INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 3, Number 3, Pages 333–347

RELIABLE AND EFFICIENT AVERAGING TECHNIQUES AS UNIVERSAL TOOL FOR A POSTERIORI FINITE ELEMENT ERROR CONTROL ON UNSTRUCTURED GRIDS

CARSTEN CARSTENSEN

(Communicated by Zhimin Zhang)

Abstract. The striking simplicity of averaging techniques in a posteriori error control of finite element methods as well as their amazing accuracy in many numerical examples over the last decade have made them an extremely popular tool in scientific computing. Given a discrete stress or flux p_h and a post-processed approximation $A(p_h)$, the a posteriori error estimator reads $\eta_A := \|p_h - A(p_h)\|$. There is not even a need for an equation to compute the estimator η_A and hence averaging techniques are employed everywhere. The most prominent example is occasionally named after Zienkiewicz and Zhu, and also called gradient recovery but preferably called averaging technique in the literature.

The first mathematical justification of the error estimator η_A as a computable approximation of the (unknown) error $\|p - p_h\|$ involved the concept of superconvergence points. For highly structured meshes and a very smooth exact solution p, the error $\|p - A(p_h)\|$ of the post-processed approximation Ap_h may be (much) smaller than $\|p - p_h\|$ of the given p_h. Under the assumption that $\|p - A(p_h)\| =$ h.o.t. is in relative terms sufficiently small, the triangle inequality immediately verifies reliability, i.e.,

$$\|p - p_h\| \le C_{\text{rel}}\, \eta_A + \text{ h.o.t.},$$

and efficiency, i.e.,

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

of the averaging error estimator η_A. However, the required assumptions on the symmetry of the mesh and the smoothness of the solution essentially contradict the use of adaptive grid refining when p is singular and the proper treatment of boundary conditions remains unclear.

This paper aims at an actual overview on the reliability and efficiency of averaging a posteriori error control for unstructured grids. New aspects are new proofs of the efficiency of *all* averaging techniques and for *all* problems.

Key Words. a posteriori error estimate, efficiency, finite element method, gradient recovery, averaging operator, mixed finite element method, non-conforming finite element method

1. Overview

The outcome of a first-order finite element method (FEM) is a globally continuous and piecewise polynomial function u_h. The corresponding flux or stress p_h is

Received by the editors November 1, 2004 and, in revised form, May 1, 2005.
2000 *Mathematics Subject Classification.* 65N30.

usually a linear operator $\mathbb{C}$ evaluated for the gradient Du_h (or its symmetric part) of the finite element function u_h,

$$p_h := \mathbb{C}Du_h \in P_0(\mathcal{T};\mathbb{M}).$$

Here and throughout, $\mathcal{T}$ is a triangulation of the computational domain Ω, $\mathbb{M}$ is a space of vectors or matrices, and $P_k(\mathcal{T};\mathbb{M})$ denotes the piecewise polynomials of degree $\le k$ [piecewise with respect to $\mathcal{T}$ and with values in $\mathbb{M}$].

Typical examples are elliptic partial differential equations of second order in Ω, namely,

$$-\operatorname{div} \mathbb{C}\, Du = f \quad \text{in } \Omega,$$

for the Poisson or Lamé equations, which give rise to a weak formulation

$$a(u, v) = b(v) \quad \text{for all } v \in V.$$

Here and throughout, a is a bounded bilinear form on the Hilbert space V (or on some larger space) and b is a bounded linear functional on V, written $b \in V^*$.

For the ease of this overview, the presentation is restricted to homogeneous Dirichlet conditions on the entire boundary. Then, the flux or stress $p := \mathbb{C}Du$ satisfies no prescribed boundary conditions and can be approximated by some globally continuous and piecewise polynomial functions which form a discrete space

$$\mathcal{Q}_h := P_1(\mathcal{T};\mathbb{M}) \cap C^0(\Omega;\mathbb{M}).$$

Given p_h, the norm $\|\cdot\|$, and the discrete space $\mathcal{Q}_h$, the minimal averaging a posteriori error estimator η_M reads

$$\eta_M := \min_{q_h \in \mathcal{Q}_h} \|p_h - q_h\|.$$

The computation of η_M involves a global minimization which can be solved by an iterative scheme which is not too costly in many applications when a (weighted) L^2 projection is involved. However, local versions appear as accurate as η_M which involves a postprocessing defined by an operator

$$A : Q \to \mathcal{Q}_h \quad \text{for } Q := \{\mathbb{C}Dv : v \in V\} \subset L := L^2(\Omega;\mathbb{M}).$$

Then the A *averaging a posteriori error estimator* η_A reads

$$\eta_A := \|p_h - A(p_h)\|.$$

One particular important example is the ZZ estimator [24]

$$\eta_Z := \|p_h - Z(p_h)\| \approx \eta_{\mathcal{E}},$$

which is equivalent to the jumps across interior element edges $\eta_{\mathcal{E}}$ (for conforming P_1 FEM). Details on the notation follow in Section 2. It is obvious that $\eta_M \le \eta_A$. The surprising converse of which will be shown below for a class of averaging operators. In fact, Section 2 establishes

$$\eta_A \approx \eta_M \approx \eta_Z \approx \eta_{\mathcal{E}}.$$

Here and throughout, the statement $a \lesssim b$ abbreviates $a \le C\, b$ for some positive generic constant C which does not depend on the meshsize in $\mathcal{T}$, and $a \lesssim b \lesssim a$ is abbreviated by $a \approx b$.

This paper discusses examples of problems and estimators and studies their *reliability*, i.e.,

$$\|p - p_h\| \lesssim \eta_M + \text{ h.o.t.},$$

(recall that h.o.t. abbreviates higher-oder terms, the meaning of which is clarified below) and their *efficiency*, i.e.,

$$\eta_M \lesssim \|p - p_h\| + \text{ h.o.t.}$$

The remaining part of this paper is organized as follows. Section 2 introduces the necessary notation to define averaging operators with respect to a general finite element mesh on a general level. Section 3 aims at a proof of efficiency of *all* averaging techniques by a discussion of the various equivalences. Up to this point, there is no need to restrict the application to some specific PDE. Section 4, however, has to focus on a model example of an elliptic boundary value problem and the reliability of η_M and hence reliability of *all* averaging techniques. The concluding Section 5 comments on various generalizations such as history, other averaging spaces, averaging on large patches, other boundary conditions, and numerical experience.

2. Averaging techniques as universal tool for a posteriori error control

This section introduces the precise notation (Subsection 2.1), some discrete spaces (Subsection 2.2), some averaging operators (Subsection 2.3), and associated estimators (Subsection 2.4). Universality is supported by the fact that no boundary value problem is considered in this and the following section: The definitions of the averaging operators and their efficiency are universal for *all problems.*

2.1. Notation and preliminaries on the FEM. The bounded Lipschitz domain $\Omega \subset \mathbb{R}^n$, $n = 1, 2, 3$, with piecewise affine boundary Γ is exactly covered by a triangulation $\mathcal{T}$, $\cup\mathcal{T} = \overline{\Omega}$. Each element $T \in \mathcal{T}$ is a compact interval $T = \text{conv}\{a, b\}$ if $n = 1$, a triangle $T = \text{conv}\{a, b, c\}$ if $n = 2$, or a tetrahedron $T = \text{conv}\{a, b, c, d\}$ if $n = 3$. The element's vertices $a, \ldots, d$ 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 $E = \text{conv}\{a, b\}$, or a face $E = \text{conv}\{a, b, c\}$; $\mathcal{E}$ 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 applies to parallelograms ($n = 2$) or parallelepipeds ($n = 3$) 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 solely for the ease of the presentation. For each node $z \in \mathcal{N}$ let $\mathcal{E}_z := \{E \in \mathcal{E} : z \in E \cap \mathcal{N}\}$ and the patch $\omega_z := \text{int}(\cup\mathcal{T}_z)$, $\mathcal{T}_z := \{T \in \mathcal{T} : z \in T \cap \mathcal{N}\}$. 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 $h_E = \text{diam}(E)$ and $|E| = \mathcal{L}^{n-1}(E)$, respectively; $\mathcal{L}^n$ denotes the n-dimensional Lebesgue measure along any affine subspace of $\mathbb{R}^n$. Similarly the length and volume of $T \in \mathcal{T}$ is denoted by $h_T = \text{diam}(T)$ and $|T| = \mathcal{L}^n(T)$, respectively.

2.2. Discrete spaces. On each element there exists a set of shape functions, namely, $P_k(T) := P_{(k)}(T)$ if T is triangular (or tetrahedral in 3D) and $P_k(T) := Q_{(k)}(T)$ if T is rectangular; $P_{(k)}(T)$ and $Q_{(k)}(T)$ denote algebraic polynomials on $T \subseteq \mathbb{R}^n$ of total and partial degree $\leq k$, respectively. Furthermore, for each $T \in \mathcal{T}$, let $P(T;\mathbb{M})$ be $\mathbb{M}$ valued and satisfy $P_0(T;\mathbb{M}) \subset P(T;\mathbb{M}) \subset P_1(T;\mathbb{M})$. Then, set

$$\begin{aligned}
P_k(\mathcal{T}) &:= \{v_h \in L^\infty(\Omega) : \forall T \in \mathcal{T},\, v_h|_T \in P_k(T)\} \text{ for } k = 0, 1,\\
\mathcal{S}^1(\mathcal{T}) &:= P_1(\mathcal{T}) \cap C^0(\Omega) = \text{span}\{\varphi_z : z \in \mathcal{N}\},\\
\mathcal{P}_h &:= P(\mathcal{T};\mathbb{M}) := \{p_h \in L^\infty(\Omega;\mathbb{M}) : \forall T \in \mathcal{T},\, p_h|_T \in P(T;\mathbb{M})\},\\
\mathcal{Q}_h &:= \mathcal{S}^1(\mathcal{T};\mathbb{M}) = \{q_h \in C^0(\Omega;\mathbb{M}) : \forall T \in \mathcal{T},\, q_h|_T \in P_1(T;\mathbb{M})\}.
\end{aligned}$$

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

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

Given any open subset ω of Ω we define the restricted spaces

$$\mathcal{P}_h|_\omega := \{p_h|_\omega : p_h \in \mathcal{P}_h\} \quad \text{and} \quad \mathcal{Q}_h|_\omega := \{q_h|_\omega : q_h \in \mathcal{Q}_h\}.$$

2.3. Averaging Operators. Given $p_h \in \mathcal{P}_h$ (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 to adapt to boundary conditions. Therefore,

$$A(p_h) := \sum_{z \in \mathcal{N}} A_z(p_h|_{\omega_z})\, \varphi_z \quad \text{and} \quad A_z : P_1(\mathcal{T}_z; \mathbb{M}) \to \mathbb{M}.$$

Recall that $P_{(1)}(\mathcal{T}_z)$ denotes the $\mathcal{T}_z$ piecewise polynomials of degree ≤ 1 and that $p_h|_{\omega_z}$ belongs to $P_{(1)}(\mathcal{T}_z)$. The linear operator

$$A_z : P_1(\mathcal{T}_z; \mathbb{M}) \to \mathbb{M}$$

describes a local averaging process and is (frequently in this paper) assumed to *preserve*

$$N(\omega_z) := (\mathcal{P}_h|_{\omega_z}) \cap (\mathcal{Q}_h|_{\omega_z}).$$

That means, there holds

$$A_z(f) = f(z) \quad \text{for all } f \in N(\omega_z) \quad \text{and all } z \in \mathcal{N}. \tag{2.1}$$

Example 2.1 (Averaging of Nodal Values). Amongst the easiest averaging techniques is the averaging at the node $z \in \mathcal{N}$ where A_z is defined as some mean of all the different values of $(f|_T)(z)$ at $z \in T$ for the different elements $T \in \mathcal{T}_z$ (notice f is, in general, discontinuous at z). That is

$$A_z(f) := \sum_{T \in \mathcal{T}_z} \lambda_{z,T} (f|_T)(z) \quad \text{for all } f \in P_1(\mathcal{T}_z),\ z \in \mathcal{N}. \tag{2.2}$$

A necessary and sufficient condition for (2.1) on the real coefficients $(\lambda_{z,T} : T \in \mathcal{T}_z)$ in (2.2) reads

$$\textstyle\sum_{T \in \mathcal{T}_z} \lambda_{z,T} = 1. \tag{2.3}$$

Example 2.2 (ZZ Averaging). The particular situation of (2.2) with

$$\lambda_{z,T} := |T|/|\omega_z| \quad \text{for all } T \in \mathcal{T},\ z \in \mathcal{N} \tag{2.4}$$

(where $|\cdot|$ denotes the area or volume) is our interpretation of a gradient recovery. For $\mathcal{P}_h = \mathcal{P}_0(\mathcal{T})^d$, this is due to Zienkiewicz and Zhu [24]. The corresponding operator $Z := A$ with $A_z := Z_z$ reads

$$Z(f) = \sum_{z \in \mathcal{N}} \left(\sum_{T \in \mathcal{T}_z} |T|/|\omega_z| (f|_T)(z) \right) \varphi_z.$$

Since the choice (2.4) immediately implies (2.3), whence (2.1).

2.4. Estimators. Given the spaces $\mathcal{P}_h$ and $\mathcal{Q}_h$ of Subsection 2.2 and the averaging operator $A : \mathcal{P}_h \to \mathcal{Q}_h$ of Subsection 2.3, we define, for any fixed $p_h \in \mathcal{P}_h$, the averaging estimators

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

For the ZZ averaging operator from Example 2.2, we define

$$\eta_Z := \|p_h - Z(p_h)\|_{L^2(\Omega)}.$$

Finally, given any $p_h \in \mathcal{P}_h$ and $E \in \mathcal{E}$ let $[p_h]|_E$ denotes the jump of p_h across the edge (if $d = 2$) or face (if $d = 3$) E with the L^2 norm $\|[p_h]|_E\|_{L^2(E)}$ along E

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

Throughout this paper, there are no Neumann boundary conditions on p and hence only interior edges are under consideration by the convention $[p_h]|_E := 0$ if $E \in \mathcal{E}_\Gamma$.

All considered estimators are applicable once a triangulation and the spaces $\mathcal{P}_h$ and $\mathcal{Q}_h$ are known.

3. Efficiency of all averaging schemes and all problems

There is no need to specify the boundary value problem in order to apply the estimators $\eta_M \leq \eta_A$. This section is devoted to prove that *all* averaging estimators are generically efficient for smooth solutions of *any* problem.

3.1. Efficiency of η_M for smooth exact solution. Given the spaces $\mathcal{P}_h$ and $\mathcal{Q}_h$ of Subsection 2.2 and η_M of Subsection 2.4, there holds efficiency in the sense of

$$\eta_M \leq \|p - p_h\| + \text{ h.o.t.} \tag{3.1}$$

The surprising fact is that p can be *any* smooth function, e.g. $p \in H^1(\Omega; \mathbb{R}^n)$, and $\|\cdot\|$ can be *any* norm such that the approximation properties of $\mathcal{P}_h$ and $\mathcal{Q}_h$ justify the notation of

$$\text{h.o.t.} := \min_{q_h \in \mathcal{Q}_h} \|p - q_h\| \tag{3.2}$$

as *higher-order terms* when compared with $\|p - p_h\|$. The proof of (3.1) is via a triangle inequality: For any $q_h \in \mathcal{Q}_h$ there holds

$$\eta_M \leq \|p_h - q_h\| \leq \|p - p_h\| + \|p - q_h\|.$$

Thus efficiency (3.1) holds without any reference to the underling boundary value problem. For instance, if p_h denotes some flux approximation in a first-order conforming or nonconforming or lowest-order mixed FEM, then the L^2 error $\|p - p_h\|$ is of first order, while (3.2) is of second order provided the exact flux p is sufficiently smooth.

For PDEs in divergence form and conforming first-order finite element methods, this efficiency result can be improved such that the higher-order terms are explicitly known as data oscillations. We refer to Subsection 4.5 below for a discussion of this point.

In this paper, the efficiency of η_A is reduced to that of η_M.

3.2. Equivalence $\eta_M \approx \eta_A$. The above very general description requires an extra condition for the claimed equivalence. For each node $z \in \mathcal{N}$ we suppose that the averaging operator A_z is exact on $N(\omega_z)$.

Theorem 3.1. *Suppose that $A : \mathcal{P}_h \to \mathcal{Q}_h$ satisfies, for any $z \in \mathcal{N}$, that*

$$A_z(f) = f(z) \quad \text{for all } f \in N(\omega_z). \tag{3.3}$$

Then there exists a mesh-size independent positive constant C_{eff} with

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

Remark 3.1. A detailed analysis for P_1 finite element methods on simplices in $\mathbb{R}^n$ with stiffness matrices, strengthened Cauchy inequalities, and some adaptation of Ascoli's lemma in [8] shows that the constant C_{eff} does not depend on the shape of the geometry. In fact,

$$C_{\text{eff}} \le \sqrt{10} \text{ for } d = 2 \quad \text{and} \quad C_{\text{eff}} \le \sqrt{15} \text{ for } d = 3.$$

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. In particular, the coefficients $p_{T,z}$ of a fixed $p_h|_T$ may depend on T (and are possibly different for different elements). Set

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

For the fixed T with set of vertices $\mathcal{N}(T) := \mathcal{N} \cap T$ let

$$\omega_T = \cup_{z \in \mathcal{N}(T)} \omega_z$$

denote the interior of the union $\cup_{z \in \mathcal{N}(T)} \mathcal{T}_z$ of all triangles which share at least some vertex with T. The expressions

$$\varrho_1(p_h) := \|p_h - Ap_h\|_{L^2(T)} \text{ and } \varrho_2(p_h) := \min_{r_h \in \mathcal{Q}_h} \|p_h - r_h\|_{L^2(\omega_T)}$$

define two seminorms on the finite-dimensional vector space

$$\mathcal{P}_h|_{\omega_T} := \{p_h|_{\omega_T} : p_h \in \mathcal{P}_h\}.$$

Suppose that $p_h \in \mathcal{P}_h$ satisfies $\varrho_2(p_h|_{\omega_T}) = 0$. Then, $p_h|_{\omega_T} = r_h|_{\omega_T}$ for some $r_h \in \mathcal{Q}_h$. In particular, for any $z \in \mathcal{N}(T)$, $f := p_h|_{\omega_z} \in N(\omega_z)$ is a proper test function in (3.3) and hence (notice p_h is continuous in ω_z) there holds $A_z(p_h|_{\omega_z}) = p_h(z)$. Since this holds for any $z \in \mathcal{N}(T)$, $p_h|_T = q_h|_T$, and so $\varrho_1(p_h|_T) = 0$. In conclusion, we have

$$\varrho_2^{-1}(\{0\}) \subseteq \varrho_1^{-1}(\{0\}).$$

As a further consequence of arguments for the equivalence of norms on finite-dimensional vector spaces, is follows that

$$\varrho_1(p_h|_{\omega_T}) \le C_T\, \varrho_2(p_h|_{\omega_T}) \qquad \text{for all } p_h \in \mathcal{P}_h.$$

A scaling argument proves that the constant $C_T \lesssim 1$ does not depend on the mesh-size (but, at this point, possibly depends on the shape) of the elements in ω_T (e.g. through their interior angles or aspect ratios).

Given any $r_h \in \mathcal{Q}_h$, there holds

$$\varrho_2(p_h|_{\omega_T}) \le \|p_h - r_h\|_{L^2(\omega_T)}.$$

Hence, the sum over all elements $T \in \mathcal{T}$ in the preceding two estimates and the fact that the covering $(\omega_T : T \in \mathcal{T})$ of Ω has a finite overlap $C_{\mathcal{T}}$ lead to

$$\|p_h - q_h\|_{L^2(\Omega)}^2 \le (\max_{K \in \mathcal{T}} C_K^2) \sum_{T \in \mathcal{T}} \|p_h - r_h\|_{L^2(\omega_T)}^2 \le C_{\text{eff}}^2 \|p_h - r_h\|_{L^2(\Omega)}^2$$

with $C_{\text{eff}} := C_{\mathcal{T}}^{1/2} \max_{K\in\mathcal{T}} C_K$ independent of the mesh-size. Since r_h is arbitrary, this concludes the proof. $\square$

3.3. Equivalence $\eta_M \approx \eta_{\mathcal{E}}$. This equivalence holds without further assumptions.

Theorem 3.2. *There holds*

$$\eta_{\mathcal{E}} \lesssim \eta_M.$$

Proof. The main argument of the proof is a local equivalence of norms. Given any node $z \in \mathcal{N}$, consider the finite-dimensional vector space $\mathcal{P}_h|_{\omega_z}$ and the two seminorms $\varrho_3, \varrho_4 : \mathcal{P}_h|_{\omega_z} \to \mathbb{R}$ defined, for any $p_h \in \mathcal{P}_h|_{\omega_z}$, by

$$\varrho_3(p_h):=\Big(\sum_{E\in\mathcal{E}_z} h_E\|[p_h]|_E\|^2_{L^2(E)}\Big)^{1/2} \text{ and } \varrho_4(p_h):=\min_{r_h\in\mathcal{Q}_h}\|p_h - r_h\|_{L^2(\omega_z)}.$$

Notice that the kernels $\varrho_j^{-1}(\{0\}) = \{f \in \mathcal{P}_h|_{\omega_z} : \varrho_j(f) = 0\}$ are equal for $j = 3, 4$,

$$\varrho_3^{-1}(\{0\}) = (\mathcal{P}_h|_{\omega_z}) \cap C(\omega_z) = (\mathcal{P}_h|_{\omega_z}) \cap (\mathcal{Q}_h|_{\omega_z}) = \varrho_4^{-1}(\{0\}).$$

A scaling argument shows that the constants in the equivalence

$$\varrho_3 \approx \varrho_4 \quad \text{on } \mathcal{P}_h|_{\omega_z}$$

do not depend on the size of the elements in $\mathcal{T}_z$ (but may depend on their shapes, e.g., through their interior angles or aspect ratios).

The sum of these estimates for all $z \in \mathcal{N}$ yields

$$\eta_{\mathcal{E}}^2 \approx \sum_{z\in\mathcal{N}} \min_{r_h\in\mathcal{Q}_h|_{\omega_z}} \|p_h - r_h\|^2_{L^2(\omega_z)}. \tag{3.4}$$

Since the covering $(\omega_z : z \in \mathcal{N})$ of Ω has a finite overlap, the preceding equivalence leads to $\eta_{\mathcal{E}} \lesssim \eta_M$. $\square$

The subsequent result implies the remaining estimate $\eta_M \lesssim \eta_{\mathcal{E}}$.

Theorem 3.3. *There holds*

$$\eta_Z \lesssim \eta_{\mathcal{E}}.$$

Proof. Adopt notation of the proof of Theorem 3.1 for $A = Z$ and one fixed T. Since

$$\|p_h - Z(p_h)\|^2_{L^2(T)} \lesssim |T|^{1/2} \max_{z\in\mathcal{N}(T)} |p_h(z) - Z_z(p_h)|,$$

one studies the seminorm ϱ_5 defined for each node $z \in \mathcal{N}$ by

$$\varrho_5(p_h) := |\omega_z|^{1/2}\, |p_h|_T(z) - Z_z(p_h)| \quad \text{for all } p_h \in \mathcal{P}_h|_{\omega_z}.$$

The proof of Theorem 3.2 involved ϱ_4 on $\mathcal{P}_h|_{\omega_z}$. The above arguments imply $\varrho_4^{-1}(\{0\}) \subseteq \varrho_5^{-1}(\{0\})$ and, by scaling, eventually lead to $\varrho_5 \lesssim \varrho_4$. The combination with the first estimate of this proof leads to

$$\|p_h - Z(p_h)\|^2_{L^2(T)} \lesssim \sum_{z\in\mathcal{N}(T)} \min_{r_h\in\mathcal{Q}_h|_{\omega_z}} \|p_h - r_h\|^2_{L^2(\omega_z)}.$$

The sum over all elements T, the finite overlap of all patches ω_z, and the already established estimate (3.4) from the end of the proof of Theorem 3.2 verifies

$$\|p_h - Z(p_h)\|^2_{L^2(\Omega)} \lesssim \sum_{z\in\mathcal{N}} \min_{r_h\in\mathcal{Q}_h|_{\omega_z}} \|p_h - r_h\|^2_{L^2(\omega_z)} \approx \eta_{\mathcal{E}}^2.$$

$\square$

4. A Posteriori Error Analysis for PDE in Divergence Form

This section is devoted to some reliability estimates for equations of divergence form. This is perhaps the most general and abstract form that allows the precise formulation of reliability and efficiency in an a posteriori error analysis.

4.1. Model Problem in Divergence Form. Given a right-hand side $f \in L^2(\Omega; \mathbb{R}^m)$, suppose that the exact flux $p \in L := L^2(\Omega; \mathbb{M})$ satisfies

$$f + \operatorname{div} p = 0 \quad \text{in } \Omega.$$

The weak form is satisfied by the discrete flux $p_h \in \mathcal{P}_h$ in the sense of

$$\int_\Omega p_h : Dv_h \, dx = \int_\Omega f \cdot v_h \, dx \qquad \text{for all } v_h \in V_h \subset V := H^1_0(\Omega; \mathbb{R}^m).$$

Here and throughout this section, $\mathcal{S}^1_0(\mathcal{T}) \subseteq V_h$, i.e., V_h is supposed to include

$$\mathcal{S}^1_0(\mathcal{T}) := \{v_h \in \mathcal{S}^1(\mathcal{T}) : v_h = 0 \text{ on } \partial\Omega\} = \mathcal{S}^1(\mathcal{T}) \cap V \subseteq V_h.$$

The first-order approximation property and the H^1 stability of a Clément type approximation operator $J : V \to V_h$ (also called quasi interpolation operator)

$$\|v - J(v)\|^2_V + \sum_{T \in \mathcal{T}} h_T^{-2} \|v - J(v)\|^2_{L^2(T)} + \sum_{E \in \mathcal{E}} h_E^{-1} \|v - J(v)\|^2_{L^2(E)} \lesssim \|v\|^2_V$$

are well-established. The paper [13] studies a simple example and gives explicit constants in this estimate.

It is understood that p exists but is not known, while the data f and Ω as well as the triangulation $\mathcal{T}$ and the discrete flux $p_h \in \mathcal{P}_h$ are given.

4.2. Data Oscillations. Given the right-hand side $f \in L^2(\Omega; \mathbb{R}^m)$ and the triangulation $\mathcal{T}$, let $f_\mathcal{T} \in P_0(\mathcal{T}; \mathbb{R}^m)$ denote the piecewise integral means defined by

$$f_\mathcal{T}|_T := f_T := |T|^{-1} \int_T f(x) \, dx \quad \text{for all } T \in \mathcal{T}.$$

Let $h_\mathcal{T} \in P_0(\mathcal{T})$ denote the piecewise constant mesh-size defined by $h_\mathcal{T}|_T := h_T := \operatorname{diam}(T)$ for all $T \in \mathcal{T}$. Then

$$\operatorname{osc}(f; \mathcal{T}) := \|h_\mathcal{T}(f - f_\mathcal{T})\|_{L^2(\Omega)} = \left(\sum_{T \in \mathcal{T}} h_T^2 \|f - f_T\|^2_{L^2(T)} \right)^{1/2}$$

denotes the *elementwise data oscillations*. Given any patch ω_z set

$$f_{\omega_z} := |\omega_z|^{-1} \int_{\omega_z} f(x) \, dx.$$

Then

$$\operatorname{osc}(f; \mathcal{N}) := \left(\sum_{z \in \mathcal{N}} \operatorname{diam}(\omega_z)^2 \|f - f_{\omega_z}\|^2_{L^2(\omega_z)} \right)^{1/2}$$

denotes the *patchwise data oscillations*.

Notice that a Poincaré inequality for smooth data f proves that the data oscillations are of quadratic order as the mesh-size tends to zero in the sense that

$$\operatorname{osc}(f; \mathcal{N}) + \operatorname{osc}(f; \mathcal{T}) \lesssim \|h_\mathcal{T}^2 Df\|_{L^2(\Omega)} \quad \text{in case } f \in H^1(\Omega; \mathbb{R}^m).$$

The subsequent theorem is essentially known and recalled here with a sketch of the proof for the reader's convenience. The proof gives a local form of the assertion which is not displayed for brevity.

Recall that $[p_h]$ is the jump of p_h accross E and ν_E is the unit normal vector for $E \in \mathcal{E}$, and define $\operatorname{div}_h p_h$ as the elementwise divergence of p_h.

Theorem 4.1. *There holds*

$$\|h_\mathcal{T}(f + \operatorname{div}_h p_h)\|_{L^2(\Omega)} \lesssim \left(\sum_{E\in\mathcal{E}} h_E \|[p_h]\nu_E\|^2_{L^2(E)}\right)^{1/2} + \operatorname{osc}(f + \operatorname{div}_h p_h; \mathcal{N}).$$

Proof. Let $\in \mathcal{N}(\Omega)$ denote an interior node and abbreviate $h_z := \operatorname{diam}(\omega_z)$ and $f_z := f_{\omega_z} \in \mathbb{R}^m$. Write $g := f + \operatorname{div}_h p_h$ and define $g_z := g_{\omega_z} \in \mathbb{R}^m$. Since $g_z\,\varphi_z \in V_h$, the discrete equilibrium condition yields

$$\int_\Omega f \cdot g_z\,\varphi_z\,dx = \int_\Omega p_h : D(g_z\varphi_z)\,dx.$$

An elementwise integration by parts shows that this equals

$$\sum_{E\in\mathcal{E}_z} \int_E (g_z\varphi_z)\cdot[p_h]\nu_E\,ds - \int_\Omega \varphi_z\,g_z \cdot \operatorname{div}_h p_h\,dx.$$

Since $g = f + \operatorname{div}_h p_h$, the above two identities combine to

$$\int_\Omega g \cdot g_z\,\varphi_z\,dx = \sum_{E\in\mathcal{E}_z} \int_E (g_z\varphi_z)\cdot[p_h]\nu_E\,ds.$$

Up to multiplicative constants, this is bounded from above by

$$\|g_z\|_{L^2(\omega_z)} \left(\sum_{E\in\mathcal{E}_z} h_E^{-1}\|[p_h]\nu_E\|^2_{L^2(E)}\right)^{1/2}.$$

On the other hand, with some elementary considerations,

$$\begin{aligned}\|g\|^2_{L^2(\omega_z)} &= \|g - g_z\|^2_{L^2(\omega_z)} + \|g_z\|^2_{L^2(\omega_z)}\\ &\lesssim \int_\Omega g\cdot g_z\,\varphi_z\,dx + \|g - g_z\|^2_{L^2(\omega_z)}.\end{aligned}$$

The final estimate and the first three identities displayed in this proof imply

$$h_z^2\|g\|^2_{L^2(\omega_z)} \lesssim \sum_{E\in\mathcal{E}_z} h_E\|[p_h]\nu_E\|^2_{L^2(E)} + \operatorname{osc}(g;\mathcal{N}_z)^2.$$

□

4.3. Explicit Residual-Based A Posteriori Error Estimates. This subsection introduces an abstract error norm $\|D^*(p - p_h)\|_{V^*}$ and adopts a reliable explicit residual-based error estimator η_R from the literature.

The first derivative $D : V \to L$ and its dual $D^* : V^* \to L$ define an error norm $\|D^*(p - p_h)\|_{V^*}$, that is

$$\|D^*(p - p_h)\|_{V^*} = \sup_{v\in V\setminus\{0\}} \frac{\int_\Omega (p - p_h) : Dv\,dx}{\|v\|_V}.$$

Theorem 4.2. *There holds*

$$\|D^*(p - p_h)\|_{V^*} \lesssim \|h_\mathcal{T}(f + \operatorname{div}_h p_h)\|_{L^2(\Omega)} + \left(\sum_{E\in\mathcal{E}} h_E\|[p_h]\nu_E\|^2_{L^2(E)}\right)^{1/2}.$$

Proof. The proof is standard in a slightly different context and hence we give only a sketch of it. With the Clément type approximation operator J from Subsection 4.1 it follows for any $v \in V$ and $w := v - J(v) \in V$ that

$$\int_\Omega (p - p_h) : Dv\,dx = \int_\Omega (p - p_h) : Dw\,dx = \int_\Omega f \cdot w\,dx - \int_\Omega p_h : Dw\,dx.$$

An elementwise integration by parts shows

$$-\int_\Omega p_h : Dw\,dx = \int_\Omega w \cdot \operatorname{div}_h p_h\,dx - \sum_{E\in\mathcal{E}} \int_E w \cdot [p_h]\nu_E\,ds.$$

The combination of the foregoing two identities with the first-order approximation properties and the H^1 stability leads to an upper bound of $\int_\Omega (p-p_h) : Dv\,dx/\|v\|_V$ and so to the assertion. This is standard [23] and we omit further details. □

The combination of Theorem 4.1 and 4.2 immediately implies the following refined version of Theorem 4.2.

Theorem 4.3. *There holds*

$$\|D^*(p - p_h)\|_{V^*} \lesssim \left(\sum_{E\in\mathcal{E}} h_E \|[p_h]\nu_E\|^2_{L^2(E)}\right)^{1/2} + \operatorname{osc}(f + \operatorname{div}_h p_h; \mathcal{N}).$$

□

4.4. Reliability of all Averaging Estimators up to Data Oscillations. For $p_h \in P_0(\mathcal{T};\mathbb{M})$ and $f \in H^1(\Omega;\mathbb{R}^m)$, $\operatorname{osc}(f + \operatorname{div}_h p_h; \mathcal{N}) \lesssim \|h_\mathcal{T}^2 Df\|_{L^2(\Omega)}$ is of higher order and so the normal components of the edge contributions in Theorem 4.3 are reliable up to higher-order terms.

Theorem 4.4. *Suppose $p_h \in P_0(\mathcal{T};\mathbb{M})$ and $f \in H^1(\Omega;\mathbb{R}^m)$. Then there holds*

$$\|D^*(p - p_h)\|_{V^*} \lesssim \left(\sum_{E\in\mathcal{E}} h_E \|[p_h]\nu_E\|^2_{L^2(E)}\right)^{1/2} + \|h_\mathcal{T}^2 Df\|_{L^2(\Omega)}.$$

□

Since the edge contributions are bounded from above by $\eta_\mathcal{E} \approx \eta_M \approx \eta_A \approx \eta_Z$, this result of dominating edge-contributions reads the *reliability of all averaging techniques* in the sense of

$$C_1\|D^*(p - p_h)\|_{V^*} - C_2\|h_\mathcal{T}^2 Df\|_{L^2(\Omega)} \le \eta_M \le \eta_A \approx \eta_Z.$$

The preceding scenario allows complete error control for conforming first-order finite element methods with $p_h = \mathbb{C}Du_h$ and $u_h \in \mathcal{S}^1(\mathcal{T};\mathbb{R}^m)$. In particular, $|[p_h]| = |[p_h]\nu_E|$ because the tangential components of jumps vanish (because of the continuity of u_h along the edge provided $\mathbb{C}$ is continuous there).

This is untrue for nonconforming first-order finite element methods with $p_h = \mathbb{C}D_h u_h$ where $u_h \in P_1(\mathcal{T};\mathbb{R}^m)$ is continuous at the midpoint of edges (and zero at midpoints of boundary edges). However, a Helmholtz decomposition of $p - p_h$ shows that we need to estimate the divergence $-\operatorname{div} = D^*$ and the curl of $p - p_h$. For $n = 2$ space dimensions, the curl is just a rotated version of the divergence and so the above result apply. Therein, the remaining crucial point is the proof of $\int_\Omega (D_h u_h) : (\operatorname{Curl} v_h)\,dx = 0$ for any (conforming) $v_h \in \mathcal{S}^1_0(\mathcal{T};\mathbb{R}^m)$. This follows from the continuity at midpoints by an elementwise integration by parts and the fact $\int_E [u_h] \cdot (\partial u_h/\partial s)\,ds = 0$. Details about those arguments can be found in [7, 11] and for three space dimensions in [12].

As the final outcome for first-order conforming or nonconforming finite element methods, there is *reliability of all averaging estimators* in the sense of

$$C_1\|p-p_h\|_L - C_2\|h_{\mathcal{T}}^2 Df\|_{L^2(\Omega)} \le \eta_M \le \eta_A \approx \eta_Z.$$

4.5. Efficiency of Averaging Estimators up to Data Oscillations. The edge contributions in Theorem 4.1 and 4.2 allow for a lower bound as well.

Theorem 4.5. *There holds*

$$\Big(\sum_{E\in\mathcal{E}} h_E\|[p_h]\nu_E\|_{L^2(E)}^2\Big)^{1/2} \lesssim \|D^*(p-p_h)\|_{V^*} + \operatorname{osc}(f;\mathcal{T}).$$

Proof. This is a modification of Verfürth's inverse estimation technique [23] and hence we present only a sketch here. Since $[p_h]\nu_E$ is affine along the fixed edge $E\in\mathcal{E}$, it is recast as the sum of two or three (in 2D or 3D) hat functions attached to the vertices of E. This function w is uniquely defined. Let b_E denote the edge-bubble function which is the product of all (two or three) aforementioned nodal basis functions which are nonzero along E. This defines $wb_E \in H_0^1(\omega_E;\mathbb{R}^m)$ with

$$h_E\|[p_h]\nu_E\|_{L^2(E)}^2 \le h_E\int_E wb_E\cdot[p_h]\nu_E\,ds = h_E\int_E v\cdot[p_h]\nu_E\,ds.$$

Herein, the function v is defined by subtraction of $\alpha_\pm b_{T_\pm}$ from w, $b_{T_\pm}$ denotes the element bubble function (defined as the product of all nodal basis function whose support includes the neighboring element $T_\pm$ of E). The coefficient $\alpha_\pm$ is chosen such that $v\in H_0^1(\omega_E;\mathbb{R}^m)$ has piecewise integral mean zero. The elementwise integration by parts of $h_E\int_E v\cdot[p_h]\nu_E\,ds$ then leads to the assertion. This is standard [23] and we omit further details. □

For first-order conforming finite element methods, Theorem 4.5 reads

$$\eta_M \le \eta_Z \approx \eta_A \approx \eta_{\mathcal{E}} \lesssim \|p-p_h\|_L + \operatorname{osc}(f;\mathcal{T}).$$

Compared to Subsection 3.1 the higher-order terms are now data oscillations $\operatorname{osc}(f;\mathcal{T})$ and do depend on the smoothness of the given right-hand side but do not depend on the smoothness of the unknown exact solution.

5. Concluding remarks

5.1. Brief Remarks on the History of Averaging Estimators. The origin of smoothening postprocessing steps dates back to the graphical representation of piecewise constant stress approximations in computational mechanics. Since a coarse piecewise stress function plot simply looked too discontinues, and the graphic programs easily allowed continuous piecewise linear approximations, it became an immediate issue to average at nodal points. The averaging operator A of this paper is precisely of that form.

The operator η_Z was then suggested by engineers amongst many other strategies for an heuristic gradient recovery [24]. However, the treatment of boundary conditions and insight in the foundation of this estimator was lacking over years. There were even attempts to discredit this estimator by laborious numerical experiments.

Superconvergene phenomena for highly symmetric meshes and very smooth functions are available far away from the boundary. For instance in the context of first-order finite element methods, the difference $I_h u - u_h$ can be much smaller than the error $u-u_h$. If so, the averaged solution on symmetric patches is very easily seen to yield an asymptotically exact estimator η_Z.

A first mathematical proof for that dates back to Rodriguez [21]. He also essentially proved Theorem 4.2 in [22] and so set the mathematical foundation for the reliability of η_Z. He also noticed (a variant of) $\eta_Z \approx \eta_{\mathcal{E}}$ with a different proof [23].

The reliability of η_Z on unstructured grids has been indicated in the literature [21, 22, 20, 6] but was not mentioned in the (otherwise comprehensive) works [1, 2, 17, 23]. The author was unaware of Rodriguez's result [22] when he started to work on the mathematical justification [16] that ended in the conclusion that all averaging techniques are reliable [11]. This is much of a theoretical evidence in support of the mentioned numerical experiments.

The fact that all averaging estimators are efficient is a newer result [8, 9] which is studied in this paper with new proofs.

5.2. Applications. The discussion in this paper is very general and outlines the relatively simple arguments. In many ways, it generalizes and complements the overview [9] with a series of explicit examples. Hence, in this paper, we can mention that reliable and efficient averaging techniques are described for general boundary value problems, the Poisson, Stokes and Lamé equations, treated with conforming, nonconforming, and even some mixed finite element methods in [11, 12, 14, 15].

Surprisingly, averaging techniques are not restricted to partial differential equations. In fact, the same techniques apply to variational inequalities as well as the affirmative results in [5, 10] on elastoplastic and obstacle problems.

Averaging for higher-order finite element methods is possible for a local variant established in [4].

Whenever one has some residual-based error estimator for a problem in divergence form, one can see through Theorem 4.1 that the edge-contributions dominate and then deduce reliability and efficiency as in Subsection 4.4-4.5.

Another averaging technique for unstructured grids is the use of averages over larger patches, cf., e.g., [19, 9], known to the experts since the eighties. For a brief sketch, suppose that we are given two meshes $\mathcal{T}_H$ and $\mathcal{T}_h$ such that (for the ease of this brief illustration) $\mathcal{T}_h$ is a uniform refinement of $\mathcal{T}_H$ such that typical mesh-sizes H and h of $\mathcal{T}_H$ and $\mathcal{T}_h$, respectively, satisfy

$$H^2 \ll h \ll H.$$

Suppose the Poisson problem has the exact solution u and discrete (fine) solution u_h. Let $\mathcal{S}^2(\mathcal{T}_H)$ denote the $\mathcal{T}_H$-piecewise quadratic polynomial subspace of $C(\Omega)$. Let u_H and u_{hH} denote the finite element approximations in $\mathcal{S}^2(\mathcal{T}_H)$ of u and u_h, respectively. Then, with h.o.t. which depend crucially on the higher smoothness of u,

$$\eta := \|Du_h - Du_{hH}\| \quad \text{is computable}$$

and (cf., e.g., [9] for the simple proof) is always efficient and reliable for small h/H in the sense that

$$\eta - \text{ h.o.t. } \leq \|p - p_h\|_L \leq \frac{\eta + \text{ h.o.t.}}{1 - C_3\, h/H}.$$

Throughout this paper, the conditions for reliability are smoothness of coefficients and right-hand sides but not necessarily of the Lipschitz domain or the (unknown) exact solution. At the moment, time-depending problems and problems with accumulated errors or with pollution are excluded. Moreover, the robustness of averaging estimators with respect to crucial parameters jumping or oscillating coefficients requires particular attention (cf. the end of Subsection 5.4 on [18] for the latter point).

5.3. Mixed Boundary Conditions. The formulae of this paper need to be modified in case of Neumann or traction boundary conditions. It is the aim of this subsection to describe the necessary modifications in a general setting [8] (e.g. for the Poisson problem) with $m = 1$ and

$$u = u_D \text{ on } \Gamma_D \quad \text{and} \quad p\nu = g \text{ on } \Gamma_N.$$

Therein, the boundary $\Gamma = \cup\mathcal{E}_\Gamma$ 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 $\overline{\Gamma}_N$. That is, the two disjoint subsets $\mathcal{E}_D$ and $\mathcal{E}_N$ of $\mathcal{E}_\Gamma$ are supposed to satisfy $\mathcal{E}_D = \emptyset$ or $\mathcal{E}_D = \{E \in \mathcal{E}_\Gamma : E \subset \Gamma_D\}$ as well as $\mathcal{E}_N = \emptyset$ or $\mathcal{E}_N = \{E \in \mathcal{E}_\Gamma : E \subset \overline{\Gamma_N}\}$.

Given $\mathcal{E}_D$ and $\mathcal{E}_N$, the boundary data $g = p\nu \in L^2(\Gamma_N)$ for the traction and $u_D \in H^{1/2}(\Gamma_D) \cap C(\Gamma_D)$ for the displacements are supposed to 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 $j = 1, \dots, n-1$ such that $(\nu_E, \tau_E^{(1)}, \dots, \tau_E^{(n-1)})$ is a Cartesian basis of $\mathbb{R}^n$. Then, $\nabla_E u_D$ denotes the tangential derivative and, given $a \in \mathbb{R}^n$, $(a)_E$ denotes the vector of all components of a in $(\tau_E^{(j)})_{j=1}^{n-1}$, e.g. $(a)_E = (\tau_E^{(1)} \cdot a, \tau_E^{(2)} \cdot a)$ for $n = 3$; $\nabla_E u_D = (\nabla u_D)_E = \partial u_D / \partial s$ for $n = 2$.

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

$$\mathcal{A}_z := \{a \in \mathbb{R}^n : \forall E \in \mathcal{E}_z \cap \mathcal{E}_N,\, g(z) = a \cdot \nu_E \text{ and } \forall E \in \mathcal{E}_z \cap \mathcal{E}_D, \nabla_E u_D(z) = (a)_E\} \tag{5.1}$$

of $\mathbb{R}^n$. Set $\mathcal{A}_z = \mathbb{R}^n$ for $z \in \mathcal{N} \cap \Omega$ and suppose $\mathcal{A}_z \neq \emptyset$ for all $z \in \mathcal{N}$. Finally, let $\pi_z : \mathbb{R}^n \to \mathbb{R}^n$ 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 $\mathbb{R}^n$. The (non-linear) orthogonal projection π_z is Lipschitz continuous with $\mathrm{Lip}(\pi_z) \le 1$ and, for each $a \in \mathbb{R}^n$, $a - \pi_z(a) \perp \mathcal{V}_z$.

The suggested averaging operator $\mathcal{A}$ then is the composition of local averaging followed by projection, i.e., for any $f \in P_1(\mathcal{T}; \mathbb{M})$,

$$\mathcal{A}(f) := \sum_{z \in \mathcal{N}} \pi_z(A_z(f|_{\omega_z}))\, \varphi_z.$$

5.4. Numerical Examples. Many numerical examples give evidence of an amazing accuracy of the ZZ averaging scheme in the context of piecewise constant fluxes and this is well-documented in many papers including [10, 11, 4, 14, 15]. Mesh perturbations in some of these papers indicate that a local symmetry is important for that. An explanation is that local superconvergence phenomena are responsible for the high precision of η_Z.

The work [11] also displays numerical results for mixed finite element methods and the estimator η_M and η_A. Their reliability is proved therein, but it is emphasized that the efficiency of η_A remains unclear because A computes the nodewise average of the integral means (and not of their nodal values) and hence may violate (2.1).

The averaging space $\mathcal{Q}_h$ could be modified to some finite element space with

$$\mathcal{Q}_h \subset P_1(\mathcal{T}; \mathbb{M}) \cap H(\mathrm{div}, \Omega; \mathbb{M}).$$

The latter set $H(\mathrm{div}, \Omega; \mathbb{M}) \subset H^1(\Omega; \mathbb{M})$ is the vector space of all $\mathbb{M}$-valued Lebesgue functions $Q \in L^2(\Omega; \mathbb{M})$ with a weak divergence in $L^2(\Omega; \mathbb{R}^m)$. For instance, mixed FEM could be employed for $\mathcal{Q}_h$. We refer to [18] for details and results in situations in which jumping coefficients suggest to relax the strict continuity conditions in the averaging functions.

References

[1] Ainsworth, M. and Oden, J.T., A posteriori error estimation in finite element analysis, Wiley-Interscience [John Wiley & Sons], New York. xx+240, 2000.

[2] Babuška, I. and Strouboulis, T., The finite element method and its reliability, The Clarendon Press Oxford University Press, New York, xii+802, 2001.

[3] Bank, R.E. and Xu, J., Asymptotically exact a posteriori error estimators. I: Grids with superconvergence.II: General unstructured grids, SIAM J. Numer. Anal., **41**, 6, 2294–2332, 2003.

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

[5] Bartels, S. and Carstensen, C., Averaging techniques yield reliable a posteriori finite element error control for abstacle problems, Numer. Math., **99**, 2, 225–249, 2004.

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

[7] Carstensen, C., Quasi-interpolation and a posteriori error analysis in finite element method, M2AN Math. Model. Numer. Anal., **33**, 6, 1187–1202, 1999.

[8] Carstensen, C., All first-order averaging techniques for a posteriori finite element error control on unstructured grids are efficient and reliable, Math. Comp., **73**, 247, 1153–1165, 2004.

[9] Carstensen, C., Some remarks on the history and future of averaging techniques in a posteriori finite element error analysis, ZAMM Z. Angew. Math. Mech., **84**, 1, 3–21, 2004.

[10] Carstensen, C. and Alberty, J., Averaging techniques for reliable a posteriori FE-error control in elastoplasticity with hardening, Comput. Methods Appl. Mech. Engrg., **192**, 11-12, 1435–1450, 2003.

[11] Carstensen, C. and Bartels, S., Each averaging technique yields reliable a posteriori error control in FEM on unstructured grids. I. Low order conforming, nonconforming and Mixed FEM, Math. Comp., **71**, 239, 945–969, 2002.

[12] Carstensen, C., Bartels, S. and Jansche, S., A posteriori error estimates for nonconforming finite element methods, Numer. Math., **92**, 2, 233–256, 2002.

[13] Carstensen, C. and Funken, S.A., Constants in Clément-interpolation error and residual based a posteriori estimates in finite element methods, East-West J. Numer. Math., **8**, 3, 153–175, 2000.

[14] Carstensen, C. and Funken, S.A., Averaging technique for FE a posteriori error control in elasticity.
I. Conforming FEM, Comput. Methods Appl. Mech. Engrg., **190**, 18-19, 2483–2498.
II. λ-independend estimates, Comput. Methods Appl. Mech. Engrg., **190**, 35–36, 4663–4675.
III. Locking-free nonconforming FEM, Comput. Methods Appl. Mech. Engrg., **191**, 8–10, 861–877, 2001.

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

[16] 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.

[17] Eriksson, K., Estep, D., Hansbo, P. and Johnson, C., Introduction to adaptive methods for differential equations, Acta Numerica, 1995, Cambridge Univ. Press, Cambridge, 105–158, 1995.

[18] Funken, S.A., Beiträge zur aposteriori Fehlerabschaetzung bei der numerischen Behandlung elliptischer partieller Differentialgleichungen -Theorie, Numerik und Anwendungen-, [Contributions to a posteriori error estimation in the numerical treatment of elliptic PDEs -theory, numerics, and applications-] Habilitation Thesis 2002 in Kiel, in German, University of Kiel, Germany, 2002.

[19] Hoffmann, W. and Schatz, A. H. and Wahlbin, L. B. and Wittum, G., Asymptotically exact a posteriori estimators for the pointwise gradient error on each element in irregular meshes.

I. A smooth problem and globally quasi-uniform meshes, Math. Comp., **70**, 235, 897–909, 2001.

[20] Nochetto, R.H., Removing the saturation assumption in a posteriori error analysis, Istit. Lombardo Accad. Sci. Lett. Rend. A, **127**, 1, 67-82 (1994), 1993.

[21] Rodriguez, R., Some remarks on Zienkiewicz-Zhu estimator, Numer. Methods Partial Differential Equations, **10**, 5, 625–635, 1994.

[22] Rodríguez, R., A posteriori error analysis in the finite element method, (Finite element methods (Jyväskylä, 1993)). Lecture Notes in Pure and Appl. Math., Dekker, New York, **164**, 389–397, 1994.

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

[24] Zienkiewicz, O.C. and Zhu, J.Z., A simple error estimator and adaptive procedure for practical engineering analysis, Internat. J. Numer. Methods Engrg., **24**, 2, 337–357, 1987.

Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany
E-mail: cc@math.hu-berlin.de