MATHEMATICS OF COMPUTATION
Volume 73, Number 247, Pages 1153–1165
S 0025-5718(03)01580-1
Article electronically published on August 12, 2003

ALL FIRST-ORDER AVERAGING TECHNIQUES FOR A POSTERIORI FINITE ELEMENT ERROR CONTROL ON UNSTRUCTURED GRIDS ARE EFFICIENT AND RELIABLE

C. CARSTENSEN

Abstract. All first-order averaging or gradient-recovery operators for lowest-order finite element methods are shown to allow for an efficient a posteriori error estimation in an isotropic, elliptic model problem in a bounded Lipschitz domain Ω in $\mathbb{R}^d$. Given a piecewise constant discrete flux $p_h \in P_h$ (that is the gradient of a discrete displacement) as an approximation to the unknown exact flux p (that is the gradient of the exact displacement), recent results verify efficiency and reliability of

$$\eta_M := \min\{\|p_h - q_h\|_{L^2(\Omega)} : q_h \in \mathcal{Q}_h\}$$

in the sense that η_M is a lower and upper bound of the flux error $\|p-p_h\|_{L^2(\Omega)}$ up to multiplicative constants and higher-order terms. The averaging space $\mathcal{Q}_h$ consists of piecewise polynomial and globally continuous finite element functions in d components with carefully designed boundary conditions. The minimal value η_M is frequently replaced by some averaging operator $A : P_h \to \mathcal{Q}_h$ applied within a simple post-processing to p_h. The result $q_h := Ap_h \in \mathcal{Q}_h$ provides a reliable error bound with $\eta_M \leq \eta_A := \|p_h - Ap_h\|_{L^2(\Omega)}$.

This paper establishes $\eta_A \leq C_{\text{eff}}\, \eta_M$ and so equivalence of η_M and η_A. This implies efficiency of η_A for a large class of patchwise averaging techniques which includes the ZZ-gradient-recovery technique. The bound $C_{\text{eff}} \leq 3.88$ established for tetrahedral P_1 finite elements appears striking in that the shape of the elements does *not* enter: The equivalence $\eta_A \approx \eta_M$ is robust with respect to anisotropic meshes. The main arguments in the proof are Ascoli's lemma, a strengthened Cauchy inequality, and elementary calculations with mass matrices.

1. Introduction

Suppose p_h is the discrete flux obtained from a conforming, nonconforming, or mixed low-order finite element method (FEM) based on a regular triangulation $\mathcal{T}$ of the domain Ω. That is, p_h is the piecewise polynomial but globally discontinuous elementwise gradient of the finite element displacement approximations u_h or a discrete flux variable (for a mixed FEM) that approximates the unknown exact flux p. It is the aim of a posteriori error control to bound the error $\|p - p_h\|_{L^2(\Omega)}$ from above and below by computable estimators [AO, BS, V]. It has recently been proven for several examples [CB, BC1, CF3, CF4] that the error $\|p - p_h\|_{L^2(\Omega)}$ in

Received by the editor July 26, 2002 and, in revised form, January 1, 2003.
2000 *Mathematics Subject Classification.* Primary 65N30; Secondary 65N15.
Key words and phrases. A posteriori error estimate, efficiency, finite element method, gradient recovery, averaging operator, mixed finite element method, nonconforming finite element method.

second-order elliptic boundary value problems is bounded by $\|p_h - q_h\|_{L^2(\Omega)}$ for any continuous and piecewise polynomial q_h in the sense that

$$\|p - p_h\|_{L^2(\Omega)} \leq C_{\text{rel}} \|p_h - q_h\|_{L^2(\Omega)} + \text{h.o.t.}$$

The boundary values are included in the set $\mathcal{Q}_h$ of possible averages q_h. The surprising aspect is that *all* averaging techniques which, given p_h, compute $q_h \in \mathcal{Q}_h$ are *reliable* in the sense that

$$\|p - p_h\|_{L^2(\Omega)} \leq C_{\text{rel}} \eta_M + \text{h.o.t.} \quad \text{for } \eta_M := \min_{q_h \in \mathcal{Q}_h} \|p_h - q_h\|_{L^2(\Omega)}.$$

The minimum η_M is frequently replaced by an upper bound η_A,

$$\eta_M \leq \eta_A := \|p_h - Ap_h\|_{L^2(\Omega)},$$

where $Ap_h \in \mathcal{Q}_h$ is computed with some local averaging operator A. One striking feature of η_M is its immediate efficiency,

$$\begin{aligned} \eta_M &= \min_{q \in \mathcal{Q}_h} \|p_h - p + p - q_h\|_{L^2(\Omega)} \\ &\leq \|p - p_h\|_{L^2(\Omega)} + \min_{q_h \in \mathcal{Q}_h} \|p - q_h\|_{L^2(\Omega)} \\ &= \|p - p_h\|_{L^2(\Omega)} + \text{h.o.t.} \end{aligned}$$

This follows from a simple triangle inequality plus some considerations of the minimal $\|p - q_h\|_{L^2(\Omega)}$. The latter argument requires smoothness of p and the correct treatment of boundary conditions that restrict the set $\mathcal{Q}_h$. Note that the multiplicative constant in the efficiency estimate

$$\eta_M \leq \|p - p_h\| + \text{h.o.t.} \tag{1.1}$$

is one; i.e. η_M is a lower bound up to higher-order terms. This is, in general, untrue for its upper bound η_A. The possible overestimation of the error $\|p - p_h\|_{L^2(\Omega)}$ by $C_{\text{rel}}\eta_A$ might be very large. In [CB, BC1] a local (edge-oriented) averaging is suggested and shown to be equivalent to η_M (cf. Theorem 3.2 in [CB]). In this paper we analyse a different and more popular averaging operator defined by

$$(Ap_h)(z) = A_z(p_h|_{\omega_z}) \in \mathbb{R}^d \quad \text{for each node } z$$

and its patch ω_z (cf. Section 2 for notation). Here, $A_z := \pi_z \circ M_z$ for some continuous averaging M_z that is exact for constants and the orthogonal projection π_z onto an affine subspace $\mathcal{A}_z \subset \mathbb{R}^d$ that carries proper boundary conditions. The main result, Theorem 4.1, reads

$$\eta_M \leq \eta_A \leq C_{\text{eff}}\, \eta_M. \tag{1.2}$$

It is remarkable that the constant C_{eff} depends only on the norm of A_z and so it holds for any unstructured grid as well as for a quite large class of averaging and finite element schemes. For the popular choice of integral means

$$M_z(p_h) := \int_{\omega_z} p_h \, dx / |\omega_z| \tag{1.3}$$

for any node z with patch ω_z of area or volume $|\omega_z|$ we establish in Corollary 5.3 for P_1 finite elements the estimates

$$1 \leq C_{\text{eff}} \leq \sqrt{10} \text{ for } d = 2 \quad \text{and} \quad 1 \leq C_{\text{eff}} \leq \sqrt{15} \text{ for } d = 3. \tag{1.4}$$

This is surprisingly sharp and does not depend on any detail of the regular triangulation with (possibly) degenerating triangles or tetrahedra.

Remark 1.1. The averaging technique (1.3) is our interpretation of the ZZ-estimator [ZZ, V] for which reliability and efficiency have been observed before [R1, R2, N, BR] (without treatment of mixed boundary conditions).

Remark 1.2. The averaging estimator η_A can be shown to be equivalent to the edge contributions

$$\eta_{\mathcal{E}} := (\sum_{E\in\mathcal{E}} h_E \|[p_h]|_E\|^2_{L^2(E)})^{1/2},$$

where $[p_h]|_E$ denotes the jump of p_h across the edge $E \in \mathcal{E}$ (with proper modifications on the boundary). Thus our qualitative results (partly) follow from reliability and efficiency of $\eta_{\mathcal{E}}$ as well [C, CV, R1, V].

Remark 1.3. The above estimates on $C_{\rm eff}$ yield lower bounds $C_{\rm eff}^{-1} \le C_{\rm rel}$ on the reliability constant (up to higher-order terms). Upper bounds on $C_{\rm rel}$ for related estimators with a best value around 1 can be found in [CF1, CF2].

Remark 1.4. As important corollaries of $\eta_M \approx \eta_A$ and (1.1) we obtain efficiency

$$\eta_A \le C_{\rm eff}\,\|p - p_h\| + \text{h.o.t.} \tag{1.5}$$

of the reliable error estimation by η_A in [CA, CB, BC1, CF3, CF4].

The remaining part of the paper is organised as follows. Section 2 presents the necessary technical notation. The preliminaries of Section 3 include Ascoli's lemma, the strengthened Cauchy inequality, and eigenvalues of mass matrices. The main result (1.2) is stated as Theorem 4.1 in Section 4 with a proof. An analysis of $C_{\rm eff}$ in a model situation of Section 5 leads to (1.4) shown in Corollary 5.3.

2. Assumptions

2.1. **Regular triangulation.** The bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$, with piecewise affine boundary Γ is exactly covered by a triangulation $\mathcal{T}$, $\bigcup \mathcal{T} = \overline{\Omega}$. Each element $T \in \mathcal{T}$ is a compact interval $T = \operatorname{conv}\{a, b\}$ if $d = 1$, a triangle $T = \operatorname{conv}\{a, b, c\}$ if $d = 2$, or a tetrahedron $T = \operatorname{conv}\{a, b, c, d\}$ if $d = 3$. The element's vertices $a, \dots, d$ are called nodes; $\mathcal{N}$ denotes the set of all nodes. Each flat boundary E of an element $T \in \mathcal{T}$ is either a point $E = \{a\}$, an edge $E = \operatorname{conv}\{a, b\}$, or a face $E = \operatorname{conv}\{a, b, c\}$; $\mathcal{E}$ denotes the set of all such E; $\mathcal{E}_\Omega$ denotes the interior edges or faces and $\mathcal{E}_\Gamma := \{E \in \mathcal{E} : E \subset \Gamma\} = \mathcal{E}\backslash\mathcal{E}_\Omega$ denotes the boundary edges. Analogous notation apply to parallelograms ($d = 2$) or parallelepipeds ($d = 3$) which are possible elements in $\mathcal{T}$ as well. Intersecting distinct elements share either one vertex, an edge, or a common face. Hanging nodes are excluded for simplicity. For each node $z \in \mathcal{N}$ let $\mathcal{E}_z := \{E \in \mathcal{E} : z \in E \cap \mathcal{N}\}$ and the patch $\omega_z := \operatorname{int}(\bigcup \mathcal{T}_z)$, $\mathcal{T}_z := \{T \in \mathcal{T} : z \in T \cap \mathcal{N}\}$. Each edge or face E is associated to a unit normal vector ν_E with fixed orientation; if $E \subseteq \partial\Omega$, set $\nu_E = \nu$, the outer unit normal along $\partial\Omega$. The length and area of $E \in \mathcal{E}$ is denoted by $h_E = \operatorname{diam}(E)$ and $|E| = \mathcal{L}^{d-1}(E)$, respectively; $\mathcal{L}^n$ denotes the n-dimensional Lebesgue measure along any affine subspace of $\mathbb{R}^d$. Similarly the length and volume of $T \in \mathcal{T}$ is denoted by $h_T = \operatorname{diam}(T)$ and $|T| = \mathcal{L}^d(T)$, respectively.

2.2. **Boundary data.** The boundary $\Gamma = \bigcup \mathcal{E}_\Gamma$ is split into a relatively closed part Γ_D and a remaining part $\Gamma_N := \Gamma \backslash \Gamma_D$ such that any edge $E \in \mathcal{E}_\Gamma$ belongs either to Γ_D or to $\overline{\Gamma}_N$. Two disjoint subsets $\mathcal{E}_D$ and $\mathcal{E}_N$ of $\mathcal{E}_\Gamma$ are supposed to satisfy

$$\mathcal{E}_D = \emptyset \quad \text{or} \quad \mathcal{E}_D = \{E \in \mathcal{E}_\Gamma : E \subset \Gamma_D\},$$
$$\mathcal{E}_N = \emptyset \quad \text{or} \quad \mathcal{E}_N = \{E \in \mathcal{E}_\Gamma : E \subset \overline{\Gamma_N}\}.$$

Given $\mathcal{E}_D$ and $\mathcal{E}_N$, the boundary data $g \in L^2(\Gamma_N)$ and $u_D \in H^{1/2}(\Gamma_D) \cap C(\Gamma_D)$ (i.e. u_D is continuous on Γ_D and can be extended to a function in $H^1(\Omega)$) satisfy $g \in C(\mathcal{E}_D)$ and $u_D \in C^1(\mathcal{E}_N)$; i.e.

$$g|_E \in C(E) \text{ for all } E \in \mathcal{E}_N \quad \text{and} \quad u_D|_E \in C^1(E) \text{ for all } E \in \mathcal{E}_D.$$

On each $E \in \mathcal{E}_D$, let $\tau_E^{(j)}$ denote a tangential unit vector for $j = 1, \ldots, d-1$ such that $(\nu_E, \tau_E^{(1)}, \ldots, \tau_E^{(d-1)})$ is a Cartesian basis of $\mathbb{R}^d$. Then, $\nabla_E u_D$ denotes the tangential derivative and, given $a \in \mathbb{R}^d$, $(a)_E$ denotes the vector of all components of a in $(\tau_E^{(j)})_{j=1}^{d-1}$, e.g. $(a)_E = (\tau_E^{(1)} \cdot a, \tau_E^{(2)} \cdot a)$ for $d = 3$; $\nabla_E u_D = (\nabla u_D)_E = \partial u_D / \partial s$ for $d = 2$.

The Dirichlet and Neumann boundary conditions on the gradient $p = \nabla u$ are asserted at each boundary node $z \in \mathcal{N}$ by $p(z) \in \mathcal{A}_z$ for the affine subspace

$$\mathcal{A}_z := \{a \in \mathbb{R}^d : \forall E \in \mathcal{E}_z \cap \mathcal{E}_N,\ g(z) = a \cdot \nu_E \text{ and } \forall E \in \mathcal{E}_z \cap \mathcal{E}_D, \nabla_E u_D(z) = (a)_E\} \tag{2.1}$$

of $\mathbb{R}^d$. Set $\mathcal{A}_z = \mathbb{R}^d$ for $z \in \mathcal{N} \cap \Omega$ and suppose $\mathcal{A}_z \neq \emptyset$ for all $z \in \mathcal{N}$. Finally, let $\pi_z : \mathbb{R}^d \to \mathbb{R}^d$ denote the orthogonal projection onto $\mathcal{A}_z$,

$$\mathcal{A}_z = \pi_z(0) + \mathcal{V}_z,$$

where $\mathcal{V}_z$ is a linear subspace of $\mathbb{R}^d$. The (nonlinear) orthogonal projection π_z is Lipschitz continuous with $\operatorname{Lip}(\pi_z) \leq 1$ and, for each $a \in \mathbb{R}^d$, $a - \pi_z(a) \perp \mathcal{V}_z$.

Remark 2.1. As an intersection of hyperplanes, $\mathcal{A}_z$ is an affine subspace of $\mathbb{R}^d$. The condition $\mathcal{A}_z \neq \emptyset$ is essentially a consistency condition on the boundary data: If $u \in C^1(\overline{\omega}_z)$ satisfies $u = u_D$ on $\Gamma_D = \bigcup \mathcal{E}_D$ and $\partial u / \partial \nu = g$ on $\overline{\Gamma}_N = \bigcup \mathcal{E}_N$, then $\nabla u(z) \in \mathcal{A}_z$.

Remark 2.2. The condition $(a)_E = \nabla_E u_D(z)$ in (2.1) is equivalent to

$$a \cdot \tau_E = \partial u_D(z) / \partial \tau_E \quad \text{for all vectors } \tau_E \in \mathbb{R}^d \text{ with } \tau_E \perp \nu_E.$$

This is a Dirichlet boundary condition $u = u_D$ on E in terms of $a = p(z) = \nabla u(z)$ at z.

Remark 2.3. In case $\mathcal{E}_D \cap \mathcal{E}_z = \emptyset$, the condition $p \in \mathcal{A}_z$ asserts Neumann boundary conditions at the node z with respect to all normals on neighbouring $\mathcal{E}_z \cap \mathcal{E}_N$. (Here, p is assumed to be a flux and not necessarily a gradient.)

Remark 2.4. The condition $p(z) \in \mathcal{A}_z$ with simultaneous Dirichlet and Neumann conditions, i.e. with $\mathcal{E}_z \cap \mathcal{E}_N \neq \emptyset \neq \mathcal{E}_z \cap \mathcal{E}_D$, is based on the interpretation of p as both a flux and a gradient. Hence, the model example is the Laplace equation with mixed boundary conditions. Nonconforming finite element methods require the case $\mathcal{E}_D \neq \emptyset$ [CB, CBJ].

Remark 2.5. It is by no means obvious that averaging concerns the fluxes and the gradients simultaneously. The positive examples in [CBJ, CF3, CF4, BC2, CA] may be seen as exceptions. In general, the flux and the gradient approximations may be averaged separately. In the latter case we encounter $\mathcal{E}_N = \emptyset$ or $\mathcal{E}_D = \emptyset$.

2.3. **Discrete spaces.** On each element there exists a set of shape functions, namely, $P_{(k)}(T) := P_k(T)$ if T is triangular and $P_{(k)}(T) := Q_k(T)$ if T is rectangular; $P_k(T)$ and $Q_k(T)$ denote algebraic polynomials on $T \subseteq \mathbb{R}^d$ of total and partial degree $\leq k$, respectively. Furthermore, for each $T \in \mathcal{T}$ let $P(T)$ satisfy $P_{(0)}(T) \subset P(T) \subset P_{(1)}(T)$. Then, set

$$\begin{aligned}
\mathcal{L}^k(\mathcal{T}) &:= \{v_h \in L^\infty(\Omega) : \forall T \in \mathcal{T},\, v_h|_T \in P_{(k)}(T)\} \text{ for } k = 0, 1,\\
\mathcal{S}^1(\mathcal{T}) &:= \mathcal{L}^1(\mathcal{T}) \cap C(\Omega) = \operatorname{span}\{\varphi_z : z \in \mathcal{N}\},\\
P_h &:= P(\mathcal{T}) := \{p_h \in L^\infty(\Omega)^d : \forall T \in \mathcal{T},\, p_h|_T \in P(T)\} \subseteq \mathcal{L}^1(\mathcal{T})^d,\\
\mathcal{Q}_h &:= \{q_h \in \mathcal{S}^1(\mathcal{T})^d : \forall z \in \mathcal{N} \cap \Gamma,\, q_h(z) \in \mathcal{A}_z\}.
\end{aligned}$$

The nodal basis functions $(\varphi_z : z \in \mathcal{N})$ are defined by $\varphi_z \in \mathcal{S}^1(\mathcal{T})$ with $\varphi_z(z) = 1$ and $\varphi_z(x) = 0$ for all $z, x \in \mathcal{N}$ with $x \neq z$. Without further explicit notice, we shall make frequent use of

$$0 \leq \varphi_z \leq 1, \quad \operatorname{supp} \varphi_z = \overline{\omega}_z, \quad \text{and} \quad \textstyle\sum_{z\in\mathcal{N}} \varphi_z = 1.$$

2.4. **Averaging operators.** Given $p_h \in P_h$ (not necessarily globally continuous), the operator $A : P_h \to \mathcal{Q}_h$ is supposed to average p_h on each patch ω_z and adopt to boundary conditions. Therefore,

$$Ap_h := \sum_{z\in\mathcal{N}} A_z(p_h|_{\omega_z})\,\varphi_z \quad \text{and} \quad A_z := \pi_z \circ M_z : \mathcal{L}^1(\mathcal{T}_z)^d \to \mathbb{R}^d.$$

Recall that $\mathcal{L}^1(\mathcal{T}_z)$ denotes the $\mathcal{T}_z$ piecewise polynomials of degree ≤ 1 and that $p_h|_{\omega_z}$ belongs to $\mathcal{L}^1(\mathcal{T}_z)$. The local operator A_z is the composition of an averaging process $M_z : \mathcal{L}^1(\mathcal{T}_z)^d \to \mathbb{R}^d$ and the orthogonal projection $\pi_z : \mathbb{R}^d \to \mathbb{R}^d$ onto the affine subspace $\mathcal{A}_z \subset \mathbb{R}^d$.

The operator M_z is supposed to be linear and exact on continuous functions in $\mathcal{P}(\mathcal{T}_z) := \{p_h \in L^\infty(\omega_z)^d : \forall T \in \mathcal{T}_z,\, p_h|_T \in P(T)\}$; i.e.

$$M_z(f) = f(z) \quad \text{for all } f \in \mathcal{P}(\mathcal{T}_z) \cap C(\omega_z)^d \text{ and } z \in \mathcal{N}. \tag{2.2}$$

The master example for M_z reads

$$M_z(f) := \sum_{T\in\mathcal{T}_z} \lambda_{z,T}(f|_T)(z) \quad \text{for all } f \in \mathcal{P}(\mathcal{T}_z),\, z \in \mathcal{N}. \tag{2.3}$$

A necessary condition for (2.2) on the real coefficients $(\lambda_{z,T} : T \in \mathcal{T}_z)$ in (2.3) reads

$$\textstyle\sum_{T\in\mathcal{T}_z} \lambda_{z,T} = 1.$$

For a practical realization of A_z and for numerical examples we refer to [CB, CF3, CF4].

2.5. **Estimators.** Given the spaces P_h and $\mathcal{Q}_h$ of subsection 2.3 and the averaging operator $A : P_h \to \mathcal{Q}_h$ of subsection 2.4, we define, for any fixed $p_h \in P_h$, the averaging estimators

$$\eta_M := \min_{r_h\in\mathcal{Q}_h} \|p_h - r_h\|_{L^2(\Omega)} \leq \eta_A := \|p_h - Ap_h\|_{L^2(\Omega)}.$$

3. Preliminaries

This section establishes some tools in an abstract frame to clarify the arguments below. Attention is on the arising constants: In contrast to earlier work based on a compactness arguments which led to unknown constants, we aim to quantify C_{eff}.

3.1. **Ascoli's lemma.** Given a linear and bounded map $L : H \to \mathbb{R}^n$ in a (real) Hilbert space H with norm $\|\cdot\|$, there holds, for $f \in H$,

$$|L(f)| \leq \|L\| \operatorname{dist}(f; \ker L). \tag{3.1}$$

Here, $\operatorname{dist}(f; \ker L) := \min\{\|f - g\| : g \in \ker L\}$ is the distance to the (closed) kernel $\ker L$ of L and

$$\|L\| := \sup_{g \in X \setminus \{0\}} |L(g)|/\|g\| \tag{3.2}$$

is the operator norm of L; $|\cdot|$ is the Euclidean norm in $\mathbb{R}^n$. The proof of (3.1) is by definition of $\|L\|$,

$$|L(f)| = |L(f-g)| \leq \|L\| \, \|f - g\| \quad \text{for all } g \in \ker L.$$

In case $n = 1$, i.e. $L \in H^*$, there even holds equality in (3.1), which is known as Ascoli's lemma. A simple proof for the converse inequality of (3.1) follows for $g \in H$ with $\|g\| = 1$, $L(g) = \|L\|$ and so with $f - g\,L(f)/\|L\| \in \ker L$ from

$$\operatorname{dist}(f; \ker L) \leq \|f - \Big(f - g\,L(f)/\|L\|\Big)\| = |L(f)|/\|L\|.$$

Suppose $n \geq 1$ again, let e_j be the jth canonical unit vector in $\mathbb{R}^n$, and set $L_j := e_j \cdot L$. Then there holds

$$|L_j(f)| = \|L_j\| \operatorname{dist}(f; \ker L_j).$$

The sum over all $j = 1, \ldots, n$ squared components shows

$$|L(f)|^2 = \textstyle\sum_{j=1}^n \|L_j\|^2 \operatorname{dist}(f; \ker L_j)^2 \quad \text{for all } f \in H. \tag{3.3}$$

Compared with (3.1), the operator norm $\|L_j\|$ in (3.3) is smaller than $\|L\|$ while the kernel of L_j is larger than $\ker L \subseteq \ker(L_j)$.

3.2. **Strengthened Cauchy inequality.** Let H be a (real) Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and norm $\|\cdot\|$ and let V and W be closed subspaces of H. Owing to the Cauchy inequality, the constant $\gamma_{V,W}$,

$$\gamma_{V,W} := \sup_{v \in V \setminus \{0\}} \sup_{w \in W \setminus \{0\}} \langle v, w \rangle / (\|v\| \, \|w\|) \leq 1, \tag{3.4}$$

defines the angle $\angle(v, w)$ between v and w by $0 \leq \cos(\angle(v, w)) = \gamma_{V,W} \leq 1$. The spaces V and W satisfy a strengthened Cauchy inequality if $\gamma_{V,W} < 1$, that is, if $\angle(V, W)$ is positive.

Lemma 3.1 ([B])**.** For a constant c with $0 < c < 1$ and $\gamma_{V,W}$ from (3.4), the following assertions (a), (b), and (c) are pairwise equivalent.

(a) $\gamma_{V,W} \leq c$;
(b) $\forall v \in V$, $\sqrt{1 - c^2}\|v\| \leq \operatorname{dist}(v; W)$;
(c) $\forall v \in V \, \forall w \in W$, $\sqrt{(1 - c^2)/2}(\|v\| + \|w\|) \leq \|v + w\|$. □

We are particularly interested in (a)⇔(b) also considered in [AO].

Lemma 3.2. Let X and Y be closed linear subspaces of a Hilbert space H with scalar product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$. Set

$$V := \{x \in X : \forall a \in X \cap Y, \langle x, a \rangle = 0\} = X \cap (X \cap Y)^{\perp} \tag{3.5}$$

and suppose that V and Y are nontrivial and that V has positive finite dimension. Set

$$\gamma_{V,Y} := \sup_{v \in V \setminus \{0\}} \sup_{y \in Y \setminus \{0\}} \langle v, y \rangle / (\|v\| \, \|y\|). \tag{3.6}$$

Then $0 \le \gamma_{V,Y} < 1$ and $\gamma_{V,Y} = \langle v, y \rangle$ for some $v \in V$ and $y \in Y$ with $\|v\| = 1 = \|y\|$. Moreover,

$$\operatorname{dist}(x; X \cap Y) \le (1 - \gamma_{V,Y}^2)^{-1/2} \operatorname{dist}(x; Y) \quad \text{for all } x \in X \tag{3.7}$$

and the factor $(1 - \gamma_{V,Y}^2)^{-1/2}$ is optimal in the sense that (3.7) fails to hold for any smaller constant.

Proof. Owing to the definition in (3.6) there exist sequences (x_j) and (y_j) in V and Y, respectively, with $\|x_j\| = 1 = \|y_j\|$ and

$$\lim_{j \to \infty} \langle x_j, y_j \rangle = \gamma_{V,Y}.$$

Since (x_j) and (y_j) are bounded in a Hilbert space, there exists a subsequence (not relabeled) with $(x_j) \to x$ and $(y_j) \rightharpoonup y$ in H. The strong convergence of (x_j) follows from the finite dimension of V. Hence, $\|x\| = 1 \ge \|y\|$ and $\lim_{j \to \infty} \langle x_j, y_j \rangle = \langle x, y \rangle$. If $y \ne 0$, we have

$$\gamma_{V,Y} = \langle x, y \rangle \le \langle x, y \rangle / (\|x\| \, \|y\|) \le \gamma_{V,Y}.$$

(The last inequality follows from (3.6) and $x \in V$, $y \in Y$.) Hence we have $\gamma_{V,Y} = \langle x/\|x\|, y/\|y\| \rangle$ for some $x/\|x\| \in V$ and $y/\|y\| \in Y$ with norm 1 if $0 < \gamma_{V,Y} < \infty$.

If $y = 0$, $\gamma_{V,Y} = 0$ and each $v \in V$ is perpendicular to Y.

In both cases, $\gamma_{V,Y} = \langle v, y \rangle$ for some $v \in V$ and $y \in Y$ with $\|v\| = \|y\| = 1$. This proves the attainment result.

A Cauchy inequality shows $\gamma_{V,Y} \le 1$. If $\gamma_{V,Y} = 1 = \langle v, y \rangle$ for $v \in V$ and $y \in Y$ with $\|v\| = 1 = \|y\|$, we have equality in the Cauchy inequality and hence $v = y$. Thus, $v \in V \cap Y \subseteq X \cap Y$ and so $\|v\|^2 = 0$ owing to (3.5). This contradicts $\|v\| = 1$ and proves $\gamma_{V,Y} \ne 1$.

It remains to apply Lemma 3.1 for V and $W := Y$. Then $\gamma_{V,Y}$ in (3.4) and (3.6) coincide and the equivalence (a)$\Leftrightarrow$(b) of Lemma 3.1 proves, first,

$$\|v\| \le (1 - \gamma_{V,Y}^2)^{-1/2} \operatorname{dist}(v; Y) \quad \text{for all } v \in V \tag{3.8}$$

and, second, that the constant factor $(1 - \gamma_{V,Y}^2)^{-1/2}$ in (3.8) cannot be smaller.

Given $x \in X$ and the closed subspace $X \cap Y$ of X, there exists an orthogonal decomposition

$$x = v + w \quad \text{with } v \in V \text{ and } w \in X \cap Y.$$

Moreover, $\operatorname{dist}(x; X \cap Y) = \|v\|$ and $\operatorname{dist}(v; Y) = \operatorname{dist}(v + w; w + Y) = \operatorname{dist}(x; Y)$. This and (3.8) conclude the proof. $\square$

TABLE 1. Mass matrices $M(T)$ (scaled with $|T|^{-1}$) for some elements T and their eigenvalues $\lambda_1 \dots, \lambda_m$ and $\lambda(T)$ of Lemma 3.3

pT	$M(T)$	Eigenvalues	$\lambda(T)$
interval	$1/6 \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$	1, 3	1
triangle	$1/12 \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$	1/12, 1/12, 1/4	12
parallelogram	$1/36 \begin{bmatrix} 4 & 2 & 1 & 2 \\ 2 & 4 & 2 & 1 \\ 1 & 2 & 4 & 2 \\ 2 & 1 & 2 & 4 \end{bmatrix}$	1/36, 1/12, 1/12, 1/4	36
tetrahedron	$1/20 \begin{bmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{bmatrix}$	1/20, 1/20, 1/20, 1/5	20

3.3. **Eigenvalues of mass matrices.** This subsection summarises a few inequalities and the constants therein. For each element $T \in \mathcal{T}$ with volume $|T| = \mathcal{L}^d(T)$ we associate $m := \operatorname{card}(\mathcal{N} \cap T)$ nodal basis functions $\varphi_1, \dots, \varphi_m$ called shape functions with

$$\int_T \varphi_j \, dx = |T|/m \quad \text{for } j = 1, \dots, m \tag{3.9}$$

(as $\sum_{j=1}^m \varphi_j = 1$ and the forms of $\varphi_1, \dots, \varphi_m$ are identical). Scaled with $|T|^{-1}$, the mass matrix reads

$$M(T) := (\int_T \varphi_j \varphi_k \, dx/|T| : \, j, k = 1, \dots, m). \tag{3.10}$$

Table 1 displays some mass matrices and their eigenvalues $\lambda_1, \dots, \lambda_m$.

Lemma 3.3. Suppose $T \in \mathcal{T}$ and $f \in P_{(1)}(T)^d$. Then

$$|T| \sum_{z \in \mathcal{N} \cap T} |f(z)|^2 \le \lambda(T) \|f\|^2_{L^2(T)} \tag{3.11}$$

where $\lambda(T) = 1/\lambda_1$ for the minimal eigenvalue λ_1 of the matrix (3.10) displayed in Table 1.

Proof. Let $f_j := e_j \cdot f$ be the jth component of f and let $\{z_1, \dots, z_m\} = \mathcal{N} \cap T$ denote the vertices of T. With the m components $\xi_k := f_j(z_k)$ of $\xi \in \mathbb{R}^m$ and a standard estimation of the Rayleigh quotient there holds

$$\lambda_1 \sum_{z \in \mathcal{N} \cap T} f_j(z)^2 = \lambda_1 |\xi|^2 \le \xi \cdot M(T)\xi = |T|^{-1} \|f_j\|^2_{L^2(T)}.$$

The sum over all components $j = 1, \dots, d$ verifies assertion (3.11). $\square$

4. Equivalence of η_M and η_A

This section is devoted to the proof of the equivalence of η_M and η_A under the present assumptions. A discussion of the constant C_{eff} follows in Section 5. Theorem 4.1 covers efficiency (1.5) for conforming, nonconforming, and mixed finite element methods [CB].

Theorem 4.1. *There exists a mesh-size independent positive constant C_{eff} with*

$$\eta_M \leq \eta_A \leq C_{\text{eff}}\, \eta_M.$$

Proof. The first inequality is obvious and the proof concerns the second. Throughout the first step and main part of the proof let T denote a fixed element. Set

$$p_h|_T = \sum_{z\in\mathcal{N}\cap T} p_z\varphi_z|_T \quad \text{and} \quad q_h := Ap_h = \sum_{z\in\mathcal{N}} q_z\varphi_z \text{ for } q_z := A_z(p_h|_{\omega_z}).$$

(Notice that the representation of p_h is local on the fixed element T; p_h may be discontinuous on Ω and so has different coefficients on different elements.) A Cauchy inequality in $\mathbb{R}^m$, $m = \text{card}(\mathcal{N}\cap T)$, shows, pointwise on T,

$$\begin{aligned} |p_h - q_h|^2 &= |\sum_{z\in\mathcal{N}\cap T} \varphi_z(p_z - q_z)|^2 \\ &\leq (\sum_{z\in\mathcal{N}\cap T}\varphi_z)(\sum_{z\in\mathcal{N}\cap T}\varphi_z|p_z - q_z|^2) \\ &= \sum_{z\in\mathcal{N}\cap T}\varphi_z|p_z-q_z|^2. \end{aligned} \tag{4.1}$$

Since $q_z = \pi_z(m_z)$ for $m_z := M_z(p_h|_{\omega_z})$ and $p_z - \pi_z(p_z)\perp\mathcal{V}_z$ in $\mathbb{R}^d$, there holds

$$|p_z - q_z|^2 = |p_z - \pi_z(p_z)|^2 + |\pi_z(p_z) - \pi_z(m_z)|^2.$$

With any $r_z \in \mathcal{A}_z = \pi_z(0) + \mathcal{V}_z$ and $\text{Lip}(\pi_z) \leq 1$, this yields

$$|p_z - q_z|^2 \leq |p_z - r_z|^2 + |p_z - m_z|^2. \tag{4.2}$$

The combination of (4.1)-(4.2) is integrated over the fixed T and shows

$$\|p_h - q_h\|^2_{L^2(T)} \leq \sum_{z\in\mathcal{N}\cap T}|p_z - r_z|^2\int_T\varphi_z\,dx + \sum_{z\in\mathcal{N}\cap T}|p_z-m_z|^2\int_T\varphi_z\,dx.$$

With (3.9) and Lemma 3.3 this gives, for $r_h := \sum_{z\in\mathcal{N}} r_z\varphi_z \in \mathcal{Q}_h$,

$$\|p_h - q_h\|^2_{L^2(T)} \leq \frac{\lambda(T)}{m}\|p_h - r_h\|^2_{L^2(T)} + \frac{|T|}{m}\sum_{z\in\mathcal{N}\cap T}|p_z - m_z|^2. \tag{4.3}$$

The second step focuses on the estimation of $p_z - m_z$ and introduces the finite-dimensional Hilbert space $X := P(\mathcal{T}_z) \subseteq \mathcal{L}^1(\mathcal{T}_z)^d$ with the inner product $\langle\cdot,\cdot\rangle$,

$$\langle f, g\rangle := \int_{\omega_z}\varphi_z\, f\cdot g\,dx/|T| \quad \text{for } f, g \in L^2(\omega_z)^d =: H. \tag{4.4}$$

Define $\delta_{T,z}(f) := (f|_T)(z)$ for all $f \in X$ and consider

$$L_{T,z} := \delta_{T,z} - M_z : X \to \mathbb{R}^d \quad \text{linear} \tag{4.5}$$

and continuous with the bound

$$\|L_{T,z}\| := \sup_{f\in P(\mathcal{T}_z)\setminus\{0\}}\Big|f|_T(z) - M_z(f)\Big|/(|T|^{-1/2}\|\varphi_z^{1/2}f\|_{L^2(\omega_z)}). \tag{4.6}$$

A scaling argument shows that $\|L_{T,z}\|$ does not depend on the diameter of ω_z because of the factor $|T|^{-1/2}$. (Details on the constant $\|L_{T,z}\|$ from (4.6) follow for specific examples after the proof.) Since M_z is exact on $\mathcal{P}(\mathcal{T}_z) \cap C(\omega_z)^d$,

$$\mathbb{R}^d \subseteq \mathcal{P}(\mathcal{T}_z) \cap C(\omega_z)^d \subseteq \ker L =: Z \subseteq X \quad \text{and} \quad Y := \mathcal{S}^1(\mathcal{T}_z)^d.$$

Ascoli's lemma (formula (3.1)) shows

$$|p_z - m_z| = |L_{T,z}(p_h)| \le \|L_{T,z}\| \operatorname{dist}(p_h|_{\omega_z}; Z). \tag{4.7}$$

Lemma 3.2 and $X \cap Y \subseteq Z$ prove $0 \le \gamma < 1$ for the constant γ of (3.6) and

$$\operatorname{dist}(p_h|_{\omega_z}; Z) \le \operatorname{dist}(p_h|_{\omega_z}; X \cap Y) \le (1-\gamma^2)^{-1/2} \operatorname{dist}(p_h|_{\omega_z}; Y). \tag{4.8}$$

(The constant (3.6) will be discussed at the end of this section for specific examples.)

Step three combines (4.3) and (4.7)-(4.8) with

$$\operatorname{dist}(p_h|_{\omega_z}; \mathcal{S}^1(\mathcal{T}_z)^d) \le |T|^{-1/2} \|\varphi_z^{1/2}(p_h - r_h)\|_{L^2(\omega_z)}$$

and (writing γ_z for γ) results in

$$\|p_h - q_h\|^2_{L^2(T)} \le \lambda(T)/m \, \|p_h - r_h\|^2_{L^2(T)} + \sum_{z \in \mathcal{N} \cap T} \|L_{T,z}\|^2 \, (1-\gamma_z^2)^{-1}/m \, \|\varphi_z^{1/2}(p_h - r_h)\|^2_{L^2(\omega_z)}.$$

In step four, the sum over all elements $T \in \mathcal{T}$ and the fact

$$\sum_{z \in \mathcal{N}} \|\varphi_z^{1/2}(p_h - r_h)\|^2_{L^2(\omega_z)} = \|p_h - r_h\|^2_{L^2(\Omega)}$$

show the assertion

$$\eta_A = \|p_h - q_h\|_{L^2(\Omega)} \le C_{\text{eff}} \, \|p_h - r_h\|_{L^2(\Omega)} \quad \text{for all } r_h \in \mathcal{Q}_h. \tag{4.9}$$

The constant C_{eff} depends on $m = m_T$, $\lambda(T)$, and $\|L_{T,z}\|^2/(1-\gamma_z^2)$ as

$$C_{\text{eff}}^2 = \max_{T \in \mathcal{T}} \big(\lambda(T) + \max_{z \in \mathcal{N} \cap T} \|L_{T,z}\|^2/(1-\gamma_z^2)\big)/m_T. \tag{4.10}$$

This concludes the proof of $\eta_A \le C_{\text{eff}} \, \eta_M$. □

5. Example

The constant C_{eff} and its possible dependence on mesh will be studied for the P_1 FEM with piecewise constant discrete fluxes. Recall that

$$X := \mathcal{P}(\mathcal{T}_z) \subseteq \mathcal{L}^1(\mathcal{T}_z)^d \subset H := L^2(\omega_z)^d$$

and $Y = \mathcal{S}^1(\mathcal{T}_z)^d$ with the scalar product (4.4) on H.

The following lemma provides coarse but uniform estimates of eigenvalues which could be computed as a function of $\operatorname{card}(\mathcal{T}_z)$.

Lemma 5.1. Suppose that $\mathcal{P}(\mathcal{T}_z) = \mathcal{L}^0(\mathcal{T}_z)^d$ and that $\mathcal{T}_z$ consists of simplices in $\mathbb{R}^d$. Then the constant $\gamma = \gamma_z \ge 0$ from (3.5)-(3.6) satisfies

$$\gamma^2 \le 5/6 \text{ for } d = 2 \quad \text{and} \quad \gamma^2 \le 9/10 \text{ for } d = 3.$$

Proof. Given any $v_h \in \mathcal{L}^0(\mathcal{T}_z)^d$ and $y_h \in \mathcal{S}^1(\mathcal{T}_z)^d$, set $v_T := v_h|_T \in \mathbb{R}^d$ and $y_T = \int_T \varphi_z\, y_h\, dx$ for $T \in \mathcal{T}_z$. Then, (3.6) reads

$$\gamma^2 = \max_{v_h, y_h} \left(\sum_{T\in\mathcal{T}_z} v_T \cdot y_T \right)^2 \Big/ \left(\left(\sum_{T\in\mathcal{T}_z} |T||v_T|^2/m \right) \left(\int_{\omega_z} \varphi_z |y_h|^2\, dx \right) \right)$$

where, by definition of V, $v_h \in V$ satisfies $\sum_{T\in\mathcal{T}_z} |T|\, v_T = 0$. Consequently, the sum

$$\textstyle\sum_{T\in\mathcal{T}_z} v_T \cdot y_T$$

does not depend on an additive constant in y_h which, therefore, is determined to minimise $\int_{\omega_z} \varphi_z\, |y_h|^2\, dx$. This results in the condition $\int_{\omega_z} \varphi_z\, y_h\, dx = 0$; i.e.

$$\textstyle\sum_{T\in\mathcal{T}_z} y_T = 0. \tag{5.1}$$

A Cauchy inequality yields

$$\gamma^2 = \max_{y_h} \left(m \sum_{T\in\mathcal{T}} |y_T|^2/|T| \right) \Big/ \int_{\omega_z} \varphi_z\, |y_h|^2\, dx \tag{5.2}$$

and equality is indeed attained for $v_T = y_T/|T|$ (compatible with $v_h \in V$ and (5.1)). Given y_h in $\mathcal{S}^1(\mathcal{T}_z)^2$ with (5.1) and nodal values $y_0 = y_h(z)$, $y_{a,T} = y_h(a)$, $y_{b,T} = y_h(b)$ on $T = \operatorname{conv}\{z, a, b\} \in \mathcal{T}_z$ for $d = 2$, a straightforward calculation shows

$$y_T = |T|(2y_0 + y_{a,T} + y_{b,T})/12 \quad \text{for } T \in \mathcal{T}_z.$$

This and (5.1) plus a Cauchy inequality yield

$$\begin{aligned} 12 \sum_{T\in\mathcal{T}_z} |y_T|^2/|T| &= \sum_{T\in\mathcal{T}_z} y_T \cdot (2y_0 + y_{a,T} + y_{b,T}) \\ &= \sum_{T\in\mathcal{T}_z} y_T \cdot (y_{a,T} + y_{b,T}) \\ &\le \left(\sum_{T\in\mathcal{T}_z} |y_T|^2/|T| \right)^{1/2} \left(\sum_{T\in\mathcal{T}_z} |T|\, |y_{a,T} + y_{b,T}|^2 \right)^{1/2} \end{aligned}$$

and so (divide by $\left(\sum_{T\in\mathcal{T}_z} |y_T|^2/|T|\right)^{1/2}$ and square) leads to

$$144 \sum_{T\in\mathcal{T}_z} |y_T|^2/|T| \le \sum_{T\in\mathcal{T}_z} |T||y_{a,T} + y_{b,T}|^2. \tag{5.3}$$

The summand on the right-hand side is

$$\begin{aligned} &|T|\, |y_{a,T} + y_{b,T}|^2 \\ &\quad \le 2|T| \left(2|y_0|^2 + |y_{a,T}|^2 + |y_{b,T}|^2 \right) \\ &\quad = 120 \int_T \varphi_z\, |y_h|^2\, dx - 288 |y_T|^2/|T|. \end{aligned} \tag{5.4}$$

The latter equality follows with lengthy but straightforward calculation with the well-known formula $\int_T \lambda_1^\alpha \lambda_2^\beta \lambda_3^\gamma\, dx = 2|T|\, \alpha!\beta!\gamma!/(2+\alpha+\beta+\gamma)!$ for the barycentric coordinates $\lambda_1, \lambda_2, \lambda_3$ on the triangle T. The combination of (5.3)-(5.4) verifies

$$\sum_{T\in\mathcal{T}_z} |y_T|^2/|T| \le 5/18 \int_{\omega_z} \varphi_z\, |y_h|^2\, dx.$$

Using this in (5.2) shows $\gamma^2 \le 5/6$. The proof for $d = 3$ follows with the same arguments modified for $y_T = |T|\,(2y_0 + y_{a,T} + y_{b,T} + y_{c,T})/20$, etc.; the details are omitted. $\square$

To study $\|L_{z,T}\|$, let M_z be given by (2.3); i.e.

$$M_z(f) = \textstyle\sum_{T\in\mathcal{T}_z} \lambda_{z,T}\, f_T \quad \text{for } f|_T = f_T \in \mathbb{R}^d,\ T \in \mathcal{T}_z, \text{ and } f \in \mathcal{L}^0(\mathcal{T}_z)^d.$$

The real coefficients $\lambda_{z,T}$ sum up to $1 = \sum_{T\in\mathcal{T}_z} \lambda_{z,T}$ (some are possibly negative). For comparison, a popular choice for the coefficient $\lambda_{z,T}$ reads

$$\mu_{z,T} := |T|/|\omega_z| \quad \text{for } T \in \mathcal{T}_z. \tag{5.5}$$

Lemma 5.2. Suppose (4.4)-(4.6) for fixed $z \in T \cap \mathcal{N}$ and that $\mathcal{P}(\mathcal{T}_z) = \mathcal{L}^0(\mathcal{T}_z)^d$ and that $\mathcal{T}_z$ consists of simplices. Then $m = d + 1$ and

$$\|L_{z,T}\|^2 = m\left((1-\lambda_{z,T})^2 + \sum_{K\in\mathcal{T}_z\setminus\{T\}} \lambda_{z,K}^2 \mu_{z,T}/\mu_{z,K}\right).$$

Proof. Given any $f \in \mathcal{L}^0(\mathcal{T}_z)^d$ (write $f_K := f_K$ for each $K \in \mathcal{T}$),

$$L_{z,T}(f) = (1-\lambda_{z,T})f_T - \textstyle\sum_{K\in\mathcal{T}_z\setminus\{T\}} \lambda_{z,K} f_K$$

is independent of a global additive constant in f. To minimise $\|\varphi_z^{1/2} f\|$, this constant is such that $\int_{\omega_z} \varphi_z\, f\, dx = 0$. Hence

$$\textstyle\sum_{K\in\mathcal{T}_z} |K|\, f_K = 0 \tag{5.6}$$

and (with an argument f in $\mathcal{L}^0(\mathcal{T}_z)^d\setminus\{0\}$ with (5.6) in the supremum)

$$\|L_{z,T}\|^2 = \sup\nolimits_f |T|\,|L_{z,T}(f)|^2 / \left(\textstyle\sum_{M\in\mathcal{T}_z} |f_M|^2\, |M|/m\right).$$

A Cauchy inequality shows

$$|L_{z,T}(f)| \le \left(\textstyle\sum_{M\in\mathcal{T}_z} |f_M|^2 |M|\right)^{1/2} \left((1-\lambda_{z,T})^2/|T| + \textstyle\sum_{K\in\mathcal{T}_z\setminus\{T\}} \lambda_{z,K}^2/|K|\right)^{1/2}.$$

Equality holds for $f_T = (1-\lambda_{z,T})/|T|\, e$ and $f_K = -\lambda_K/|K|\, e$ for any other $K \in \mathcal{T}_z\setminus\{T\}$ and some fixed unit vector $e \in \mathbb{R}^d$. Since this choice of f satisfies (5.6), there holds

$$\|L_{z,M}\|^2 = m|T|\left((1-\lambda_{z,T})^2/|T| + \textstyle\sum_{K\in\mathcal{T}_z\setminus\{T\}} \lambda_{z,K}^2/|K|\right). \quad \square$$

The following consequence gives an estimate for the choice (5.5) and indicates that this choice is optimal.

Corollary 5.3. Under the assumptions of the preceding two lemmas (satisfied for all $z \in \mathcal{N}$) and for $\mu_{z,T} = \lambda_{z,T}$ there holds

$$\eta_M \le \eta_A \le \sqrt{10}\,\eta_M \text{ for } d = 2 \quad \text{and} \quad \eta_M \le \eta_A \le \sqrt{15}\,\eta_M \text{ for } d = 3.$$

Proof. The estimates follow from Theorem 4.1 and Lemmas 5.1 and 5.2 with $\|L_{z,T}\|^2 = m(1-\mu_{z,T}) \le m$ and $\lambda(T) = 12,\ 20$ for $d = 2, 3$. $\square$

References

[AO] M. Ainsworth, J.T Oden: *A posteriori error estimation in finite element analysis*, John Wiley & Sons, New York, 2001. MR **2003b:**65001

[BS] I. Babuška, T. Strouboulis: *The Finite Element Method and its Reliability.* Oxford University Press, 2001. MR **2002k:**65001

[BC1] S. Bartels, C. Carstensen: Each averaging technique yields reliable a posteriori error control in FEM on unstructured grids. Part II: Higher order FEM. *Math. Comp.* **71** (2002) 971-994. MR **2003e:**65207

[BC2] S. Bartels, C. Carstensen: Averaging techniques yield reliable a posteriori finite element error control for obstacle problems. *Numer. Math.* (2003) to appear.

[BR] R. Becker, R. Rannacher: A feed-back approach to error control in finite element methods: basic analysis and examples. *East-West Journal of Numerical Mathematics* **4** Number 4 (1996) 237-264. MR **98m:**65185

[B] D. Braess: Enhanced assumed strain elements and locking in membrane problems, *Comp. Meths. Appl. Mech. Engrg.* **165** (1998) 155-174. MR **2000j:**74084

[C] C. Carstensen: Quasi-interpolation and a posteriori error analysis in finite element method. *M2AN* **33** (1999) 1187–1202. MR **2001a:**65135

[CA] C. Carstensen, J. Alberty: Averaging techniques for reliable a posteriori FE-error control in elastoplasticity with hardening. *Comput. Methods Appl. Mech. Engrg.* **192** (2003) 1435–1450.

[CB] C. Carstensen, S. Bartels: Each averaging technique yields reliable a posteriori error control in FEM on unstructured grids, part I: Low order conforming, nonconforming, and mixed FEM. *Math. Comp.* **71** (2002) 945-969. MR **2003e:**65212

[CBJ] C. Carstensen, S. Bartels, S. Jansche: A posteriori error estimates for nonconforming finite element methods. *Numer. Math.* **92** (2002) 233–256. MR

[CF1] C. Carstensen, S.A. Funken: Constants in Clément-interpolation error and residual based a posteriori estimates in finite element methods, *East-West Journal of Numerical Analysis*, **8**, 3, 153–256. MR **2002a:**65173

[CF2] C. Carstensen, S.A. Funken: Fully reliable localised error control in the FEM, *SIAM J. Sci. Comp.*, **21**, 4, 1465–1484. MR **2000k:**65205

[CF3] C. Carstensen, S.A. Funken: Averaging technique for FE - a posteriori error control in elasticity. Part I: Conforming FEM. *Comput. Methods Appl. Mech. Engrg. 190 (2001)*, pp. 2483–2498, Part II: λ-independent estimates. *Comput. Methods Appl. Mech. Engrg.* **190** (2001) 4663–4675. Part III: Locking-free nonconforming FEM. *Comput. Methods Appl. Mech. Engrg.* **191** (2001), no. 8-10, 861–877. MR **2002a:**74114, MR **2002d:**65140, MR **2002j:**65106

[CF4] C. Carstensen, S.A. Funken: A posteriori error control in low-order finite element discretisations of incompressible stationary flow problems. *Math. Comp.* **70** (2001) 1353–1381. MR **2002f:**65157

[CV] C. Carstensen, R. Verfürth: Edge residuals dominate a posteriori error estimates for low order finite element methods, *SIAM J. Numer. Anal.* **36**, 5,(1999) 1571–1587. MR **2000g:**65115

[N] R. Nochetto: Removing the saturation assumption in a posteriori error analysis. *Rend., Sci. Mat. Appl.*, A 127, 67-82 (1994). MR **95c:**65187

[R1] R. Rodriguez: Some remarks on Zienkiewicz–Zhu estimator. *Int. J. Numer. Meth. in PDE* **10** (1994) 625–635. MR **95e:**65103

[R2] R. Rodriguez: A posteriori error analysis in the finite element method. *Finite element methods. 50 years of the Courant element. Conference held at the University of Jyvaeskylae, Finland, 1993.* Inc. Lect. Notes Pure Appl. Math. 164, 389-397 (1994). MR **95g:**65158

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

[ZZ] O.C. Zienkiewicz, J.Z. Zhu: *A simple error estimator and adaptive procedure for practical engineering analysis*, *Int. J. Numer. Meth. Engrg.*, **24** (1987) 337–357. MR **87m:**73055

Institute for Applied Mathematics and Numerical Analysis, Vienna University of Technology, Wiedner Hauptstrasse 8-10, A-1040 Vienna, Austria

E-mail address: `Carsten.Carstensen@tuwien.ac.at`