. and Carstensen, C. (2002). *Each averaging technique yields reliable a posteriori error control in FEM on unstructured grids. II. Higher order FEM.* Math. Comp., **71**, 239, 971–994.

[BR1] Becker, R. and Rannacher, R. (1996). *A feed-back approach to error control in finite element methods: basic analysis and examples*, East-West J. Numer. Math, **4**, 237–264.

[BR2] Becker, R. and Rannacher, R. (2001). *An optimal control approach to a posteriori error estimation in finite element methods.* Acta Numerica, Cambridge University Press, 2001, 1–102.

[B] Braess, D. (1997). *Finite Elements*, Cambridge University Press.

[BCR] Braess, D. and Carstensen, C. and Reddy, B.D. (2004). *Uniform Convergence and a posteriori error estimators for the enhanced strain finite element method.* Numer. Math, **96**, 461–479.

[BS] Brenner, S.C. and Scott, L.R. (1994). *The mathematical theory of finite element methods*, Springer Verlag, New York, Texts in Applied Mathematics, **15**.

[BF] Brezzi, F. and Fortin, M. (1991). *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, New York.

[C] Carstensen, C. (1997). *A posteriori error estimate for the mixed finite element method.* Math. Comp., **66**, 465–476.

[CBJ] Carstensen, C., Bartels, S,. Dolzmann, G. (2002). *A posteriori error estimates for nonconforming finite element methods*, Numer. Math, **92**, Issue 2, 233-256

[CBK] Carstensen, C., Bartels, S. and Klose, R. (2001). *An experimental survey of a posteriori Courant finite element error control for the Poisson equation.* Adv. Comput. Math., **15**, 1-4, 79–106 (2002).

[CD] Carstensen, C. and Dolzmann, G. (1998). *A posteriori error estimates for mixed FEM in elasticity.* Numer. Math., **81**, 2, 187–209.

[CDFH] Carstensen, C., Dolzmann, G., Funken, S.A. and Helm, D.S. (2000). *Locking-free adaptive mixed finite element methods in linear elasticity.* Comput. Methods Appl. Mech. Engrg., **190**, 13–14, 1701–1718.

[CF1] Carstensen, C. and Funken, S. A. (2001). *A posteriori error control in low-order finite element discretisations of incompressible stationary flow problems.* Math. Comp., **70**, 236, 1353–1381.

[CF2] C. Carstensen and S. A. Funken. *Averaging technique for FE-a posteriori error control in elasticity. Part I: Conforming FEM.* Comput. Methods Appl. Mech. Engrg., **190** (2001) 2483-2498. *Part II: λ-independent estimates.* Comput. Methods Appl. Mech. Engrg., **190** (2001) 4663—4675. *Part III: Locking-free conforming FEM.* Comput. Methods Appl. Mech. Engrg., **191** (2001), no. 8-10, 861–877.

[CV] Carstensen, C. and Verfürth, R. (1999). *Edge residuals dominate a posteriori error estimates for low order finite element methods*, SIAM J. Numer. Anal., **36**, 5, 1571–1587.

[DDVP] Dari, E., Duran, R., Padra, C., Vampa, V. (1996). *A posteriori error estimators for nonconforming finite element methods*, RAIRO Modl. Math. Anal. Numr., **30**, no. 4, 385–400.

[DDP] Dari, E., Durn, R., Padra, C. (1995). *Error estimators for nonconforming finite element approximations of the Stokes problem*, Math. Comp. **64**, no. 211, 1017–1033.

[EJ] Eriksson, K. and Johnson, C. (1991). *Adaptive finite element methods for parabolic problems I. A linear model problem*, SIAM J. Numer. Anal., **28**, 43–77.

[HW] Hoppe, R.H.W. and Wohlmuth, B. (1996). *Element-orientated and edge-orientated local error estimates for nonconforming finite element methods*, Math. Modeling Numer. Anal., **30**, 237–263.

[KS] Kouhia, R. and Stenberg, R. (1995). *A linear nonconforming finite element method for nearly incompressible elasticity and Stokes flow.* Comput. Methods Appl. Mech. Engrg., **124**, 3, 195–212.

[S] Stenberg, R. (1988). *A family of mixed finite elements for the elasticity problem.* Numer. Math., **53**, 5, 513–538.

[V1] Verfürth, R. (1989). *A posteriori error estimators for the Stokes equations*, Numer. Math., **55**, no. 3, 309–325.

[V2] Verfürth, R. (1996). *A review of a posteriori error estimation and adaptive mesh-refinement techniques*, Wiley-Teubner.