

**EACH AVERAGING TECHNIQUE
YIELDS RELIABLE A POSTERIORI ERROR CONTROL
IN FEM ON UNSTRUCTURED GRIDS.
PART I: LOW ORDER CONFORMING, NONCONFORMING,
AND MIXED FEM**

CARSTEN CARSTENSEN AND SÖREN BARTELS

ABSTRACT. Averaging techniques are popular tools in adaptive finite element methods for the numerical treatment of second order partial differential equations since they provide efficient a posteriori error estimates by a simple post-processing. In this paper, their reliability is shown for conforming, nonconforming, and mixed low order finite element methods in a model situation: the Laplace equation with mixed boundary conditions. Emphasis is on possibly unstructured grids, nonsmoothness of exact solutions, and a wide class of averaging techniques. Theoretical and numerical evidence supports that the reliability is up to the smoothness of given right-hand sides.

1. INTRODUCTION

Error control and efficient mesh-design in finite element simulations of computational engineering and scientific computing finite element simulations is frequently based on a posteriori error estimates. One of the more popular techniques is local or global averaging, e.g., in form of the ZZ -error indicator [ZZ]. Efficiency and reliability of this estimator were known only for very structured grids and for solutions of higher regularity and then we have even asymptotic exactness [V]. Numerical experiments in [Baetal] showed that averaging techniques were quite more reliable on irregular meshes than expected. For homogeneous Dirichlet conditions and conforming finite element methods, the reliability and efficiency of the ZZ -estimator is proven on unstructured, merely shape-regular grids [R2].

This work is devoted to give theoretical and numerical support for the robust reliability of *all* averaging techniques, robust with respect to violated (local) symmetry of meshes and superconvergence and robust with respect to *other boundary conditions* or *other finite element methods*.

For a more precise description of averaging techniques, let us discuss a discretisation of a conservation equation

$$(1.1) \quad f + \operatorname{div} p = 0$$

Received by the editor August 25, 1999.

2000 *Mathematics Subject Classification.* Primary 65N30, 65R20, 74B20, 74G99, 74H99.

Key words and phrases. A posteriori error estimates, residual based error estimate, adaptive algorithm, reliability, finite element method, mixed finite element method, nonconforming finite element method.

with a given right-hand side $f \in L^2(\Omega)$ and a known approximation $p_h \in L^2(\Omega)^d$ to the unknown exact flux $p \in H(\operatorname{div}; \Omega)$ in the bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$ with piecewise flat boundary. The test function finite element space should include the continuous piecewise affines $\mathcal{S}_D^1(\mathcal{T})$ (with homogeneous Dirichlet boundary conditions) based on a regular (in the sense of [Ci], cf. Section 2) triangulation \mathcal{T} of Ω . Suppose a Galerkin property for $p - p_h$ with $\mathcal{S}_D^1(\mathcal{T})$, i.e.,

$$(1.2) \quad \int_{\Omega} p_h \cdot \nabla v_h \, dx = \int_{\Omega} f v_h \, dx \quad \text{for all } v_h \in \mathcal{S}_D^1(\mathcal{T}).$$

What can be said about the error $\|p - p_h\|_{L^2(\Omega)}$ when we regard p as an unknown and p_h as a known variable?

In averaging techniques, the error estimator is based on a smoother approximation, e.g., in $\mathcal{S}^1(\mathcal{T})^d$, the continuous \mathcal{T} -piecewise linears, to the (components of the) discrete solution p_h . For instance,

$$(1.3) \quad \eta_Z := \min_{q_h \in \mathcal{S}^1(\mathcal{T})^d} \|p_h - q_h\|_{L^2(\Omega)}$$

may serve as a computable error estimator and the elementwise contributions as local error indicators in an adaptive mesh-refining algorithm.

The triangle inequality shows that η_Z is efficient with constant 1 up to higher order terms of the exact solution p , indeed,

$$(1.4) \quad \eta_Z \leq \|p - p_h\|_{L^2(\Omega)} + \min_{q_h \in \mathcal{S}^1(\mathcal{T})^d} \|p - q_h\|_{L^2(\Omega)}.$$

The last term converges as $O(h^2)$ (provided p is smooth enough and h denotes the maximal mesh-size in \mathcal{T}) and so, generically, is of higher order than the error $\|p - p_h\|_{L^2(\Omega)} = O(h)$ in the lowest order finite element method. If the second term in the right-hand side of (1.4) fails to be of higher order, one can still prove efficiency of η_Z using equivalence of global and local averaging (cf. Theorem 3.2) and that local averaging is equivalent to weighted jumps across interelement boundaries. An efficiency estimate with higher order terms that depend on local smoothness of right-hand sides but with unknown constants then follows as in [V].

In practise, we may apply an averaging operator $\mathcal{A} : L^2(\Omega)^d \rightarrow \mathcal{S}^1(\mathcal{T})^d$ to p_h and compute the upper bound $\|p_h - \mathcal{A}p_h\|_{L^2(\Omega)}$ of η_Z . Then, efficiency depends strongly on the approximation properties of \mathcal{A} and deserves further investigation.

In this paper, the focus is on the reliability of η_Z , i.e., we investigate under which conditions an estimate

$$(1.5) \quad \|p - p_h\|_{L^2(\Omega)} \leq c_1 \eta_Z + \text{h.o.t.} \leq c_1 \|p_h - \mathcal{A}p_h\|_{L^2(\Omega)} + \text{h.o.t.}$$

holds, we study what the constant $c_1 > 0$ depends on, what affects the higher-order contributions “h.o.t.”, how to modify the definition of η_Z in the presence of mixed boundary conditions, and how to modify the general setting presented for nonconforming and mixed lowest order finite element methods.

Recall from (1.5) that *any* averaging technique, described by \mathcal{A} , then is reliable up to higher order terms. We also prove equivalence to local modifications of η_Z where the minimisation is over smaller domains, e.g., patches of nodes or edges.

The outline of the paper is as follows. Preliminaries and notation are introduced in Section 2 where we state and prove stability and first order estimates for a certain approximation operator $\mathcal{J} : H_D^1(\Omega) \rightarrow \mathcal{S}_D^1(\mathcal{T})$ essentially designed to yield further local orthogonality properties as in [CV, C2]. Basic estimates are provided in Section 3 for a local and global averaging technique and their equivalence. The

subsequent Sections 4, 5, and 6 display the consequences to averaging techniques in a posteriori error control for first order conforming, nonconforming and mixed finite element schemes. Numerical evidence, reported in Section 7, supports the theoretical results for adaptively refined and evenly perturbed meshes. Although asymptotic exactness is not claimed in this paper, our numerical experiments illustrate that η_Z is a very good approximate to $\|p - p_h\|_{L^2(\Omega)}$ even on perturbed grids.

The proofs are given for a simple elliptic model example with mixed boundary conditions for conforming, nonconforming, and mixed finite elements in two dimensions for notational simplicity. More interesting examples such as higher-order schemes, the application to the Stokes problem or the Navier-Lamé equations without incompressibility-locking will appear elsewhere [BC, CF2, CF3, CF4, CF5].

2. APPROXIMATION IN FINITE ELEMENT SPACES

The Lipschitz boundary $\Gamma = \partial\Omega$ of the bounded domain Ω is split into a closed Dirichlet part Γ_D with positive surface measure and a remaining, relatively open and possibly empty, Neumann part $\Gamma_N := \Gamma \setminus \Gamma_D$. Suppose \mathcal{T} be a regular triangulation of the domain $\Omega \subseteq \mathbb{R}^d$, $d = 1, 2, 3$, in the sense of Ciarlet [BS, Ci] (no hanging node, domain is matched exactly) with piecewise affine Lipschitz boundary $\Gamma = \partial\Omega = \Gamma_D \cup \Gamma_N$, i.e., \mathcal{T} consists of a finite number of closed subsets of $\bar{\Omega}$, that cover $\bar{\Omega} = \cup \mathcal{T}$. Each element $T \in \mathcal{T}$ is either an interval $T = \text{conv}\{a, b\}$ if $d = 1$, a triangle $T = \text{conv}\{a, b, c\}$, or a parallelogram $T = \text{conv}\{a, b, c, d\}$ if $d = 2$. The extremal points a, b, c are called vertices, the faces $E \subseteq \partial T$, e.g., $E = \text{conv}\{a, b\}$, are called edges. The set of all vertices and all edges appearing for some T in \mathcal{T} are denoted as \mathcal{N} and \mathcal{E} . Two distinct and intersecting T_1 and T_2 share either an entire edge or a vertex. Each edge $E \in \mathcal{E}$ on the boundary Γ belongs either to Γ_D , written $E \in \mathcal{E}_D$, or to $\bar{\Gamma}_N$, written $E \in \mathcal{E}_N$. Therefore the set of edges is partitioned into $\mathcal{E}_\Omega := \{E \in \mathcal{E} : E \not\subseteq \Gamma\}$, \mathcal{E}_D , and \mathcal{E}_N . We stress that $\cup \mathcal{E}$, the union of all edges, denotes the skeleton of edges in \mathcal{T} , i.e., the set of all points x that belong to some boundary $x \in \partial T$ of some element $T \in \mathcal{T}$. Finally, $\mathcal{K} := \mathcal{N} \setminus \Gamma_D$ denotes the set of free nodes.

We do not explicitly distinguish between nodes and vertices when we consider conforming finite elements (and avoid these concepts for nonconforming or mixed finite element schemes).

For $T \in \mathcal{T}$, let $P_T^k := \mathcal{P}_k(T)$ if T is a triangle or $P_T^k := \mathcal{Q}_k(T)$ if T is a parallelogram. Here, $\mathcal{P}_k(K)$, resp. $\mathcal{Q}_k(K)$, denotes the set of algebraic polynomials in d variables on K of total, resp. partial, degree $\leq k$. The space $\mathcal{L}^k(\mathcal{T})$ of (possibly discontinuous) \mathcal{T} -piecewise polynomials of degree $\leq k$ is the set of all $U \in L^\infty(\Omega)$ with $U|_T \in P_T^k$ for all T in \mathcal{T} . Set

$$\mathcal{S}^k(\mathcal{T}) := \mathcal{L}^k(\mathcal{T}) \cap C(\Omega) \quad \text{and} \quad \mathcal{S}_D^1(\mathcal{T}) := \{u_h \in \mathcal{S}^1(\mathcal{T}) : u_h|_{\Gamma_D} = 0\}.$$

Let $(\varphi_z|_z \in \mathcal{N})$ denote the nodal basis of $\mathcal{S}^1(\mathcal{T})$, i.e., $\varphi_z \in \mathcal{S}^1(\mathcal{T})$ satisfies $\varphi_z(x) = 0$ if $x \in \mathcal{N} \setminus \{z\}$ and $\varphi_z(z) = 1$. Note that $(\varphi_z|_z \in \mathcal{N})$ is a partition of unity and the open patches

$$(2.1) \quad \omega_z := \{x \in \Omega : 0 < \varphi_z(x)\}$$

form an open cover $(\omega_z : z \in \mathcal{N})$ of Ω with finite overlap.

In order to define a weak interpolation operator $\mathcal{J} : H_D^1(\Omega) \rightarrow \mathcal{S}_D^1(\mathcal{T})$, we modify $(\varphi_z|_z \in \mathcal{K})$ to a partition of unity $(\psi_z|_z \in \mathcal{K})$. For each fixed node $z \in \mathcal{N} \setminus \mathcal{K}$, we

choose a node $\zeta(z) \in \mathcal{K}$ and let $\zeta(z) := z$ if $z \in \mathcal{K}$. In this way, we define a partition of \mathcal{N} into $\text{card}(\mathcal{K})$ classes $I(z) := \{\tilde{z} \in \mathcal{N} : \zeta(\tilde{z}) = z\}$, $z \in \mathcal{K}$. For each $z \in \mathcal{K}$ set

$$(2.2) \quad \psi_z := \sum_{\zeta \in I(z)} \varphi_\zeta$$

and notice that $(\psi_z|_{z \in \mathcal{K}})$ is a partition of unity. It is required that

$$(2.3) \quad \Omega_z := \{x \in \Omega : 0 < \psi_z(x)\}$$

is connected and that $\psi_z \neq \varphi_z$ implies that $\Gamma_D \cap \partial\Omega_z$ has a positive surface measure.

For $g \in L^1(\Omega)$ and $z \in \mathcal{K}$ let $g_z \in \mathbb{R}$ be

$$(2.4) \quad g_z := \frac{\int_{\Omega_z} g \psi_z \, dx}{\int_{\Omega_z} \varphi_z \, dx},$$

and then define

$$(2.5) \quad \mathcal{J}g := \sum_{z \in \mathcal{K}} g_z \varphi_z \in \mathcal{S}_D^1(\mathcal{T}).$$

The local mesh-sizes are denoted by $h_{\mathcal{T}}$ and $h_{\mathcal{E}}$, where $h_{\mathcal{T}} \in \mathcal{L}^0(\mathcal{T})$ denotes the element-size, $h_{\mathcal{T}}|_T := h_T := \text{diam}(T)$ for $T \in \mathcal{T}$, and the edge-size $h_{\mathcal{E}} \in L^\infty(\cup \mathcal{E})$ is defined on the union or skeleton $\cup \mathcal{E}$ of all edges E in \mathcal{E} by $h_{\mathcal{E}}|_E := h_E := \text{diam}(E)$. The patch-size $h_z := \text{diam}(\Omega_z)$ is defined for each node $z \in \mathcal{K}$ separately.

Theorem 2.1. *There exist $(h_{\mathcal{T}}, h_{\mathcal{E}})$ -independent constants $c_2, c_3, c_4, c_5 > 0$ such that for all $g \in H_D^1(\Omega)$ and $f \in L^2(\Omega)$ there holds*

$$(2.6) \quad \|\nabla \mathcal{J}g - \nabla g\|_{L^2(\Omega)} \leq c_2 \|\nabla g\|_{L^2(\Omega)},$$

$$(2.7) \quad \int_{\Omega} f(g - \mathcal{J}g) \, dx \leq c_3 \|\nabla g\|_{L^2(\Omega)} \left(\sum_{z \in \mathcal{K}} h_z^2 \min_{f_z \in \mathbb{R}} \|f - f_z\|_{L^2(\Omega_z)}^2 \right)^{1/2},$$

$$(2.8) \quad \|h_{\mathcal{T}}^{-1}(g - \mathcal{J}g)\|_{L^2(\Omega)} \leq c_4 \|\nabla g\|_{L^2(\Omega)},$$

$$(2.9) \quad \|h_{\mathcal{E}}^{-1/2}(g - \mathcal{J}g)\|_{L^2(\Gamma_N)} \leq c_5 \|\nabla g\|_{L^2(\Omega)}.$$

The constants c_2, c_3, c_4, c_5 only depend on $\Omega, \Gamma_D, \Gamma_N$ and the shape of the elements and patches (not on their sizes).

Remark 2.1. The assertion of the theorem holds verbatim for three space dimensions where \mathcal{T} consists of tetrahedra or parallelepipeds with the same proof.

Proof. In this proof and at similar occasions, \lesssim abbreviates an inequality \leq up to a constant $(h_{\mathcal{T}}, h_{\mathcal{E}})$ -independent factor. Also, $\|\cdot\|_{p,K}$ abbreviates $\|\cdot\|_{L^p(K)}$ and we neglect K if Ω is meant, i.e., $\|\cdot\|_2 := \|\cdot\|_{2,\Omega}$. Hence, e.g., (2.6) could be phrased as $\|\nabla \mathcal{J}g - \nabla g\|_2 \lesssim \|\nabla g\|_2$.

The key estimate for the stability and the approximation property of \mathcal{J} will be

$$(2.10) \quad \|g_z \varphi_z - g \psi_z\|_{2,\Omega_z} \lesssim h_z \|\nabla g\|_{2,\Omega_z} \quad (z \in \mathcal{K}).$$

For the proof of (2.10), let \bar{g}_z denote the integral mean of g on Ω_z . Then, using the definition (2.4) for the coefficients g_z , Cauchy's and Young's inequality, we infer,

with $c_6 := \|1\|_{2,\Omega_z}/\|\varphi_z^{1/2}\|_{2,\Omega_z}$,

(2.11)

$$\begin{aligned} \|\varphi_z^{1/2}(g_z - \bar{g}_z)\|_{2,\Omega_z}^2 &= \int_{\Omega_z} \varphi_z(\bar{g}_z - g)(\bar{g}_z - g_z) dx + \int_{\Omega_z} (\psi_z - \varphi_z)g(g_z - \bar{g}_z) dx \\ &\leq \frac{1}{4}\|\varphi_z^{1/2}(g_z - \bar{g}_z)\|_{2,\Omega_z}^2 + \|g - \bar{g}_z\|_{2,\Omega_z}^2 + \frac{1}{4c_6^2}\|g_z - \bar{g}_z\|_{2,\Omega_z}^2 \\ &\quad + c_6^2\|(\psi_z - \varphi_z)g\|_{2,\Omega_z}^2. \end{aligned}$$

Absorbing $\frac{1}{4c_6^2}\|g_z - \bar{g}_z\|_{2,\Omega_z}^2 \leq \frac{1}{4}\|\varphi_z^{1/2}(g_z - \bar{g}_z)\|_{2,\Omega_z}^2$, we deduce

$$(2.12) \quad \|\varphi_z^{1/2}(g_z - \bar{g}_z)\|_{2,\Omega_z}^2 \lesssim \|g - \bar{g}_z\|_{2,\Omega_z}^2 + \|(\psi_z - \varphi_z)g\|_{2,\Omega_z}^2.$$

A Poincaré inequality yields

$$\|g - \bar{g}_z\|_{2,\Omega_z} \lesssim h_z\|\nabla g\|_{2,\Omega_z}.$$

Note that $(\psi_z - \varphi_z)g$ is nonzero only if $\Gamma_D \cap (\partial\Omega_z)$ has positive surface measure. Since g vanishes there, Friedrichs' inequality shows

$$\|g\|_{2,\Omega_z} \lesssim h_z\|\nabla g\|_{2,\Omega_z}.$$

Therefore, (2.12) results in

$$(2.13) \quad \|\varphi_z^{1/2}(g_z - \bar{g}_z)\|_{2,\Omega_z} \lesssim h_z\|\nabla g\|_{2,\Omega_z}.$$

To prove (2.10), we use the triangle inequality, (2.13), and again Cauchy's and Friedrichs' inequality to verify

$$(2.14) \quad \begin{aligned} \|g_z\varphi_z - g\psi_z\|_{2,\Omega_z} &\leq \|(g_z - \bar{g}_z)\varphi_z\|_{2,\Omega_z} + \|(g - \bar{g}_z)\varphi_z\|_{2,\Omega_z} \\ &\quad + \|(\psi_z - \varphi_z)g\|_{2,\Omega_z} \lesssim h_z\|\nabla g\|_{2,\Omega_z}. \end{aligned}$$

To prove (2.7), we use that $(\psi_z|_{z \in \mathcal{K}})$ is a partition of unity and obtain with (2.10), (2.4) for any $f_z \in \mathbb{R}$ that

$$(2.15) \quad \begin{aligned} \int_{\Omega} f(g - \mathcal{J}g) dx &= \sum_{z \in \mathcal{K}} \int_{\Omega_z} f(g\psi_z - g_z\varphi_z) dx = \sum_{z \in \mathcal{K}} \int_{\Omega_z} (f - f_z)(g\psi_z - g_z\varphi_z) dx \\ &\lesssim \sum_{z \in \mathcal{K}} \|f - f_z\|_{2,\Omega_z} h_z \|\nabla g\|_{2,\Omega_z} \lesssim \left(\sum_{z \in \mathcal{K}} h_z^2 \|f - f_z\|_{2,\Omega_z}^2 \right)^{1/2} \|\nabla g\|_{2,\Omega}. \end{aligned}$$

In the last step we used that $(\varphi_z|_{z \in \mathcal{K}})$ has a finite overlap that depends on the shape of the elements only. This concludes the proof of (2.7).

The remaining part of the proof uses standard arguments and is therefore sketched for brevity. To prove (2.8) we let $f := h_T^{-2}(g - \mathcal{J}g)$ and $f_z = 0$, $z \in \mathcal{K}$, in (2.7). To verify (2.6) we use $\sum_{z \in \mathcal{K}} \psi_z = 1$ and $\sum_{z \in \mathcal{K}} \nabla \psi_z = 0$ and repeat the triangle inequality several times for

$$(2.16) \quad \|\nabla g - \nabla \mathcal{J}g\|_2^2 \lesssim \sum_{z \in \mathcal{K}} \|\nabla(\psi_z g - \varphi_z g_z)\|_2^2.$$

With Friedrichs' and Poincaré's inequality we infer

$$(2.17) \quad \begin{aligned} \|\nabla(g\psi_z - g_z\varphi_z)\|_{2,\Omega_z} &\leq \|(\psi_z - \varphi_z)\nabla g\|_{2,\Omega_z} + \|\nabla(\varphi_z(g_z - \bar{g}_z))\|_{2,\Omega_z} \\ &\quad + \|\nabla(\varphi_z(\bar{g}_z - g))\|_{2,\Omega_z} + \|g\nabla(\psi_z - \varphi_z)\|_{2,\Omega_z} \lesssim \|\nabla g\|_{2,\Omega_z} + \|\nabla(\varphi_z(g_z - \bar{g}_z))\|_{2,\Omega_z}. \end{aligned}$$

It remains to estimate $\|\nabla(\varphi_z(g_z - \bar{g}_z))\|_{2,\Omega_z}$ with (2.10), Friedrichs' inequality, and the above arguments. A trace inequality [BS, Cl, CF1] of the form

$$(2.18) \quad \|w\|_{2,E} \lesssim h_E^{-1/2} \|w\|_{2,T} + h_E^{1/2} \|\nabla w\|_{2,T}$$

for $E \in \mathcal{E}_N$ and $T \in \mathcal{T}$ with $E \subset \partial T \cap \bar{\Gamma}_N$ together with (2.6) and (2.8) implies (2.9). □

3. BASIC ESTIMATES

In this section we first derive with the approximation operator \mathcal{J} a global error estimate for a posteriori error control by averaging processes in an abstract setting. We then show the equivalence of local and global averaging techniques. The estimates of this section are then specified, and thereby proved to be substantial, in the subsequent sections to conforming, nonconforming, and mixed finite element methods.

Theorem 3.1. *Suppose $p, q \in H(\text{div}; \Omega)$ and $p_h \in \mathcal{L}^k(\mathcal{T})^d$ with $p \cdot n, q \cdot n \in L^2(\Gamma_N)$ and*

$$(3.1) \quad \int_{\Omega} (p - p_h) \cdot \nabla w_h \, dx = 0 \quad \text{for all } w_h \in \mathcal{S}_D^1(\mathcal{T}).$$

Then there holds

$$(3.2) \quad \begin{aligned} & \sup_{\substack{w \in H_D^1(\Omega) \\ \|\nabla w\|_{L^2(\Omega)}=1}} \int_{\Omega} (p - p_h) \cdot \nabla w \, dx \\ & \leq c_2 \|p_h - q\|_{L^2(\Omega)} + c_3 \left(\sum_{z \in \mathcal{K}} h_z^2 \min_{f_z \in \mathbb{R}} \|\text{div}(p - q) - f_z\|_{L^2(\Omega_z)}^2 \right)^{1/2} \\ & \quad + c_5 \|h_{\mathcal{E}}^{1/2} (p - q) \cdot n\|_{L^2(\Gamma_N)}. \end{aligned}$$

Proof. According to (3.1), (2.6), Cauchy's inequality, and an integration by parts we have, for each $w \in H_D^1(\Omega)$ with $\|\nabla w\|_{L^2(\Omega)} = 1$, that

$$(3.3) \quad \begin{aligned} \int_{\Omega} (p - p_h) \cdot \nabla w \, dx &= \int_{\Omega} (p - p_h) \cdot \nabla(w - \mathcal{J}w) \, dx \\ &= \int_{\Omega} (p - q) \cdot \nabla(w - \mathcal{J}w) \, dx + \int_{\Omega} (q - p_h) \cdot \nabla(w - \mathcal{J}w) \, dx \\ &\leq \int_{\Gamma_N} (w - \mathcal{J}w) (p - q) \cdot n \, dx - \int_{\Omega} (w - \mathcal{J}w) \text{div}(p - q) \, dx + c_2 \|p_h - q\|_{2,\Omega} \end{aligned}$$

since w and $\mathcal{J}w$ vanish on $\partial\Omega \setminus \Gamma_N$. Owing to (2.7) and (2.9) in Theorem 2.1, we conclude (3.2) from (3.3) and Cauchy's inequality. □

The second result justifies local averaging. For each edge $E \in \mathcal{E}_{\Omega}$, let $\omega_E := \text{int}(T_1 \cup T_2)$ and $\mathcal{T}_E := \{T_1, T_2\}$ for the two distinct elements $T_1, T_2 \in \mathcal{T}$ with $E = T_1 \cap T_2$ and for each edge $E \in \mathcal{E}_N$, let $\omega_E := \text{int}(T)$ and $\mathcal{T}_E := \{T\}$ for the element $T \in \mathcal{T}$ with $E = T \cap \bar{\Gamma}_N$. Let $\mathcal{L}^k(\mathcal{E}_N)$ denote the (possibly discontinuous) \mathcal{E}_N -piecewise polynomials of degree $\leq k$ on Γ_N and let $\mathcal{S}^k(\mathcal{T}_E) := \mathcal{L}^k(\mathcal{T}_E) \cap C(\omega_E)$.

Theorem 3.2. *There exists an $(h_{\mathcal{T}}, h_{\mathcal{E}})$ -independent constant $c_7 > 0$ which depends on the shape of the elements in \mathcal{T} and on the polynomial degree $k \geq 1$,*

$c_8 = \max_{T \in \mathcal{T}} \text{card} \{E \in \mathcal{E}_\Omega \cup \mathcal{E}_N : E \subseteq \partial T\}$, such that, for all $(p_h, g_h) \in \mathcal{L}^{k-1}(\mathcal{T})^d \times \mathcal{L}^k(\mathcal{E}_N)$, we have

$$(3.4) \quad \begin{aligned} & c_7 \min_{q_h \in \mathcal{S}^k(\mathcal{T})^d} \left(\|p_h - q_h\|_{L^2(\Omega)}^2 + \|h_{\mathcal{E}}^{1/2}(g_h - q_h \cdot n)\|_{L^2(\Gamma_N)}^2 \right) \\ & \leq \sum_{E \in \mathcal{E}_\Omega \cup \mathcal{E}_N} \min_{q_E \in \mathcal{S}^k(\mathcal{T}_E)^d} \left(\|p_h - q_E\|_{L^2(\omega_E)}^2 + h_E \|g_h - q_E \cdot n\|_{L^2(E \cap \Gamma_N)}^2 \right) \\ & \leq c_8 \min_{q_h \in \mathcal{S}^k(\mathcal{T})^d} \left(\|p_h - q_h\|_{L^2(\Omega)}^2 + \|h_{\mathcal{E}}^{1/2}(g_h - q_h \cdot n)\|_{L^2(\Gamma_N)}^2 \right). \end{aligned}$$

Proof. The upper estimate follows from $q_E := q_h|_{\omega_E} \in \mathcal{S}^k(\mathcal{T}_E)^d$ for all $q_h \in \mathcal{S}^k(\mathcal{T})^d$ and a rearrangement of the sums over edges and elements. To verify the lower estimate in (3.4) we consider a subspace $\tilde{\mathcal{S}}^k(\mathcal{T})$ of $\mathcal{S}^k(\mathcal{T})$,

$$\tilde{\mathcal{S}}^k(\mathcal{T}) := \left\{ \sum_{z \in \mathcal{N}} q_z \varphi_z : q_z \in \mathcal{S}^{k-1}(\mathcal{T}_z)^d \right\} \subseteq \mathcal{S}^k(\mathcal{T}),$$

where $\mathcal{T}_z = \{T \in \mathcal{T} : T \subseteq \bar{\omega}_z\}$ denotes the restriction of the triangulation \mathcal{T} to ω_z . Since $\{(q_h, q_h \cdot n|_{\Gamma_N}) : q_h \in \tilde{\mathcal{S}}^k(\mathcal{T})^d\}$ is a closed convex subset of $L^2(\Omega)^d \times L^2(\Gamma_N)$, the best-approximation problem

$$(3.5) \quad \min_{q_h \in \tilde{\mathcal{S}}^k(\mathcal{T})^d} \left(\|p_h - q_h\|_2^2 + \|h_{\mathcal{E}}^{1/2}(g_h - q_h \cdot n)\|_{2, \Gamma_N}^2 \right)$$

defines an orthogonal relation, namely, for all $q_z \in \mathcal{S}^{k-1}(\mathcal{T}_z)^d$,

$$(3.6) \quad \int_{\Omega} (p_h - \tilde{q}_h) \cdot q_z \varphi_z \, dx + \int_{\Gamma_N} h_{\mathcal{E}} (g_h - \tilde{q}_h \cdot n) q_z \cdot n \varphi_z \, ds = 0,$$

where $\tilde{q}_h = \sum_{z \in \mathcal{N}} \tilde{q}_z \varphi_z \in \tilde{\mathcal{S}}^k(\mathcal{T})^d$, $\tilde{q}_z \in \mathcal{S}^{k-1}(\mathcal{T}_z)^d$, denotes the minimiser in (3.5). From $\sum_{z \in \mathcal{N}} \varphi_z = 1$, (3.6), and Cauchy's inequality we deduce, for arbitrary $q_z \in \mathcal{S}^{k-1}(\mathcal{T}_z)^d$,

$$(3.7) \quad \begin{aligned} & \|p_h - \tilde{q}_h\|_2^2 + \|h_{\mathcal{E}}^{1/2}(g_h - \tilde{q}_h \cdot n)\|_{2, \Gamma_N}^2 \\ & = \sum_{z \in \mathcal{N}} \left(\int_{\Omega} (p_h - \tilde{q}_h) \varphi_z (p_h - \tilde{q}_z) \, dx + \int_{\Gamma_N} h_{\mathcal{E}} (g_h - \tilde{q}_h \cdot n) \varphi_z (g_h - \tilde{q}_z \cdot n) \, ds \right) \\ & = \sum_{z \in \mathcal{N}} \left(\int_{\Omega} (p_h - \tilde{q}_h) \varphi_z (p_h - q_z) \, dx + \int_{\Gamma_N} h_{\mathcal{E}} (g_h - \tilde{q}_h \cdot n) \varphi_z (g_h - q_z \cdot n) \, ds \right) \\ & \leq (\|p_h - \tilde{q}_h\|_2 + \|h_{\mathcal{E}}^{1/2}(g_h - \tilde{q}_h \cdot n)\|_{2, \Gamma_N}) \\ & \quad \times \left(\sum_{z \in \mathcal{N}} \left(\|\varphi_z^{1/2}(p_h - q_z)\|_{2, \omega_z}^2 + \|h_{\mathcal{E}}^{1/2} \varphi_z^{1/2}(g_h - q_z \cdot n)\|_{2, \Gamma_N}^2 \right) \right)^{1/2}. \end{aligned}$$

For each $z \in \mathcal{N}$, we consider the semi-norms on a finite dimensional subspace of $L^2(\omega_z)^d \times L^2((\partial\omega_z) \cap \Gamma_N)$

$$\begin{aligned} \|(p_h, g_h)\|_{z,1} & := \min_{q_z \in \mathcal{S}^{k-1}(\mathcal{T}_z)^d} \left(\|\varphi_z^{1/2}(p_h - q_z)\|_{2, \omega_z} + \|\varphi_z^{1/2} h_{\mathcal{E}}^{1/2}(g_h - q_z \cdot n)\|_{2, \Gamma_N} \right), \\ \|(p_h, g_h)\|_{z,2} & := \left(\sum_{\substack{E \in \mathcal{E} \\ z \in E}} \min_{q_E \in \mathcal{S}^k(\mathcal{T}_E)^d} \left(\|p_h - q_E\|_{2, \omega_E}^2 + h_E \|g_h - q_E \cdot n\|_{2, \Gamma_N \cap E}^2 \right) \right)^{1/2}. \end{aligned}$$

Then, (3.7) and $\tilde{\mathcal{S}}^k(\mathcal{T}) \subseteq \mathcal{S}^k(\mathcal{T})$ yield

$$(3.8) \quad \min_{q_h \in \mathcal{S}^k(\mathcal{T})^d} (\|p_h - q_h\|_2^2 + \|h_{\mathcal{E}}^{1/2}(g_h - q_h \cdot n)\|_{2,\Gamma_N}^2) \lesssim \sum_{z \in \mathcal{N}} \|(p_h, g_h)\|_{z,1}^2.$$

We claim $\|\cdot\|_{z,1} \lesssim \|\cdot\|_{z,2}$. For a proof, suppose $\|(p_h, g_h)\|_{z,2} = 0$. Then, for each E that is an inner edge of ω_z , we have $p_h = q_E$ on the open set ω_E for some $q_E \in \mathcal{S}^k(\mathcal{T}_E)^d$. Since $p_h \in \mathcal{L}^{k-1}(\mathcal{T})^d$, we find that $p_h|_{\omega_E} \in \mathcal{S}^{k-1}(\mathcal{T}_E)$. The set of all such ω_E is a cover of ω_z and there is a sequence E_1, \dots, E_m of inner edges such that $\omega_{E_j} \cap \omega_{E_{j+1}} \neq \emptyset$, so that we deduce $p_h|_{\omega_z} \in \mathcal{S}^{k-1}(\mathcal{T}_z)$. Moreover, $g_h = p_h \cdot n$ on each edge $E \subseteq \bar{\Gamma}_N$ with $z \in E$, while for edges $E \subseteq \partial\omega_z \cap \bar{\Gamma}_N$ with $z \notin E$ we have $\varphi_z|_E = 0$. Altogether, we deduce $\|(p_h, g_h)\|_{z,1} = 0$. A compactness and scaling argument then shows our claim

$$(3.9) \quad \|\cdot\|_{z,1} \lesssim \|\cdot\|_{z,2} \quad \text{on } \mathcal{L}^{k-1}(\mathcal{T}_z)^d \times \mathcal{L}^k(\{E \in \mathcal{E} : E \subseteq \partial\omega_z\}).$$

Utilizing (3.9) in (3.8), we conclude

$$(3.10) \quad \begin{aligned} & \min_{q_h \in \mathcal{S}^k(\mathcal{T})^d} (\|p_h - q_h\|_2^2 + \|h_{\mathcal{E}}^{1/2}(g_h - q_h \cdot n)\|_{2,\Gamma_N}^2) \\ & \lesssim \sum_{z \in \mathcal{N}} \|(p_h, g_h)\|_{z,1}^2 \lesssim \sum_{z \in \mathcal{N}} \|(p_h, g_h)\|_{z,2}^2 \\ & \lesssim \sum_{E \in \mathcal{E}} \min_{q_E \in \mathcal{S}^k(\mathcal{T}_E)^d} (\|p_h - q_E\|_{2,\omega_E}^2 + h_E \|g_h - q_E \cdot n\|_{2,\Gamma_N \cap E}^2). \quad \square \end{aligned}$$

Remark 3.1. The assertions of Theorems 3.1 and 3.2 hold verbatim for three space dimensions where \mathcal{T} consists of tetrahedra or parallelepipeds with the same proofs.

4. APPLICATIONS TO CONFORMING FINITE ELEMENT SCHEMES

Given right-hand sides $f \in L^2(\Omega)$, $g \in L^2(\Gamma_N)$, and $u_D \in H^1(\Gamma_D)$, let $u \in H^1(\Omega)$ denote the unique weak solution to

$$(4.1) \quad -\Delta u = f \quad \text{in } \Omega,$$

$$(4.2) \quad u = u_D \quad \text{on } \Gamma_D,$$

$$(4.3) \quad \partial u / \partial n = g \quad \text{on } \Gamma_N.$$

Suppose a finite element scheme, based on a regular triangulation \mathcal{T} , provided a discrete flux $p_h := \nabla u_h$ to the exact flux $p := \nabla u \in H(\text{div}; \Omega)$ such that $u_h \in \mathcal{S}^1(\mathcal{T})$, $u_h(z) = u_D(z)$ for all $z \in \mathcal{N} \cap \Gamma_D$ and

$$(4.4) \quad \int_{\Omega} \nabla u_h \cdot \nabla w_h \, dx = \int_{\Omega} f w_h \, dx + \int_{\Gamma_N} g w_h \, ds \quad \text{for all } w_h \in \mathcal{S}_D^1(\mathcal{T}).$$

Theorem 4.1. *There exists an $(h_{\mathcal{T}}, h_{\mathcal{E}})$ -independent constant $c_9 > 0$ (that depends on k and the shape of the elements and patches) such that*

$$(4.5) \quad \begin{aligned} \|\nabla(u - u_h)\|_{L^2(\Omega)} & \leq \min_{q_h \in \mathcal{S}^k(\mathcal{T})^d} (c_9 \|\nabla u_h - q_h\|_{L^2(\Omega)} + 2c_5 \|h_{\mathcal{E}}^{1/2}(g - q_h \cdot n)\|_{L^2(\Gamma_N)}) \\ & \quad + \inf_{v|_{\Gamma_D} = u_D} \|\nabla(u_h - v)\|_{L^2(\Omega)} + 2c_3 \left(\sum_{z \in \mathcal{K}} h_z^2 \min_{f_z \in \mathbb{R}} \|f - f_z\|_{L^2(\Omega_z)}^2 \right)^{1/2}. \end{aligned}$$

In the infimum, “ $v|_{\Gamma_D} = u_D$ ” stands for all $v \in H^1(\Omega)$ with $v = u_D$ on Γ_D .

Proof. Abbreviate $e := u - u_h$ and let $q_h \in \mathcal{S}^k(\mathcal{T})^d$. Assume that $v \in H^1(\Omega)$ satisfies $v = u_D$ on Γ_D and $\|\nabla(u_h - v)\|_2 \leq \|\nabla e\|_2$. Recall $p = \nabla u$ and $p_h = \nabla u_h$. Then (4.1)-(4.4) imply (3.1). Hence, we may choose $q = q_h$ and $w = u - v$ in Theorem 3.1 to obtain with Cauchy's inequality for the second term that

$$\begin{aligned}
 \|\nabla e\|_2^2 &= \int_{\Omega} \nabla e \cdot \nabla w \, dx + \int_{\Omega} \nabla e \cdot \nabla(v - u_h) \, dx \\
 &\leq \|\nabla w\|_2 \left(c_2 \|p_h - q_h\|_2 + c_5 \|h_{\mathcal{E}}^{1/2} (g - q_h \cdot n)\|_{2, \Gamma_N} \right. \\
 &\quad \left. + c_3 \left(\sum_{z \in \mathcal{K}} h_z^2 \min_{f_z \in \mathbb{R}} \|f + \operatorname{div} q_h - f_z\|_{2, \Omega_z}^2 \right)^{1/2} \right) \\
 &\quad + \|\nabla(u_h - v)\|_2 \|\nabla e\|_2.
 \end{aligned} \tag{4.6}$$

Since $\|\nabla w\|_2 \leq \|\nabla e\|_2 + \|\nabla(u_h - v)\|_2 \leq 2\|\nabla e\|_2$, we can divide (4.6) by $\|\nabla e\|_2$ to verify

$$\begin{aligned}
 \|\nabla e\|_2 &\leq 2c_2 \|p_h - q_h\|_2 + 2c_5 \|h_{\mathcal{E}}^{1/2} (g - q_h \cdot n)\|_{2, \Gamma_N} + \|\nabla(u_h - v)\|_2 \\
 &\quad + 2c_3 \left(\sum_{z \in \mathcal{K}} h_z^2 \min_{f_z \in \mathbb{R}} \|f + \operatorname{div} q_h - f_z\|_{2, \Omega_z}^2 \right)^{1/2}.
 \end{aligned} \tag{4.7}$$

Let $\operatorname{div}_{\mathcal{T}}$ denote the \mathcal{T} -piecewise action of the div -operator. The triangle inequality in the last summand in (4.7) and $h_z \lesssim h_T$ for $z \in T \cap \mathcal{N}$ and $T \in \mathcal{T}$ and a summation over elements show

$$\begin{aligned}
 &\sum_{z \in \mathcal{K}} h_z^2 \min_{f_z \in \mathbb{R}} \|f + \operatorname{div} q_h - f_z\|_{2, \Omega_z}^2 \\
 &\lesssim \|h_{\mathcal{T}} \operatorname{div}_{\mathcal{T}}(p_h - q_h)\|_2^2 + \sum_{z \in \mathcal{K}} h_z^2 \min_{f_z \in \mathbb{R}} \|f + \operatorname{div}_{\mathcal{T}} p_h - f_z\|_{2, \Omega_z}^2.
 \end{aligned} \tag{4.8}$$

Note that $\operatorname{div}_{\mathcal{T}} p_h = \Delta_{\mathcal{T}} u_h = 0$ for our choices of $u_h \in \mathcal{S}^1(\mathcal{T})$. A \mathcal{T} -elementwise inverse estimate shows $\|h_{\mathcal{T}} \operatorname{div}_{\mathcal{T}}(p_h - q_h)\|_2 \lesssim \|p_h - q_h\|_2$ (with a constant that depends on the shape of the finite elements only). Utilising this in (4.7)–(4.8), we deduce (4.5). \square

Remark 4.1. In the proof of Theorem 4.1 we used the assumption that u_h is of lowest order, i.e., $\nabla u_h \in \mathcal{L}^0(\mathcal{T})^d$, for the purpose of estimating $\|h_{\mathcal{T}} \operatorname{div} q_h\|_{L^2(\Omega)}$ by $\|q_h - \nabla u_h\|_{L^2(\Omega)}$. We refer to [BC] for related error estimates for higher order methods.

The subsequent lemma shows that $\inf_{v|_{\Gamma_D} = u_D} \|\nabla(u_h - v)\|_{L^2(\Omega)}$ is a higher order term.

Lemma 4.1. *Suppose that $u_h(z) = u_D(z)$ for all $z \in \mathcal{N} \cap \Gamma_D$. Then there exists an $h_{\mathcal{E}}$ -independent constant $c_{10} > 0$ (that depends on the shapes of the elements only) such that*

$$\inf_{v|_{\Gamma_D} = u_D} \|\nabla(u_h - v)\|_{L^2(\Omega)} \leq c_{10} \|h_{\mathcal{E}}^{1/2} \partial(u_h - u_D) / \partial s\|_{L^2(\Gamma_D)}. \tag{4.9}$$

If $u_D \in H^2(\mathcal{E}_D) := \{v \in L^2(\Gamma_D) : \forall E \in \mathcal{E}_D, v|_E \in H^2(E)\}$, we have

$$\inf_{v|_{\Gamma_D} = u_D} \|\nabla(u_h - v)\|_{L^2(\Omega)} \leq c_{10} \|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}}^2 u_D / \partial s^2\|_{L^2(\Gamma_D)}. \tag{4.10}$$

Proof. Let $E \in \mathcal{E}_D$ belong to some $T \in \mathcal{T}$ and denote $\gamma := \Gamma_D \cap \partial T$. We determine $w \in H^1(T)$ by extending the boundary values $w|_\gamma = u_h - u_D$ and $w|_{\partial T \setminus \gamma} = 0$. Note that w is continuous on ∂T since u_h interpolates $u_D = v$ at each node on Γ_D . An harmonic extension of $w|_{\partial T}$ to $w \in H^1(T)$ yields

$$(4.11) \quad \|\nabla w\|_{2,T} \lesssim \|w\|_{H^{1/2}(\partial T)} \lesssim \|w\|_{2,\partial T}^{1/2} \|\partial w / \partial s\|_{2,\partial T}^{1/2},$$

where we applied an interpolation estimate. A one-dimensional integration argument shows $\|w\|_{2,\partial T} \leq h_T \|\partial w / \partial s\|_{2,\partial T}$. Consequently,

$$(4.12) \quad \|\nabla w\|_{2,T} \lesssim h_T^{1/2} \|\partial w / \partial s\|_{2,\partial T} = h_T^{1/2} \|\partial(u_h - u_D) / \partial s\|_{2,\gamma}.$$

A scaling argument guarantees that the constant in (4.12) is h_T -independent. Defining v by $u_h - w$ on elements on Γ_D and by zero on other elements then shows the lemma. The second estimate follows from $\|w\|_{2,\partial T} \leq h_T^2 \|\partial^2 w / \partial s^2\|_{2,\partial T}$. \square

Lemma 4.2. *Suppose $g \in H^1(\mathcal{E}_N)$ and, for each node $z \in \mathcal{N} \cap \bar{\Gamma}_N$ where the outer unit normal n on Γ_N is continuous (hence constant in a neighbourhood of z as Γ_N is a polygon), let g be continuous. Then, the set*

$$(4.13) \quad \mathcal{S}_N^1(\mathcal{T}, g) := \{q_h \in \mathcal{S}^1(\mathcal{T})^d : \forall E \in \mathcal{E}_N \forall z \in E \cap \mathcal{N}, q_h(z) \cdot n_E = g(z)\}$$

is nonvoid and, for each $q_h \in \mathcal{S}_N^1(\mathcal{T}, g)$,

$$(4.14) \quad \|h_\mathcal{E}^{1/2}(g - q_h \cdot n)\|_{L^2(\Gamma_N)} \leq \|h_\mathcal{E}^{3/2} \partial_\mathcal{E} g / \partial s\|_{L^2(\Gamma_N)}.$$

Proof. Elementary estimates on each edge on Γ_N verify (4.14); the proof of $\mathcal{S}_N^1(\mathcal{T}, g) \neq \emptyset$ follows from an explicit construction in Example 4.1. \square

Example 4.1. We define an operator $\mathcal{A} : L^2(\Omega)^2 \rightarrow \mathcal{S}_N^1(\mathcal{T}, g)$ by

$$(4.15) \quad \mathcal{A}p := \sum_{z \in \mathcal{N}} p_z \varphi_z,$$

where $p_z := \int_{\omega_z} p \, dx := \frac{1}{|\omega_z|} \int_{\omega_z} p \, dx \in \mathbb{R}^2$ for $z \in \mathcal{N} \setminus \bar{\Gamma}_N$ while we incorporate $\mathcal{A}p(z) \cdot n_E = g(z)$ for $z \in \mathcal{N} \cap \bar{\Gamma}_N$. In case $z = E_1 \cap E_2$ for two distinct edges $E_1, E_2 \in \mathcal{E}_N$ with distinct outer unit normals n_{E_1}, n_{E_2} on E_1, E_2 at a corner z we choose $p_z \in \mathbb{R}^2$ to be the unique solution of the 2×2 linear system

$$(4.16a) \quad n_{E_1} \cdot p_z = g|_{E_1}(z) \quad \text{and} \quad n_{E_2} \cdot p_z = g|_{E_2}(z).$$

In the remaining cases $z \in E_1 \cap \Gamma_D$ for $E_1 \in \mathcal{E}_N$ or $z = E_1 \cap E_2$ with two parallel edges $E_1, E_2 \in \mathcal{E}_N$ with the unit tangent vector t_{E_1} let $p_z \in \mathbb{R}^2$ solve

$$(4.16b) \quad n_{E_1} \cdot p_z = g|_{E_1}(z) \quad \text{and} \quad t_{E_1} \cdot p_z = \int_{\omega_z} t_{E_1} \cdot p \, dx. \quad \square$$

The following corollary is (1.5) with a constant $c_1 = c_9$ as in Theorem 4.1 and with specified higher order terms from Lemma 4.1 and 4.2 and a Poincaré inequality.

Corollary 4.1. *Under the conditions of Theorem 4.1 and Lemmas 4.1 and 4.2 we have for $f \in H^1(\Omega)$ that*

$$(4.17) \quad \|\nabla(u - u_h)\|_{L^2(\Omega)} \leq c_9 \min_{q_h \in \mathcal{S}_N^1(\mathcal{T}, g)} \|\nabla u_h - q_h\|_{L^2(\Omega)} + c_{11} \left(\|h_\mathcal{E}^{3/2} \partial_\mathcal{E}^2 u_D / \partial s^2\|_{L^2(\Gamma_D)} + \|h_\mathcal{E}^{3/2} \partial_\mathcal{E} g / \partial s\|_{L^2(\Gamma_N)} + \|h_T^2 \nabla f\|_{L^2(\Omega)} \right).$$

The $(h_{\mathcal{T}}, h_{\mathcal{E}})$ -independent constant $c_{11} > 0$ depends on the shape of the elements and patches only. □

Remark 4.2. Let us emphasise that the derivatives along Γ are required only \mathcal{E} -piecewisely while f needs to be patch-wise (not only elementwise) in H^1 and so $f \in H^1(\Omega)$. For a nonsmooth right-hand side f , $\|h_{\mathcal{T}}^2 \nabla f\|_{L^2(\Omega)}$ may be replaced by a patch-wise L^2 -best approximation error in the approximation through constants of f (cf. (2.7)).

The global averaging process might be too expensive or its approximation may be inefficient and hence a local averaging process of interest. Recall that ω_E is the (interior of the) union of all elements in \mathcal{T} that share the edge $E \in \mathcal{E}$.

Corollary 4.2. *Under the conditions of Theorem 4.1 and Lemmas 4.1 and 4.2 we have for $f \in H^1(\Omega)$ that*

$$\begin{aligned}
 (4.18) \quad \|\nabla(u - u_h)\|_{L^2(\Omega)} &\leq c_{12} \left(\sum_{E \in \mathcal{E}} \min_{q_E \in \mathcal{S}^1(\mathcal{T}_E)^d} \left(\|\nabla u_h - q_E\|_{L^2(\omega_E)}^2 \right. \right. \\
 &\quad \left. \left. + h_E \|g_h - q_E \cdot n\|_{L^2(E \cap \Gamma_N)}^2 \right) \right)^{1/2} \\
 &\quad + c_{11} \left(\|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}}^2 u_D / \partial s^2\|_{L^2(\Gamma_D)} \right. \\
 &\quad \left. + \|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}} g / \partial s\|_{L^2(\Gamma_N)} + \|h_{\mathcal{T}}^2 \nabla f\|_{L^2(\Omega)} \right).
 \end{aligned}$$

The $(h_{\mathcal{T}}, h_{\mathcal{E}})$ -independent constant $c_{12} = \max\{c_9, 2c_5\}/c_7$ depends on the shape of the elements and patches only.

Proof. Theorem 4.1, Lemma 4.1, an approximation g_h of g as in Lemma 4.2, and a Poincaré inequality show

$$\begin{aligned}
 (4.19) \quad \|\nabla(u - u_h)\|_2 &\lesssim \min_{q_h \in \mathcal{S}^1(\mathcal{T})^d} (\|\nabla u_h - q_h\|_2 + \|h_{\mathcal{E}}^{1/2} (g_h - q_h \cdot n)\|_{2, \Gamma_N}) + \|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}} g / \partial s\|_{2, \Gamma_N} \\
 &\quad + \|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}}^2 u_D / \partial s^2\|_{2, \Gamma_D} + \|h_{\mathcal{T}}^2 \nabla f\|_2.
 \end{aligned}$$

This and the first inequality of Theorem 3.2 imply the assertion. □

Remark 4.3. The results of this section hold also in three dimensions where \mathcal{T} consists of tetrahedra or parallelepipeds. The proofs of some details as Lemma 4.1 or Lemma 4.2 require much more technical preparations and so are omitted in this overview.

Remark 4.4. It is shown in [CV, C2] that the edge-contributions (jump differences in the normal fluxes components across edges) dominate in standard residual a posteriori error estimates [BaR, B, BS, CF1, EEHJ, V]. Arguing as in [R1, R2, DMR], one can hence derive alternative proofs of (4.18) and then of (4.17).

Remark 4.5. In an L^∞ -estimate of [HSWW] it is suggested to average over a domain of size $O(h \log(1/h))$ instead of merely over patches or the entire domain to obtain asymptotic exact results.

5. APPLICATIONS TO NONCONFORMING FINITE ELEMENT SCHEMES

In the Laplace problem with mixed boundary conditions (4.1)–(4.3), we suppose that the discrete flux $p_h := \nabla_{\mathcal{T}} u_h \in \mathcal{L}^0(\mathcal{T})^d$, where $\nabla_{\mathcal{T}}$ denotes the \mathcal{T} -piecewise application of the gradient, satisfies

$$(5.1) \quad \int_{\Omega} \nabla_{\mathcal{T}} u_h \cdot \nabla w_h \, dx = \int_{\Omega} f w_h \, dx + \int_{\Gamma_N} g w_h \, ds \quad \text{for all } w_h \in \mathcal{S}_D^1(\mathcal{T}).$$

The usual conformity conditions read for all $E \in \mathcal{E}_{\Omega} \cup \mathcal{E}_D$,

$$(5.2) \quad \int_E [u_h] \, ds = 0,$$

where $[u_h]|_E$ denotes the jump of u_h across $E \in \mathcal{E}_{\Omega}$ and denotes $u_D - u_h$ on Γ_D . Those conditions are satisfied by construction for Crouzeix–Raviart finite elements of lowest order.

Remark 5.1. It is stressed that $\mathcal{S}_D^1(\mathcal{T})$ is a conforming test function space which is included in the nonconforming finite element spaces for triangles or tetrahedra. For parallelograms, (5.1) means that the polynomial degrees are at least of second order to include the conforming term $x_1 x_2$. This technical detail could actually be dropped since the contribution from an enhanced finite element space leads to a higher order term [KS]. We restrict our analysis to triangles or tetrahedra for simplicity.

Theorem 5.1. *Suppose that Γ_N is connected and that Γ_D belongs to only one connectivity component of $\partial\Omega$. Then, there exists an $(h_{\mathcal{T}}, h_{\mathcal{E}})$ -independent constant $c_{13} > 0$ (that depends on $k \geq 1$ and the shape of the elements and patches) such that*

$$(5.3) \quad \begin{aligned} & \|\nabla_{\mathcal{T}}(u - u_h)\|_{L^2(\Omega)} \\ & \leq \min_{q_h \in \mathcal{S}^k(\mathcal{T})^d} (c_{13} \|\nabla_{\mathcal{T}} u_h - q_h\|_{L^2(\Omega)} + c_5 \|h_{\mathcal{E}}^{1/2}(g - q_h \cdot n)\|_{L^2(\Gamma_N)} \\ & \quad + c_5 \|h_{\mathcal{E}}^{1/2}(q_h \cdot t - \partial u_D / \partial s)\|_{L^2(\Gamma_D)}) \\ & \quad + c_3 \left(\sum_{z \in \mathcal{K}} h_z^2 \min_{f_z \in \mathbb{R}} \|f - f_z\|_{L^2(\Omega_z)}^2 \right)^{1/2}. \end{aligned}$$

Here, $t \in \mathcal{L}^0(\mathcal{E}_D)^d$ denotes the unit tangent vector on Γ_D .

Remark 5.2. The following lemma is based on the Helmholtz decomposition of a vector field. The decomposition is available in three dimensions as well (cf., e.g., [GR]) but the notation is more involved so we restrict the discussion to the two-dimensional setting for brevity.

Lemma 5.1. *For all $p - p_h \in L^2(\Omega)^2$, there exist $\alpha, \beta \in H^1(\Omega)$ that satisfy boundary conditions $\alpha|_{\Gamma_D} = 0$ and $\beta|_{\Gamma_N}$ is constant such that*

$$(5.4) \quad p - p_h = \nabla \alpha + \text{Curl} \beta \quad \text{and} \quad \|p - p_h\|_{L^2(\Omega)}^2 = \|\nabla \alpha\|_{L^2(\Omega)}^2 + \|\nabla \beta\|_{L^2(\Omega)}^2.$$

Proof. The lemma follows from the Helmholtz decomposition where $\alpha \in H_D^1(\Omega)$ solves $\Delta \alpha = \text{div}(p - p_h)$ and $\Delta \beta = \text{curl}(p - p_h)$ with proper boundary conditions (cf., e.g., [GR]). □

Proof of Theorem 5.1. For $p = \nabla u$ and $p_h = \nabla_{\mathcal{T}} u_h$, Lemma 5.1 yields

$$(5.5) \quad \|p - p_h\|_2^2 = \int_{\Omega} (p - p_h) \cdot \nabla \alpha \, dx + \int_{\Omega} (p - p_h) \cdot \text{Curl} \beta \, dx.$$

Since Γ_N is connected, we may and will assume without loss of generality that $\beta = 0$ on Γ_N . According to (4.1)–(4.3) and (5.1), we infer (3.1) and hence may choose $q = q_h \in \mathcal{S}^k(\mathcal{T})^d$ and, in case $\alpha \neq 0$, $w = \alpha/\|\alpha\|_{2,\Omega}$ in Theorem 3.1 to obtain

$$(5.6) \quad \int_{\Omega} (p - p_h) \cdot \nabla \alpha \, dx \leq \|\nabla \alpha\|_2 \left(c_2 \|p_h - q_h\|_2 + c_5 \|h_{\mathcal{E}}^{1/2} (g - q_h \cdot n)\|_{2,\Gamma_N} \right. \\ \left. + c_3 \left(\sum_{z \in \mathcal{K}} h_z^2 \min_{f_z \in \mathbb{R}} \|f + \text{div} \, q_h - f_z\|_{2,\Omega_z}^2 \right)^{1/2} \right).$$

The estimate of the last term in (5.5) will follow from Theorem 3.1 as well once we establish an analogy to (3.1), namely

$$(5.7) \quad \int_{\Omega} (p - p_h) \cdot \text{Curl} \, w_h \, dx = 0 \quad \text{for all } w_h \in \mathcal{S}_N^1(\mathcal{T}),$$

where $\mathcal{S}_N^1(\mathcal{T}) := \{v_h \in \mathcal{S}^1(\mathcal{T}) : v_h = 0 \text{ on } \Gamma_N\}$. It is essential to notice that $\partial w_h / \partial s$ is constant and $[u_h]$ has a vanishing integral on any edge. An elementwise integration by parts on the left-hand side of (5.7) yields volume terms $(u - u_h) \text{div}_{\mathcal{T}} \text{Curl} \, w_h = 0$ and edge terms $[(u - u_h) \partial w_h / \partial s] = [u_h] \partial w_h / \partial s$ whose integral vanishes on any E (the case $E \in \mathcal{E}_{\Omega}$ is indicated and the assertion is true for $E \in \mathcal{E}_D$ as well; $w_h = 0$ on Γ_N shows it for $E \in \mathcal{E}_N$). In this way we establish (5.7).

To employ Theorem 3.1, we interchange components, writing in this proof $Q(a_1, a_2) := (-a_2, a_1)$ for vectors, and we interchange the role of the boundaries and adopt Theorem 2.1 and (3.1) where $\tilde{\Gamma}_D = \Gamma_N$ acts as the Dirichlet boundary and $\tilde{\Gamma}_N = \Gamma_D$ acts as the Neumann boundary. Writing $\tilde{p} = Qp$ and $\tilde{p}_h = Qp_h$, (5.7) reads $\int_{\Omega} (\tilde{p} - \tilde{p}_h) \cdot \nabla w_h \, dx = 0$ for all $w_h \in \tilde{\mathcal{S}}_D^1(\mathcal{T}) = \mathcal{S}_N^1(\mathcal{T})$ and this is (3.1). Reading Theorem 3.1 in the present notation, we obtain

$$(5.8) \quad \int_{\Omega} (p - p_h) \cdot \text{Curl} \beta \, dx \\ \leq \|\nabla \beta\|_2 \left(c_2 \|p_h - q_h\|_2 + c_5 \|h_{\mathcal{E}}^{1/2} (\partial u_D / \partial s - q_h \cdot t)\|_{2,\Gamma_D} \right. \\ \left. + c_3 \left(\sum_{z \in \mathcal{K}} h_z^2 \|\text{curl}_{\mathcal{T}}(q_h - p_h)\|_{2,\Omega_z}^2 \right)^{1/2} \right)$$

with $\text{curl}_{\mathcal{T}}(q_h - p_h) := \text{div}_{\mathcal{T}} Q(q_h - p_h)$. In the second last term, $t = Qn$ denotes the unit tangent vector and in the last term we used that $\text{curl}_{\mathcal{T}} p = 0 = \text{curl}_{\mathcal{T}} p_h$.

The remaining arguments are similar to those in the proof of Theorem 4.1 and hence are omitted. □

In contrast to the conforming situation, Theorem 5.1 demands averaging functions to satisfy some conditions on the Dirichlet boundary.

Lemma 5.2. *Suppose $u_D \in H^2(\mathcal{E}_D)$ and, for each node $z \in \mathcal{N} \cap \Gamma_D$ where the outer unit normal n on Γ is continuous, let $\partial u_D / \partial s$ be continuous. Then, the set*

$$(5.9) \quad \mathcal{S}_D^1(\mathcal{T}, u_D) := \{q_h \in \mathcal{S}^1(\mathcal{T})^d : \forall E \in \mathcal{E}_D \forall z \in E \cap \mathcal{N}, q_h(z) \cdot t_E = \partial u_D / \partial s(z)\}$$

is nonvoid and, for each $q_h \in \mathcal{S}_D^1(\mathcal{T}, u_D)$,

$$(5.10) \quad \|h_{\mathcal{E}}^{1/2}(q_h \cdot t_{\mathcal{E}} - \partial u_D / \partial s)\|_{L^2(\Gamma_D)} \leq \|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}}^2 u_D / \partial s^2\|_{L^2(\Gamma_D)}.$$

Proof. Similar to (4.10) in Lemma 4.1 or 4.2. \square

Example 5.1. Assume the conditions of Lemmas 4.2 and 5.2 on the data g and u_D . We define an operator $\mathcal{A} : L^2(\Omega)^2 \rightarrow \mathcal{S}_N^1(\mathcal{T}, g) \cap \mathcal{S}_D^1(\mathcal{T}, u_D)$ by (4.15) and $p_z := \int_{\omega_z} p \, dx$ for $z \in \mathcal{N} \setminus \Gamma$. In case $z \in (\mathcal{N} \cap \bar{\Gamma}_N) \setminus \Gamma_D$ we proceed as in (4.16a), resp. (4.16b). In case $z \in (\mathcal{N} \cap \Gamma_D) \setminus \bar{\Gamma}_N$ we consider the analogous 2×2 systems

$$(5.11a) \quad t_{E_1} \cdot p_z = \partial u_D|_{E_1} / \partial s(z) \quad \text{and} \quad t_{E_2} \cdot p_z = \partial u_D|_{E_2} / \partial s(z)$$

(cf. (4.16a) and notation from Example 4.1), resp., as an analog to (4.16b),

$$(5.11b) \quad t_{E_1} \cdot p_z = \partial u_D|_{E_1} / \partial s(z) \quad \text{and} \quad n_{E_1} \cdot p_z = \int_{\omega_z} n_{E_1} \cdot p \, dx.$$

For $z \in \Gamma_D \cap \bar{\Gamma}_N \subseteq \mathcal{N}$ we require a compatibility condition if $n_{E_1} = t_{E_2}$, namely $g(z) = \partial u_D / \partial s(z)$. Then, we define $p_z \in \mathbb{R}^2$ as in (5.11b) when $E_1 \subseteq \Gamma_D$ (the case $E_2 \subseteq \Gamma_D$ is analogous). For $n_{E_1} \neq \pm t_{E_2}$, we need no further compatibility of the data and solve the 2×2 linear system

$$(5.12) \quad t_{E_1} \cdot p_z = \partial u_D|_{E_1} / \partial s(z) \quad \text{and} \quad n_{E_2}(z) \cdot p_z = g(z).$$

(Here $E_1 \subseteq \Gamma_D$; the case $E_2 \subseteq \Gamma_D$ is analogous.) \square

The modification of (1.5) in the nonconforming setting is a direct consequence of Theorem 5.1, Lemma 5.2 and Example 5.1. Note that Corollary 4.1 is a special case apart from the different treatment of the Dirichlet boundary conditions.

Corollary 5.1. *Under the conditions of Theorem 5.1 and Lemmas 4.2 and 5.2, we have for $f \in H^1(\Omega)$ that*

$$(5.13) \quad \|\nabla_{\mathcal{T}}(u - u_h)\|_{L^2(\Omega)} \leq c_{13} \min_{q_h \in \mathcal{S}_N^1(\mathcal{T}, g) \cap \mathcal{S}_D^1(\mathcal{T}, u_D)} \|\nabla_{\mathcal{T}}(u_h - q_h)\|_{L^2(\Omega)} \\ + c_{14} (\|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}}^2 u_D / \partial s^2\|_{L^2(\Gamma_D)} + \|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}} g / \partial s\|_{L^2(\Gamma_N)} + \|h_{\mathcal{T}}^2 \nabla f\|_{L^2(\Omega)}). \quad \square$$

The analog to Corollary 4.2 concludes this section on lowest order Raviart–Crouzeix finite elements.

Corollary 5.2. *Under the conditions of Theorem 5.1 and Lemmas 4.2 and 5.2, there exists a constant $c_{15} > 0$ such that we have for $f \in H^1(\Omega)$*

$$(5.14) \quad \|\nabla_{\mathcal{T}}(u - u_h)\|_{L^2(\Omega)} \\ \leq c_{15} \left(\sum_{E \in \mathcal{E}} \min_{q_E \in \mathcal{S}^1(\mathcal{T}_E)^d} (\|\nabla_{\mathcal{T}} u_h - q_E\|_{L^2(\omega_E)}^2 \\ + h_E \|g_h - q_E \cdot n\|_{L^2(E \cap \Gamma_N)}^2 + h_E \|q_E \cdot t - u'_{D,h}\|_{L^2(E \cap \Gamma_D)}^2) \right)^{1/2} \\ + c_{14} (\|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}}^2 u_D / \partial s^2\|_{L^2(\Gamma_D)} + \|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}} g / \partial s\|_{L^2(\Gamma_N)} + \|h_{\mathcal{T}}^2 \nabla f\|_{L^2(\Omega)}).$$

Here, $u'_{D,h}$ denotes an approximation of $\partial u_D / \partial s$ as in Lemma 5.2, i.e., $u'_{D,h} = q_h \cdot t_{\mathcal{E}}$ on Γ_D for some $q_h \in \mathcal{S}_D^1(\mathcal{T}, u_D)$. \square

Remark 5.3. The results of this section can be generalised to three space dimensions as all the required tools such as a Helmholtz decomposition are available then as well. Details on the three-dimensional case are omitted for notational simplicity.

Remark 5.4. Arguing as in [CV, C2], one can prove that edge contributions (jumps in the fluxes across edges) dominate the residual based error estimates from [DDPV, C2, KS]. Arguing in the spirit of [R1, R2, DMR], one can hence derive alternative proofs of (5.14) and then of (5.13).

6. APPLICATIONS TO MIXED FINITE ELEMENT SCHEMES

In the Laplace problem with mixed boundary conditions (4.1)–(4.3), we suppose that the discrete flux $p_h \in H(\operatorname{div}, \Omega) \cap \mathcal{L}^k(\mathcal{T})^2$ and the displacement approximation $u_h \in \mathcal{L}^k(\mathcal{T})$ satisfy, for all $q_h \in \operatorname{Curl} \mathcal{S}^1(\mathcal{T})$ with $q_h \cdot n = 0$ on Γ_N , and for all $T \in \mathcal{T}$ and $E \in \mathcal{E}_N$ that

$$(6.1) \quad \int_{\Omega} (p_h \cdot q_h + u_h \operatorname{div} q_h) dx = \int_{\Gamma_D} u_D q_h \cdot n ds,$$

$$(6.2) \quad \int_T (f + \operatorname{div} p_h) dx = 0,$$

$$(6.3) \quad \int_E (g - p_h \cdot n) ds = 0.$$

Remark 6.1. Standard mixed finite element methods of any order, such as Raviart–Thomas (RT), Brezzi–Douglas–Marini (BDM), or Brezzi–Douglas–Fortin–Marini (BDFM) elements (cf. [BF] for details), provide (6.1)–(6.3) [C1].

Theorem 6.1. *Suppose that Γ_N is connected and that Γ_D belongs to only one connectivity component of $\partial\Omega$ and let $f \in H^1(\mathcal{T})$, i.e., $f|_T \in H^1(T)$ for all $T \in \mathcal{T}$. Then, there exists an $(h_{\mathcal{T}}, h_{\mathcal{E}})$ -independent constant $c_{16} > 0$ (that depends on $k \geq 1$ and the shape of the elements and patches) such that*

$$(6.4) \quad \begin{aligned} \|p - p_h\|_{L^2(\Omega)} \leq & \min_{q_h \in \mathcal{S}^k(\mathcal{T})^2} \left(c_2 \|p_h - q_h\|_{L^2(\Omega)} + c_5 \|h_{\mathcal{E}}^{1/2} (\partial u_D / \partial s - q_h \cdot t)\|_{L^2(\Gamma_D)} \right. \\ & \left. + c_3 \left(\sum_{z \in \mathcal{K}} h_z^2 \min_{f_z \in \mathbb{R}} \|\operatorname{curl}_{\mathcal{T}} q_h - f_z\|_{L^2(\Omega_z)}^2 \right)^{1/2} \right) \\ & + c_{16} \|h_{\mathcal{T}}^2 \nabla_{\mathcal{T}}(f + \operatorname{div} p_h)\|_{2,\Omega} + c_{16} \|h_{\mathcal{E}}^{1/2} (g - p_h \cdot n)\|_{2,\Gamma_N}. \end{aligned}$$

Proof. Lemma 5.1 provides (5.5), and we may and will assume without loss of generality that $\beta = 0$ on Γ_N . An integration by parts and (6.2)–(6.3) show for the \mathcal{T} -piecewise integral mean $\alpha_{\mathcal{T}} \in \mathcal{L}^0(\mathcal{T})$ of $\alpha \in H_D^1(\Omega)$ that

$$(6.5) \quad \begin{aligned} \int_{\Omega} (p - p_h) \cdot \nabla \alpha dx &= \int_{\Omega} (f + \operatorname{div} p_h) \alpha dx + \int_{\Gamma_N} (g - p_h \cdot n) \alpha dx \\ &= \int_{\Omega} (f + \operatorname{div} p_h) (\alpha - \alpha_{\mathcal{T}}) dx + \int_{\Gamma_N} (g - p_h \cdot n) (\alpha - \alpha_{\mathcal{T}}) dx. \end{aligned}$$

The second last term is estimated with an elementwise Poincaré inequality while the last term in (6.5) involves a trace theorem [BS, CF1, Cl], namely

$$(6.6) \quad \|a\|_{2,E} \lesssim h_E^{-1/2} \|a\|_{2,T_E} + h_E^{1/2} \|\nabla a\|_{2,T_E}$$

for $a = \alpha - \alpha_{\mathcal{T}} \in H^1(T_E)$ on the triangle $T_E \in \mathcal{T}$ and the edge $E \in \mathcal{E}_N$, $E \subset \partial T_E$. With a second application of Poincaré’s inequality, (6.6), and Cauchy’s inequality

we show that

$$\begin{aligned}
 (6.7) \quad \int_{\Omega} (p - p_h) \cdot \nabla \alpha \, dx &\lesssim \|h_{\mathcal{T}}(f + \operatorname{div} p_h)\|_2 \|\nabla \alpha\|_2 \\
 &\quad + \|h_{\mathcal{E}}^{1/2}(g - p_h \cdot n)\|_{2, \Gamma_N} \|h_{\mathcal{E}}^{-1/2}(\alpha - \alpha_{\mathcal{T}})\|_{2, \Gamma_N} \\
 &\lesssim \|\nabla \alpha\|_2 \left(\|h_{\mathcal{T}}^2 \nabla_{\mathcal{T}}(f + \operatorname{div} p_h)\|_2 + \|h_{\mathcal{E}}^{1/2}(g - p_h \cdot n)\|_{2, \Gamma_N} \right).
 \end{aligned}$$

The second contribution on the right-hand side of (5.5) is analysed with Theorem 3.1, where, as in the proof of Theorem 5.1, we interchange components and the role of the boundary conditions. As already employed in [C1, C2], $\operatorname{Curl} w_h \in H(\operatorname{div}, \Omega)$ for all $w_h \in \mathcal{S}_N^1(\mathcal{T})$. Moreover, $\operatorname{Curl} w_h \cdot n = \partial w_h / \partial s = 0$ on Γ_N . Hence, (6.1) and an integration by parts for p yield (5.7) because of (4.2). Arguing as in the proof of Theorem 5.1, we deduce for arbitrary $q_h \in \mathcal{S}^k(\mathcal{T})^2$ that

$$\begin{aligned}
 (6.8) \quad \int_{\Omega} (p - p_h) \cdot \operatorname{Curl} \beta \, dx &\leq \|\nabla \beta\|_2 \left(c_2 \|p_h - q_h\|_2 + c_5 \|h_{\mathcal{E}}^{1/2}(\partial u_D / \partial s - q_h \cdot t)\|_{2, \Gamma_D} \right. \\
 &\quad \left. + c_3 \left(\sum_{z \in \mathcal{K}} h_z^2 \min_{f_z \in \mathbb{R}} \|\operatorname{curl}_{\mathcal{T}} q_h - f_z\|_{2, \Omega_z}^2 \right)^{1/2} \right).
 \end{aligned}$$

The remaining details are analogous to the proof of Theorem 5.1 and hence are omitted. □

The precise version of (1.5) for lowest order mixed finite element methods is summarised as follows.

Corollary 6.1. *Suppose that the discrete flux p_h satisfies $\operatorname{curl}_{\mathcal{T}} p_h = 0$, $\operatorname{div}_{\mathcal{T}} p_h \in \mathcal{L}^0(\mathcal{T})$ and $p_h \cdot n \in \mathcal{L}^0(\mathcal{E}_N)$. Then,*

$$\begin{aligned}
 (6.9) \quad \|p - p_h\|_{L^2(\Omega)} &\leq \min_{q_h \in \mathcal{S}_D^1(\mathcal{T}, u_D)} c_{17} \|p_h - q_h\|_{L^2(\Omega)} \\
 &\quad + c_{18} \left(\|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}}^2 u_D / \partial s^2\|_{L^2(\Gamma_D)} + \|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}} g / \partial s\|_{L^2(\Gamma_N)} + \|h_{\mathcal{T}}^2 \nabla_{\mathcal{T}} f\|_{L^2(\Omega)} \right).
 \end{aligned}$$

Proof. Combine Theorem 6.1 and Lemmas 4.1 and 5.2, and use an inverse estimate to prove

$$\begin{aligned}
 \sum_{z \in \mathcal{K}} h_z^2 \|\operatorname{curl}_{\mathcal{T}} q_h\|_{2, \Omega_z}^2 &= \sum_{z \in \mathcal{K}} h_z^2 \|\operatorname{curl}_{\mathcal{T}}(q_h - p_h)\|_{2, \Omega_z}^2 \\
 &\lesssim \sum_{z \in \mathcal{K}} \|q_h - p_h\|_{2, \Omega_z}^2 \lesssim \|q_h - p_h\|_2. \quad \square
 \end{aligned}$$

Remark 6.2. The assumptions in Corollary 6.1 are satisfied for lowest order Raviart–Thomas and Brezzi–Douglas–Fortin–Marini finite elements.

Example 6.1. Assume the conditions of Lemma 5.2 on the data u_D . We define an operator $\mathcal{A} : L^2(\Omega)^2 \rightarrow \mathcal{S}_D^1(\mathcal{T}, u_D)$ by (4.15) and $p_z := \int_{\omega_z} p \, dx$ for $z \in \mathcal{N} \setminus \Gamma_D$. In case $z \in \mathcal{N} \cap \Gamma_D$ we consider 2×2 systems

$$(6.10a) \quad t_{E_1} \cdot p_z = \partial u_D|_{E_1} / \partial s(z) \quad \text{and} \quad t_{E_2} \cdot p_z = \partial u_D|_{E_2} / \partial s(z)$$

(cf. (4.16a) and notation from Example 4.1), resp., as an analog to (4.16b),

$$(6.10b) \quad t_{E_1} \cdot p_z = \partial u_D|_{E_1} / \partial s(z) \quad \text{and} \quad n_{E_1} \cdot p_z = \int_{\omega_z} n_{E_1} \cdot p \, dx. \quad \square$$

A local version follows from Theorem 3.2 and concludes this section on mixed finite element methods.

Corollary 6.2. *Under the conditions of Theorem 6.1 and Corollary 6.1 we have*

$$\begin{aligned} \|p - p_h\|_{L^2(\Omega)} &\leq c_{19} \left(\sum_{E \in \mathcal{E}} \min_{q_E \in \mathcal{S}^1(\mathcal{T}_E)^2} \|p_h - q_E\|_{L^2(\omega_E)}^2 + h_E \|u'_{D,h} - q_E \cdot t\|_{L^2(\Gamma_D)}^2 \right)^{1/2} \\ &+ c_{18} (\|h_\mathcal{E}^{3/2} \partial_\mathcal{E}^2 u_D / \partial s^2\|_{L^2(\Gamma_D)} + \|h_\mathcal{E}^{3/2} \partial_\mathcal{E} g / \partial s\|_{L^2(\Gamma_N)} + \|h_\mathcal{T}^2 \nabla_\mathcal{T} f\|_{L^2(\Omega)}). \quad \square \end{aligned}$$

Remark 6.3. The results of this section could be generalised to three space dimensions. Details are omitted for brevity.

Remark 6.4. For related residual based a posteriori error estimates we refer to [A, BV, C1, C2, HW].

7. NUMERICAL EXPERIMENTS

The theoretical results of this paper are supported by numerical experiments. In this section, we report on two examples of the problem (4.1)–(4.3) on uniform, adapted, and perturbed meshes for conforming, nonconforming, and mixed finite element methods.

Example 7.1. Let $f := 0$ on the L-shaped domain $\Omega := (-1, 1)^2 \setminus [0, 1] \times [-1, 0]$, $u_D := 0$ on the Dirichlet boundary $\Gamma_D := \{0\} \times [-1, 0] \cup [0, 1] \times \{0\}$, and on the Neumann boundary $\Gamma_N := \partial\Omega \setminus \Gamma_D$,

$$g(r, \varphi) := 2/3 r^{-1/3} (-\sin(\varphi/3), \cos(\varphi/3)) \cdot n$$

using polar coordinates (r, φ) . The exact solution $u(r, \varphi) := r^{2/3} \sin(2\varphi/3)$ of (4.1)–(4.3) has a typical corner singularity at the origin. In this example, the right-hand sides are smooth, but the solution is not. The coarsest triangulation \mathcal{T}_0 consists of three squares halved by diagonals parallel to the vector $(1, 1)$ (cf. Figure 1).

Example 7.2. Let $f := -\Delta u$ for the function

$$u(x, y) := x(1 - x)y(1 - y) \arctan(60(r - 1)), \quad r^2 := (x - 1.25)^2 + (y + 0.25)^2$$

on the unit square $\Omega := (0, 1)^2$ and set $u_D := 0$ on the entire boundary $\Gamma_D := \partial\Omega$ ($\Gamma_N = \emptyset$). The solution u to (4.1)–(4.3) is H^2 -regular but f (although theoretically smooth) has huge gradients on the circle with radius 1 around $(1.25, -0.25)$. The coarsest triangulation \mathcal{T}_0 consists of four squares halved by diagonals parallel to the vector $(1, 1)$ (cf. Figure 3).

The following adaptive algorithm generates all the sequences of meshes $\mathcal{T}_0, \mathcal{T}_1, \mathcal{T}_2, \dots$ in this paper which are uniform for $\Theta = 0$ or adapted for $\Theta = 1/2$ in (7.2). Since the resulting meshes might show local symmetries, we considered meshes that are either unperturbed (relative to \mathcal{T}_0) for $\vartheta = 0$ and randomly perturbed for $\vartheta = 1$ in step (e). The implementation was performed in Matlab in the spirit of [ACF] with a direct solution of linear systems of equations. For details on the red-blue-green refinements we refer to [V].

Algorithm (A_{Θ}^{ϑ}) .

- (a) Start with a coarse mesh \mathcal{T}_0 , $k = 0$.
- (b) Compute the discrete solution p_h on the actual mesh \mathcal{T}_k .
- (c) Compute error indicators

$$(7.1) \quad \eta_{Z,T} := \|p_h - \mathcal{A}p_h\|_{L^2(T)}$$

for all $T \in \mathcal{T}_k$ and plot energy error $e_N := \|p - p_h\|_{L^2(\Omega)}$ and its estimator $\eta_N^2 := \sum_{T \in \mathcal{T}} \eta_{Z,T}^2$ versus the degree of freedom N of the triangulation \mathcal{T}_k .

- (d) Mark the element T for *red*-refinement provided

$$(7.2) \quad \eta_{Z,T} \geq \Theta \max_{T' \in \mathcal{T}_k} \eta_{Z,T'}.$$

- (e) Mark further elements (*red-blue-green* refinement) to avoid hanging nodes. Generate a new triangulation $\tilde{\mathcal{T}}_{k+1}$ using edge-midpoints if $\vartheta = 0$ and points on the edges at a random distance at most $0.3 h_E$ from the edge-midpoints if $\vartheta = 1$. Perturbe the nodes $z \in \mathcal{N}_{k+1}$ of the mesh $\tilde{\mathcal{T}}_{k+1}$ at random with values taken uniformly from a ball around z of radius $\vartheta 2^{-k}/15$. Correct boundary nodes by orthogonal projection onto that boundary piece they are expected such that $\Omega, \Gamma_D, \Gamma_N$ are matched by the resulting mesh \mathcal{T}_{k+1} exactly. Update k and go to (b).

7.1. Results for conforming finite element methods. In the conforming finite element scheme, we use operator \mathcal{A} from Example 4.1 in (7.1) of Algorithm (A_{Θ}^{ϑ}) and report on results obtained for $(\Theta, \vartheta) = (0, 0)$ (uniform), $(1/2, 0)$ (adaptive), and $(1/2, 1)$ (adaptive, perturbed).

Some meshes obtained for Example 7.1 are shown in Figure 1 and illustrate a high automatic mesh-refinement of the adapted meshes towards the origin, which is expected to improve the convergence rate of $2/3$ possibly to the optimal value 1. The result of the perturbation in step (e) of Algorithm $(A_{1/2}^1)$ is seen in the right half of

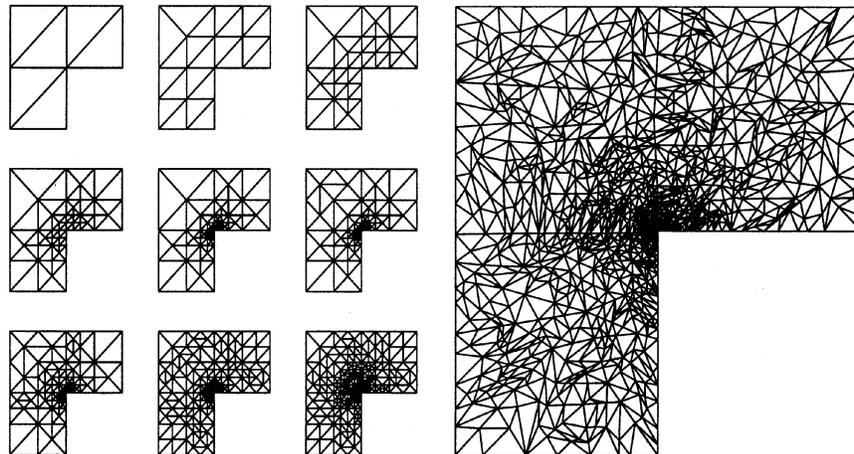


FIGURE 1. Adaptively refined meshes \mathcal{T}_0 (left upper) to \mathcal{T}_8 (right lower) (left) and perturbed triangulation \mathcal{T}_{16} with 1157 free nodes (right) in Example 7.1 for the conforming finite element scheme.

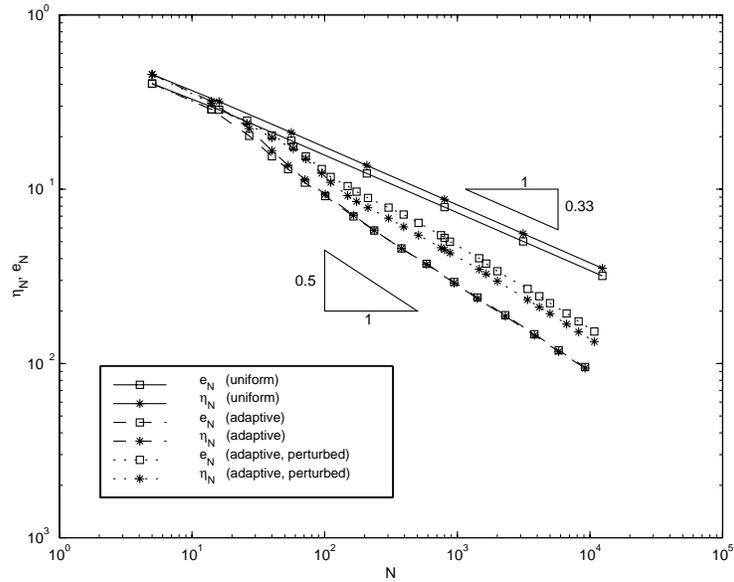


FIGURE 2. Error and error estimator for uniform, adaptive, and perturbed adaptive mesh-refinement in the conforming finite element scheme in Example 7.1.

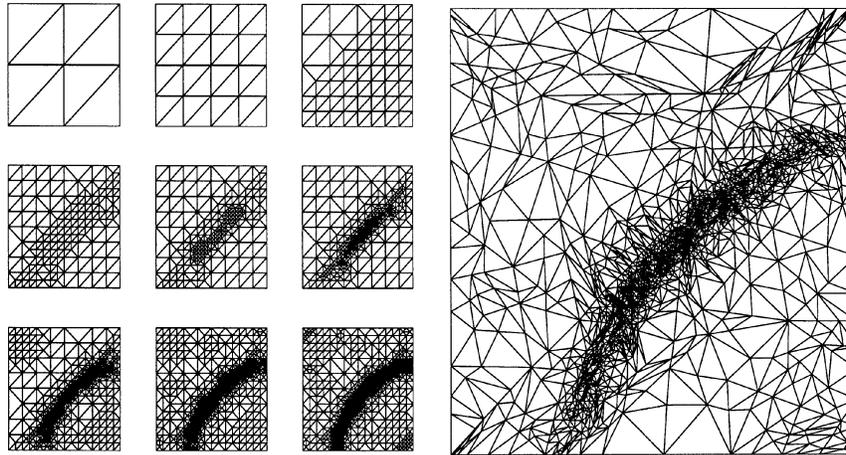


FIGURE 3. Adaptively refined meshes \mathcal{T}_0 (left upper) to \mathcal{T}_8 (right lower) (left) and perturbed triangulation \mathcal{T}_{12} with 1909 free nodes (right) for the conforming finite element scheme in Example 7.2.

Figure 1. We believe that the meshes generated by Algorithm $(A_{1/2}^1)$ have less local symmetry than that by $(A_{1/2}^0)$. According to local extrapolation, symmetry could cause superconvergence phenomena. To check the practical convergence behaviour, we plotted in Figure 2 for each mesh \mathcal{T}_k an entry (N, e_N) and (N, η_N) . A log-scaling on both axes allows a slope $-\alpha$ of a straight line in the plot that connects two

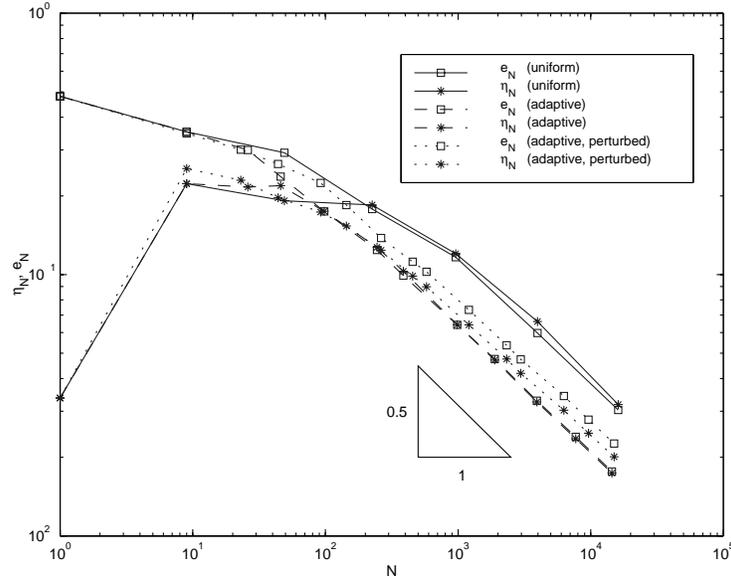


FIGURE 4. Error and error estimator for uniform, adaptive, and perturbed adaptive mesh-refinement in the conforming finite element scheme in Example 7.2.

subsequent entries for a series of meshes $\mathcal{T}_0, \mathcal{T}_1, \mathcal{T}_2, \dots$ generated by Algorithm (A_Θ^ϑ) , to be interpreted as an experimental convergence rate 2α (owing to $N \propto h^{-2}$ in two dimensions). We observe experimental convergence rates $2/3$, resp. 1 , for uniform, resp. adapted, meshes (generated by Algorithm (A_Θ^ϑ) for $\Theta = 0$, resp. $\Theta = 1/2$). Furthermore, even for coarse meshes, η_N appears to be a very good approximation to e_N ; corresponding entries almost coincide for $(\Theta, \vartheta) = (1/2, 0)$. If these meshes are perturbed (cf. Figure 1), the quotient η_N/e_N is almost a constant very close to 1. Numerical checks with different numerical quadrature rules (used to evaluate e_N) convinced us that, in general, η_N behaves not asymptotically exact in practise but is very accurate.

In Example 7.2 we obtained meshes and experimental convergence rates displayed in Figures 3 and 4. Although u belongs to $H^2(\Omega)$ and we expect linear convergence, u has huge second order derivatives along a circular arc where f is steep. We observe high refinements in the adapted meshes towards this arc. In this example the preasymptotic range is very large, an experimental convergence rate 1 can be observed only for $N \geq 300$ for all refinement strategies. In this regime the estimator η_Z appears as a good approximate for e_N and the entries (N, e_N) and (N, η_N) almost coincide for $(\Theta, \vartheta) = (0, 0), (1/2, 0)$. This is not the case for $(\Theta, \vartheta) = (1/2, 1)$, but the quotient $\eta_N/e_N \approx 0.9$ is still close to 1.

7.2. Results for nonconforming finite element methods. The operator \mathcal{A} from Example 5.1 serves in (7.1) to define $\eta_{Z,T}$ for first order Crouzeix–Raviart finite elements (cf., e.g., [BS, Ci]) in Algorithm (A_Θ^ϑ) . The generated meshes look similar to those shown in Figure 1, resp. Figure 3, and therefore are not displayed in this paper. The experimental convergence rates for Example 7.1, resp. 7.2, are illustrated in Figure 5, resp. Figure 6. The overall picture appears similarly to

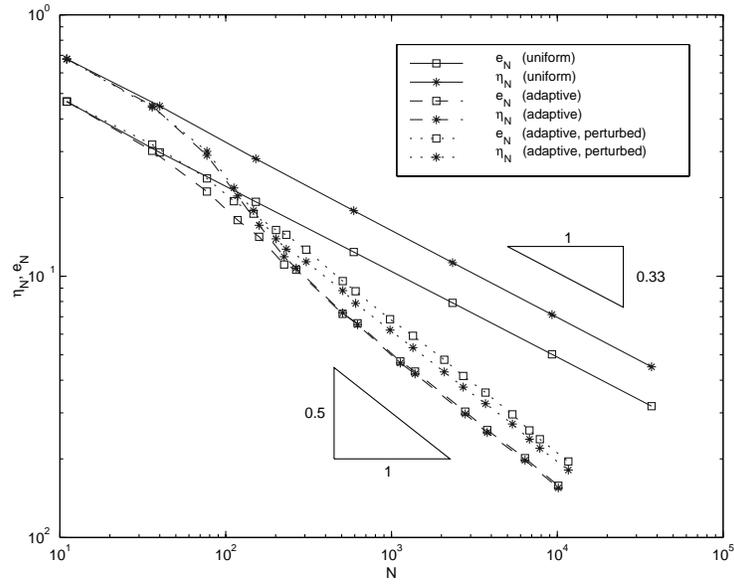


FIGURE 5. Error and error estimator for uniform, adaptive, and perturbed adaptive mesh-refinement in the nonconforming finite element scheme in Example 7.1.

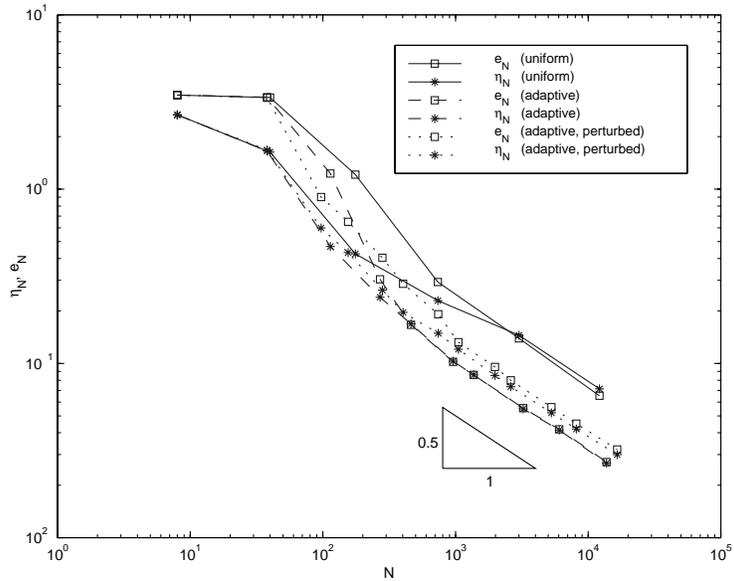


FIGURE 6. Error and error estimator for uniform, adaptive, and perturbed adaptive mesh-refinement in the nonconforming finite element scheme in Example 7.2.

the above discussions and we draw the same conclusions. For uniform meshes, the quotient η_N/e_N is nearly constant 1.2 but significantly larger than 1.1 in Figure 2.

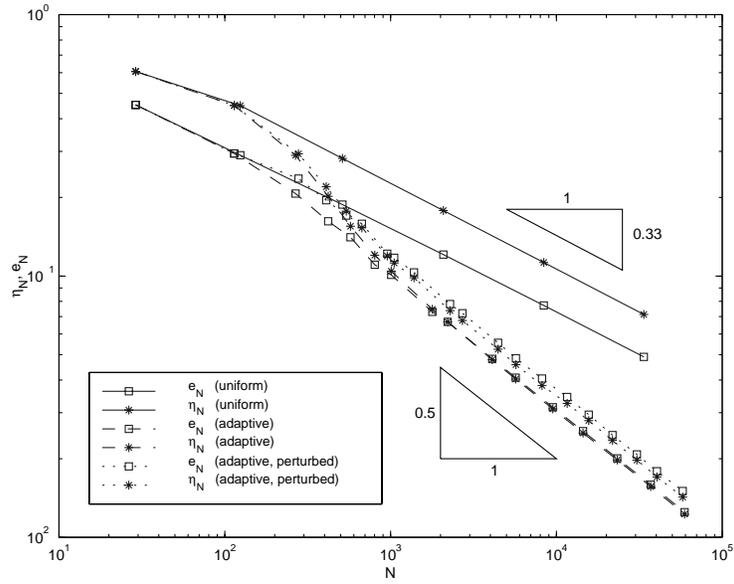


FIGURE 7. Error and error estimator for uniform, adaptive, and perturbed adaptive mesh-refinement in the mixed finite element scheme in Example 7.1.

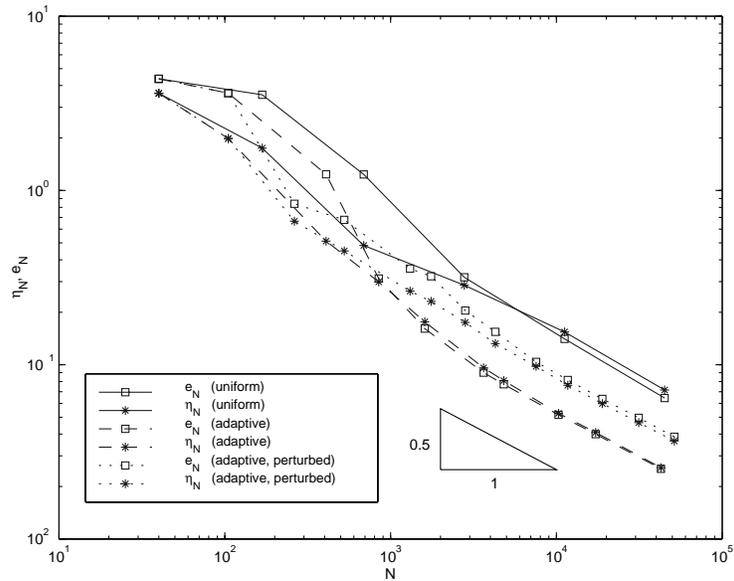


FIGURE 8. Error and error estimator for uniform, adaptive, and perturbed adaptive mesh-refinement in the mixed finite element scheme in Example 7.2.

7.3. Results for mixed finite element methods. For the Raviart–Thomas finite element method (cf. [B, BF, BS]) we use the operator \mathcal{A} from Example 6.1 to define $\eta_{Z,T}$. The adapted meshes look similar to those shown in Figures 1 and

3 and therefore are not displayed in this paper. Figures 7 and 8 display the error and the estimator for the mixed finite element scheme in Examples 7.1 and 7.2 obtained from Algorithm (A_{Θ}^{ϑ}) and $(\Theta, \vartheta) = (0, 0), (1/2, 0), (1/2, 1)$. We obtain the same experimental convergence rates as in the previous methods.

7.4. Remarks.

- (i) Our overall experience with Algorithm (A_{Θ}^{ϑ}) and other (e.g., residual-based) adaptive algorithms supports that all such adaptive algorithms yield a considerable convergence improvement.
- (ii) Although asymptotic exactness of η_N is not observed, the reliability constant c_1 in (1.5) and the efficiency constant are experimentally very close to 1 since η_N is a very good approximation to e_N for very fine meshes (i.e., when h.o.t. is negligible, say, for $N \geq 100$).
- (iii) Note that the efficiency constant is not known to be one as η_N (based on the averaging operator \mathcal{A}) is different from

$$(7.3) \quad \eta_Z := \min_{q_h \in \mathcal{S}_N^1(\mathcal{T}, g)} \|p_h - q_h\|_{L^2(\Omega)}.$$

For conforming linear triangular finite elements, the efficiency of η_N with an averaging operator follows from [R1, R2].

- (iv) Instead of the averaging operator \mathcal{A} , we tested the error estimator $\eta_Z \leq \eta_N$ from (7.3) and found that sometimes the performance is poorer than that of η_N : In Figure 4, for instance, the results of η_Z are much smaller than those of $\eta_N \approx e_N$. The averaging technique suggested in [HSWW]; average over a domain of size $O(h \log(1/h))$, is expected to give values between η_N and e_N .
- (v) The error estimation in Example 7.1 is very accurate even for very coarse meshes. Hence the higher order terms do not seem to be important here although u is nonsmooth. This agrees with our theoretical prediction in (1.5) since f and u_D are zero and g is piecewise analytic. For a generic corner singularity of u , we expect

$$\min_{q_h \in \mathcal{S}_N^1(\mathcal{T}, g)} \|p - q_h\|_{L^2(\Omega)} \ll \min_{q_h \in \mathcal{L}^0(\mathcal{T})} \|p - q_h\|_{L^2(\Omega)} \leq \|p - p_h\|_{L^2(\Omega)}$$

and so the h.o.t. in (1.4) are not expected to be dominant even for coarse meshes.

- (vi) For coarse meshes in Example 7.2, higher order terms may cause the overall observation that η_N is much smaller than e_N . Assuming $c_1 \approx 1$ (which is seen for fine meshes, whence for negligible h.o.t.) nonsmooth data (∇f is large) indicate that $e_N \leq c_1 \eta_N + \text{h.o.t}$ cannot be improved to $e_N \leq c_1 \eta_N$.

Acknowledgments. The second author (S.B.) thankfully acknowledges the partial support by the German research foundation at the Graduiertenkolleg 357 “Effiziente Algorithmen und Mehrskalenmethoden”.

REFERENCES

- [ACF] J. ALBERTY, C. CARSTENSEN, S.A. FUNKEN: Remarks around 50 lines of Matlab: short finite element implementation. *Numer. Algorithms* **20** (1999) 117-137. CMP 2001:01
- [A] A. ALONSO: Error estimators for a mixed method. *Numer. Math.* **74** (1996), 385-395. MR **97g**:65212
- [BaR] I. BABUŠKA, W.C. RHEINBOLDT: Error estimates for adaptive finite element computations. *SIAM J. Numer. Anal.* **15** (1978) 736-754. MR **58**:3400

- [Baetal] I. BABUŠKA, T. STROUBOULIS, C.S. UPADHYAY, S.K. GANGARAJ, K. COPPS: Validation of a posteriori error estimators by numerical approach. *Int. J. Numer. Meth. Engrg.* **37** (1994) 1073–1123. MR **95e**:65096
- [BC] 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.*, posted on February 4, 2002, PII S 0025-5718(02)01412-6 (to appear in print).
- [BeR] R. BECKER, R. RANNACHER: A feed-back approach to error control in finite element methods: basic analysis and examples. *East-West J. Numer. Math.*, **4**, No. 4 (1996) 237–264. MR **98m**:65185
- [B] D. BRAESS: *Finite Elements*. Cambridge University Press (1997). MR **98f**:65002
- [BV] D. BRAESS, R. VERFÜRTH: A posteriori error estimators for the Raviart-Thomas element. *SIAM J. Numer. Anal.* **33** (1996) 2431–2444. MR **97m**:65201
- [BS] S.C. BRENNER, L.R. SCOTT: *The mathematical theory of finite element methods*. Texts Appl. Math. 15, Springer, New-York (1994). MR **95f**:65001
- [BF] F. BREZZI, M. FORTIN: *Mixed and hybrid finite element methods*. Springer-Verlag (1991). MR **92d**:65187
- [C1] C. CARSTENSEN: A posteriori error estimate for the mixed finite element method. *Math. Comp.* **66** (1997) 465–476. MR **98a**:65162
- [C2] C. CARSTENSEN: Quasi interpolation and a posteriori error analysis in finite element method. *M2AN Math. Model Numer. Anal.* **33** (1999) 1187–1202. MR **2001a**:65135
- [CF1] C. CARSTENSEN, S.A. FUNKEN: Constants in Clément-interpolation error and residual based a posteriori error estimates in Finite Element Methods. *East-West J. Numer. Math.* **8** (3) (2000) 153–175. CMP 2001:07
- [CF2] 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.
- [CF3] C. CARSTENSEN, S.A. FUNKEN: Averaging technique for FE—a posteriori error control in elasticity. Part I: conforming FEM, *Comp. Meth. Appl. Mech. Engrg.* **190** (2001) 2483–2498. MR **2002a**:74114
- [CF4] C. CARSTENSEN, S.A. FUNKEN: Averaging technique for FE—a posteriori error control in elasticity. Part II: λ -independent estimates, *Comput. Methods Appl. Mech. Engrg.*, **190** (2001), 4663–4675.
- [CF5] C. CARSTENSEN, S.A. FUNKEN: Averaging technique for FE—a posteriori error control in elasticity. Part III: Locking-free nonconforming FEM, *Comput. Methods Appl. Mech. Engrg.* (to appear).
- [CV] C. CARSTENSEN, R. VERFÜRTH: Edge residuals dominate a posteriori error estimates for low order finite element methods. *SIAM J. Numer. Anal.* **36** No. 5 (1999) 1571–1587. MR **2000g**:65115
- [Cl] P. CLÉMENT: Approximation by finite element functions using local regularization. *RAIRO Sér. Rouge Anal. Numér.* **R-2** (1975) 77–84. MR **53**:4569
- [Ci] P.G. CIARLET: *The finite element method for elliptic problems*. North-Holland, Amsterdam (1978). MR **58**:25001
- [DDPV] E. DARI, R. DURAN, C. PADRA, AND V. VAMPA: A posteriori error estimators for nonconforming finite element methods. *RAIRO Model. Math. Anal. Numer.* **30** (1996) 385–400. MR **97f**:65066
- [DMR] R. DURAN, M.A. MUSCHIETTI, R. RODRIGUEZ: On the asymptotic exactness of error estimators for linear triangular elements. *Numer. Math.* **59** (1991) 107–127. MR **92b**:65086
- [EEHJ] K. ERIKSSON, D. ESTEP, P. HANSBO, C. JOHNSON: Introduction to adaptive methods for differential equations. *Acta Numer.* (1995) 105–158. MR **96k**:65057
- [GR] V. GIRAULT, P.A. RAVIART: *Finite Element Methods for Navier-Stokes Equations*. Springer, Berlin (1986). MR **88b**:65129
- [HSWW] W. HOFFMANN, A.H. SCHATZ, L.B. WAHLBIN, G. WITTUM: Asymptotically exact a posteriori error estimators for the pointwise gradient error on each element in irregular meshes. Part I: A smooth problem and globally quasi-uniform meshes. *Math. Comp.* **70** (2001) 897–909.

- [HW] R.H.W. HOPPE, B. WOHLMUTH: Element-oriented and edge-oriented local error estimators for nonconforming finite element methods. *RAIRO Model. Math. Anal. Numer.* **30** (1996) 237–263. MR **97e**:65124
- [KS] G. KANSCHAT, F.-T. SUTTMEIER: A posteriori error estimates for nonconforming finite element schemes. *Calcolo* **36**, No.3 (1999) 129–141. MR **2000k**:65208
- [R1] R. RODRIGUEZ: Some remarks on Zienkiewicz–Zhu estimator. *Numer. Methods Partial Differential Equations* **10** (1994) 625–635. MR **95de**: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 Jyväskylä, 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*. Wiley-Teubner (1996).
- [ZZ] O.C. ZIENKIEWICZ, J.Z. ZHU: A simple error estimator and adaptive procedure for practical engineering analysis. *Int. J. Numer. Methods 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`

MATHEMATISCHES SEMINAR, CHRISTIAN-ALBRECHTS-UNIVERSITÄT ZU KIEL LUDEWIG-MEYN-STR. 4, D-24098 KIEL, FRG.

E-mail address: `sba@numerik.uni-kiel.de`