A Synthesis of A Posteriori Error Estimation Techniques for Conforming, Non-Conforming and Discontinuous Galerkin Finite Element Methods

Mark Ainsworth

ABSTRACT. A posteriori error estimation for conforming, non-conforming and discontinuous finite element schemes are discussed within a single framework. By dealing with three ostensibly different schemes under the same umbrella, the same common underlying principles at work in each case are highlighted leading to a clearer understanding of the issues involved. The ideas are presented in the context of piecewise affine finite element approximation of a second-order elliptic problem. It is found that the framework leads to three different known a posteriori error estimators: the equilibrated residual method in the case of conforming Galerkin FEM; the estimator of Ainsworth [**3**] in the case of the Crouzeix-Raviart scheme, and a new estimator [**1**] recently derived in case of discontinuous Galerkin approximation. In all cases one has computable upper bounds on the error measured in the energy norm and corresponding local lower bounds showing the efficiency of the schemes.

1. Introduction

A posteriori error estimation for standard conforming finite element approximation has reached a degree of maturity as witnessed by the results outlined in the books [**4, 6, 7, 15, 21**]. However, the situation regarding non-standard finite element schemes such as *non-conforming* and the currently popular *discontinuous Galerkin* finite element schemes is in its infancy.

One of the first works to address a posteriori estimation for the non-conforming finite element scheme of Crouzeix and Raviart [**11**] was the important paper of Dari et. al. [**13**] who obtained two sided bounds on the error measured in the energy norm up to generic constants using a technique based on the Helmholtz decomposition. These ideas were later extended to non-conforming mixed finite element approximation of Stokes flow [**12**]. Hierarchical basis type estimators were

1991 *Mathematics Subject Classification.* 65N50. Secondary 65N15, 65N50, 76S05.

Key words and phrases. a posteriori error estimation, non-conforming finite element, discontinuous Galerkin method.

Support for this work from the Leverhulme Trust through a Leverhulme Trust Fellowship is gratefully acknowledged.

explored in [**16**], whilst [**10**] derived estimators based on gradient averaging techniques. These works all give an a posteriori error upper bound involving an unknown, generic coefficient and as such do not provide actual numerical bounds. An alternative approach to the a posteriori error estimation for non-conforming scheme was presented in [**3**] and shown to provide two-sided bounds on the error with, significantly, an upper bound that does not involve *any* generic constants, meaning that one has a guaranteed computable upper bound on the error measured in the energy norm. The basic idea proved quite general and was subsequently extended to non-conforming approximation of the Stokes equations by Dörfler and Ainsworth [**14**] and in [**2**] was extended to the non-conforming rotated $\mathbb{Q}_1$ element of Rannacher and Turek [**19**] where computable upper bounds were provided for the case of mapped bilinear quadrilateral elements.

The situation regarding the discontinuous finite element method is even less developed. An explicit residual type estimator for the error measured in the energy norm was derived by Becker et. al. [**8, 9**] while Rivière and Wheeler [**20**] derived residual estimators for the error in the L_2-norm. Subsequently, Karakashian and Pascal [**18**] discussed several different estimators for a discontinuous finite element scheme. Once again, it is found that all of the estimators involve generic unknown constants in the upper bound, so that one does not obtain actual numerical bounds on the error. Recently, the ideas used in [**3**] were extended to the analysis of the discontinuous Galerkin method to obtain an actual numerical upper bound on the error.

In the present work, we depart from the usual practice whereby specific schemes are analysed in isolation. Instead, we discuss a posteriori error estimation for conforming, non-conforming and discontinuous finite element schemes in the same framework. In dealing with three ostensibly different schemes under the same umbrella, we aim to highlight the same common underlying principles at work in each case which we hope will lead to a clearer and deeper understanding of the nature of a posteriori error estimation. The ideas are presented in the context of piecewise affine finite element approximation of a second-order elliptic problem. It is found that the underlying framework leads to three different known a posteriori error estimators: the equilibrated residual method [**4**] in the case of conforming Galerkin FEM; the estimator of [**3**] in case of the Crouzeix-Raviart scheme, and the estimator recently derived in [**1**] in case of discontinuous Galerkin approximation. In all cases one has computable upper bounds on the error measured in the energy norm and corresponding local lower bounds showing the efficiency of the schemes.

2. Preliminaries

2.1. Model Problem. Consider the following model problem

$$\left.\begin{array}{rcl} -\operatorname{\mathbf{div}}\boldsymbol{\sigma}(u) &=& f \\ \boldsymbol{\sigma}(u) - a\,\mathbf{grad}\,u &=& 0 \end{array}\right\} \text{ in } \Omega \tag{2.1}$$

subject to $u = 0$ on Γ_D and $\sigma_\nu(u) = \boldsymbol{\nu}\cdot\boldsymbol{\sigma}(u) = g$ on Γ_N, where Ω is a plane polygonal domain, the disjoint sets Γ_D and Γ_N form a partitioning of the boundary $\Gamma = \partial\Omega$ of the domain and a is strictly positive. For ease of exposition, it will be assumed that f and a are piecewise constant on sub-domains of Ω, and g is piecewise constant on Γ_N.

The standard variational formulation of the problem consists of seeking $u \in H^1_E(\Omega)$ such that

$$(a\,\mathbf{grad}\,u, \mathbf{grad}\,v) = (f, v) + \int_{\Gamma_N} gv\,\mathrm{d}s \quad \forall v \in H^1_E(\Omega) \tag{2.2}$$

where $H^1_E(\Omega)$ denotes the space $\{v \in H^1(\Omega) : v = 0 \text{ on } \Gamma_D\}$. Here, we use the notation $(\cdot,\cdot)_\omega$ to denote the integral inner product over a region ω, and omit the subscript in the case where ω is the physical domain Ω.

The space $H^1_E(\Omega)$ is the natural space over which to pose the problem. A conforming Galerkin discretisation of this problem would consist of introducing a finite dimensional subspace X_E of $H^1_E(\Omega)$ and seeking $u_X \in X_E$ such that

$$(a\,\mathbf{grad}\,u_X, \mathbf{grad}\,v) = (f, v) + \int_{\Gamma_N} gv\,\mathrm{d}s \quad \forall v \in X_E. \tag{2.3}$$

Although conforming approximation schemes of this type are used extensively, there are certain advantages in weakening the smoothness requirements on the finite dimensional space and allowing the use of spaces $X_{\mathcal{P}}$ that are non-conforming in the sense that they are *not* subspaces of the natural space $H^1_E(\Omega)$ for the problem.

The non-conforming spaces will be constructed on families of partitions $\{\mathcal{P}\}$ of the domain Ω into the union of non-overlapping, shape regular triangular elements such that the non-empty intersection of a distinct pair of elements is a single common node or single common edge. The family of partitions is assumed to be locally quasi-uniform in the sense that the ratio of the diameters of any pair of neighbouring elements is uniformly bounded above and below over the whole family. The set of all edges of the elements is denoted by $\mathcal{E}$, which we partition into subsets $\mathcal{E}_D$, $\mathcal{E}_N$ and $\mathcal{E}_I$ consisting of edges lying on the Dirichlet boundary Γ_D, the Neumann boundary Γ_N and the interior edges respectively.

2.2. Broken Variational Formulation. We wish to develop a variational formulation of problem (2.2) over a suitable *broken space* $H^m(\mathcal{P})$, $m > 1$ defined by

$$H^m(\mathcal{P}) = \{v : \Omega \to \mathbb{R}, \quad v_{|K} \in H^m(K) \quad \forall K \in \mathcal{P}\}. \tag{2.4}$$

The space is broken in the sense that it consists of functions that are piecewise H^1-regular over the elements of the partition. However, the lack of any formal inter-element continuity requirements or application of essential boundary conditions on functions belong to the space $H^1(\mathcal{P})$ means that the variational formulation must be modified appropriately to ensure that such conditions are applied in a weak or variational sense.

Before proceeding to the variational formulation, it will be useful to introduce some notation and conventions to describe jumps and averages of functions associated with the broken spaces across edges. For each element $K \in \mathcal{P}$, let $\mu_K : \partial K \to \{+1, -1\}$ denote a function that is piecewise constant on the edges of element K and chosen such that $\mu_K + \mu_{K'} = 0$ on $\partial K \cap \partial K'$ and $\mu_K = 1$ on $\partial K \cap \partial\Omega$. For $v \in H^1(\mathcal{P})$, we define the jump and average value of v on the edges $\mathcal{E}$ by

$$[v] = \begin{cases} \mu_K v_K + \mu_{K'} v_{K'} & \text{on } \gamma = \partial K \cap \partial K' \\ v_K & \text{on } \gamma = \partial K \cap \Gamma_D \end{cases} \tag{2.5}$$

and

$$\langle v \rangle = \begin{cases} \dfrac{1}{2}(v_K + v_{K'}) & \text{on } \gamma = \partial K \cap \partial K' \\ \qquad\quad v_K & \text{on } \gamma = \partial K \cap \Gamma_D. \end{cases} \tag{2.6}$$

The sign functions $\{\mu_K\}$ may be used to define a unique unit normal vector $\boldsymbol{\nu}$ on any given edge $\gamma \in \mathcal{E}$ using the formula

$$\boldsymbol{\nu} = \mu_K \boldsymbol{\nu}_K, \quad \gamma \subset \partial K \tag{2.7}$$

where K is *any* element of which γ is an edge. The jump and average in the flux of a function $v \in H^m(\mathcal{P})$ on individual edges may then be defined for $m > 1$ by

$$[\sigma_\nu(v)] = [\boldsymbol{\nu} \cdot \boldsymbol{\sigma}(v)] \tag{2.8}$$

and

$$\langle \sigma_\nu(v) \rangle = \langle \boldsymbol{\nu} \cdot \boldsymbol{\sigma}(v) \rangle \,. \tag{2.9}$$

For a given positive constant κ, we define the bilinear form $\mathcal{B}_\mathcal{P} : H^m(\mathcal{P}) \times H^m(\mathcal{P}) \to \mathbb{R}$ for $m > 1$ by

$$\begin{aligned} \mathcal{B}_\mathcal{P}(v,w) = &\sum_{K\in\mathcal{P}} (a\,\mathbf{grad}_\mathcal{P}\, v, \mathbf{grad}_\mathcal{P}\, w)_K \\ &- \sum_{\gamma\in\mathcal{E}_I\cup\mathcal{E}_D} \int_\gamma \left(\langle \sigma_\nu(v)\rangle\,[w] + [v]\,\langle\sigma_\nu(w)\rangle\right) \mathrm{d}s + \sum_{\gamma\in\mathcal{E}_I\cup\mathcal{E}_D} \frac{\kappa}{h_\gamma}\int_\gamma [v]\,[w]\;\mathrm{d}s \end{aligned} \tag{2.10}$$

and the linear form $\mathcal{L}_\mathcal{P} : H^m(\mathcal{P}) \to \mathbb{R}$ by

$$\mathcal{L}_\mathcal{P}(w) = \sum_{K\in\mathcal{P}} (f, w_K)_K + \sum_{\gamma\in\mathcal{E}_N} \int_\gamma g w \,\mathrm{d}s \tag{2.11}$$

Here, an integrals of products of averages and jumps over an individual edge is interpreted as a suitable duality pairing.

Suppose that there exists a smooth function $U \in H^m(\mathcal{P})$ satisfying the variational problem

$$\mathcal{B}_\mathcal{P}(U,v) = \mathcal{L}_\mathcal{P}(v) \quad \forall v \in H^m(\mathcal{P}). \tag{2.12}$$

Then, integrating the volume term in $\mathcal{B}_\mathcal{P}$ by parts and rearranging, we arrive at the identity

$$\begin{aligned} \mathcal{L}_\mathcal{P}(v) - \mathcal{B}_\mathcal{P}(U,v) &= \sum_{K\in\mathcal{P}} (f + \mathbf{div}_\mathcal{P}\,\boldsymbol{\sigma}(U), v)_K \\ &+ \sum_{\gamma\in\mathcal{E}_D} \int_\gamma U\left(\boldsymbol{\nu}\cdot\boldsymbol{\sigma}(v) - c h_\gamma^{-1} v\right)\,\mathrm{d}s \\ &+ \sum_{\gamma\in\mathcal{E}_N} \int_\gamma (g - \boldsymbol{\nu}\cdot\boldsymbol{\sigma}(U)) v \,\mathrm{d}s \\ &+ \sum_{\gamma\in\mathcal{E}_I} \int_\gamma [U]\left(\langle\sigma_\nu(v)\rangle + \kappa h_\gamma^{-1}\,[v]\right) - [\sigma_\nu(U)]\,\langle v\rangle\;\mathrm{d}s \end{aligned}$$

for all $v \in H^m(\mathcal{P})$. Choosing a test function v supported on element K and applying a standard density argument lead to the conclusion that U satisfies the condition

$$-\,\mathbf{div}\,\boldsymbol{\sigma}(U) = f \text{ on } K$$

for all $K \in \mathcal{P}$. Likewise, choosing the test function v to be non-zero on a given edge γ and using density shows that U satisfies the boundary conditions

$$U = 0 \text{ on } \gamma \subset \Gamma_D; \quad \boldsymbol{\nu} \cdot \boldsymbol{\sigma}(U) = g \text{ on } \gamma \subset \Gamma_N$$

and the jump conditions

$$[U] = [\sigma_\nu(U)] = 0 \text{ on } \gamma \in \mathcal{E}_I.$$

We refer the reader to Arnold et. al. [**5**] for further details on the broken formulation of elliptic problems.

2.3. Three Finite Element Schemes. The conclusion to be drawn from the above arguments is that any function satisfying the variational equation (2.12) offers a consistent approximation of the solution u of the original variational problem (2.2). Equally well, by introducing a finite dimensional subspace $X_\mathcal{P} \subset H^m(\mathcal{P})$, we may construct an approximation of the function u as follows: find $U_\mathcal{P} \in X_\mathcal{P}$ such that

$$\mathcal{B}_\mathcal{P}(U_\mathcal{P}, v) = \mathcal{L}_\mathcal{P}(v) \quad \forall v \in X_\mathcal{P}. \tag{2.13}$$

By choosing the parameter κ sufficiently large, one can arrange that the only function $v \in X_\mathcal{P}$ satisfying $\mathcal{B}_\mathcal{P}(v, v) = 0$ is the trivial function and hence, in view of the fact that $X_\mathcal{P}$ is assumed finite dimensional, there is a unique solution of the equation.

We wish to consider three alternative schemes based on differing choices of the space $X_\mathcal{P}$. In all cases, the basic space is based on piecewise affine finite element approximation on the underlying partitioning $\mathcal{P}$ of the domain into triangular elements. The difference between the schemes lies in the conditions applied to the jumps $[v]$ on interior edges and at the Dirichlet boundary:

Discontinuous Galerkin FEM:

$$X_\mathcal{P} = X_{\mathsf{DG}} = \{v \in H^1(\mathcal{P}) : v_{|K} \in \mathbb{P}_1(K) \quad \forall v \in \mathcal{P}\}$$

Crouzeix-Raviart FEM:

$$X_\mathcal{P} = X_{\mathsf{CR}} = \{v \in X_{\mathsf{DG}} : \int_\gamma [v] \, \mathrm{d}s = 0 \quad \forall \gamma \in \mathcal{E}_I \cup \mathcal{E}_D\}$$

Conforming FEM:

$$X_\mathcal{P} = X_{\mathsf{CG}} = \{v \in X_{\mathsf{DG}} : [v]_\gamma = 0 \quad \forall \gamma \in \mathcal{E}_I \cup \mathcal{E}_D\}.$$

The space X_{DG} carries no constraints on the jumps while at the opposite extreme all jumps are required to vanish in the case of X_{CG} and the scheme is conforming in the sense that $X_{\mathsf{CG}} \subset H^1_E(\Omega)$. The Crouzeix-Raviart space X_{CR} lies in between these extremes

$$X_{\mathsf{CG}} \subset X_{\mathsf{CR}} \subset X_{\mathsf{DG}} \subset H^1(\mathcal{P}) \tag{2.14}$$

and is non-conforming in the sense that $X_{\mathsf{CG}} \not\subset H^1_E(\Omega)$ but average values of jumps are required to vanish. An equivalent characterisation of the space X_{CR} which is more useful for the purpose of practical implementation of the method is that the functions are continuous at the midpoints of interior edges and vanish at the midpoints of edges on the Dirichlet portion of the boundary.

The discrete variational problem (2.13) may be used in conjunction with any of these choices to yield an approximation $U_\mathcal{P} \in X_\mathcal{P}$. The issue that we wish to discuss in the present work is: *how can one obtain computable a posteriori estimates* *for the error* $u - U_{\mathcal{P}}$*?* Traditionally, the a posteriori error analysis of the different schemes has been carried out in isolation. By treating all three schemes within a single framework, we hope to demonstrate that the same basic guiding principles are applicable in all cases and we show that three existing a posteriori estimators all naturally arise from the application of the framework.

3. A Posteriori Error Analysis

The error $e = u - U_{\mathcal{P}}$ in the finite element approximation schemes described in the previous section need not necessarily belong to the natural space $H^1_E(\Omega)$ for the original variational problem (2.2). A natural idea is to split the error into components corresponding to that part of the error that does belong to the space $H^1_E(\Omega)$, and the remainder. Thus, we consider the projection $\phi \in H^1_E(\Omega)$ of the total error onto the conforming space $H^1_E(\Omega)$ defined by

$$(a\,\mathbf{grad}\,\phi, \mathbf{grad}\,v) = (a\,\mathbf{grad}_{\mathcal{P}}\,e, \mathbf{grad}\,v) \quad \forall v \in H^1_E(\Omega). \tag{3.1}$$

The function ϕ will be referred to as the *conforming error* and is well-defined by the above condition. As a matter of fact, one can even write down a boundary value problem for ϕ by making use of the original equation (2.2):

$$\begin{aligned}(a\,\mathbf{grad}\,\phi, \mathbf{grad}\,v) &= \\ &(f, v) + \int_{\Gamma_N} gv\,\mathrm{d}s - (a\,\mathbf{grad}_{\mathcal{P}}\,U_{\mathcal{P}}, \mathbf{grad}\,v) \quad \forall v \in H^1_E(\Omega).\end{aligned} \tag{3.2}$$

It is worth observing that the equation involves only known data and in particular is independent of the unknown true solution u. Thus, at least in principle, one could compute ϕ explicitly from this characterisation. Of course, this is not a viable practical proposition since it is equivalent to computing the solution of the original problem.

The remaining part of the error $\rho = e - \phi$ is non-conforming and orthogonal to the space $H^1_E(\Omega)$:

$$0 = (a\,\mathbf{grad}_{\mathcal{P}}\,\rho, \mathbf{grad}\,v) \quad \forall v \in H^1_E(\Omega). \tag{3.3}$$

We make use of this equation by splitting the bilinear form into contributions from individual elements, applying integration by parts and regrouping terms to obtain

$$0 = -\sum_{K\in\mathcal{P}} (\mathbf{div}_{\mathcal{P}}\,\boldsymbol{\sigma}(\rho), v)_K + \sum_{\gamma\in\mathcal{E}_I} \int_\gamma v\,[\sigma_\nu(\rho)]\,\mathrm{d}s + \sum_{\gamma\in\mathcal{E}_N} \int_\gamma v\boldsymbol{\nu}\cdot\boldsymbol{\sigma}(\rho)\,\mathrm{d}s \quad \forall v \in H^1_E(\Omega)$$

since the test function v vanishes on Γ_D. By choosing a test function v supported on a single element K, we arrive at the conclusion that $\boldsymbol{\sigma}(\rho)_{|K}$ is solenoidal and hence can be expressed as a rotation:

$$\boldsymbol{\sigma}(\rho)_{|K} = \mathbf{curl}\,\psi \quad \forall K \in \mathcal{P}$$

where $\psi \in H^1(\mathcal{P})$. By choosing a non-zero test function v on an interior edge and inserting the above form for $\boldsymbol{\sigma}(\rho)$, we conclude that the traces of the normal components of $\mathbf{curl}\,\psi$ are continuous across interior edges. Equally well, ψ has continuous traces across interior edges and it follows that in fact $\psi \in H^1(\Omega)$. Finally, choosing a test function supported on an edge lying on the Neumann portion of the boundary, we find that the tangential component of ψ must be constant on such edges. Summarising, we have deduced that $\boldsymbol{\sigma}(\rho)$ may be expressed as the rotation of a function ψ belonging to the space

$$\mathcal{H} = \{w \in H^1(\Omega) : \partial w/\partial s = 0 \text{ on } \Gamma_N\}. \tag{3.4}$$

Overall we have the splitting of the total error

$$\boldsymbol{\sigma}_{\mathcal{P}}(e) = \boldsymbol{\sigma}(\phi) + \mathbf{curl}\,\psi \tag{3.5}$$

where ϕ is defined above, and $\psi \in \mathcal{H}$ satisfies the variational equation

$$(a^{-1}\,\mathbf{curl}\,\psi, \mathbf{curl}\,w) = (a^{-1}\boldsymbol{\sigma}_{\mathcal{P}}(e), \mathbf{curl}\,w) = (\mathbf{grad}_{\mathcal{P}}\,e, \mathbf{curl}\,w) \quad \forall w \in \mathcal{H} \tag{3.6}$$

where we have exploited the fact that

$$(a^{-1}\boldsymbol{\sigma}(\phi), \mathbf{curl}\,w) = (\mathbf{grad}\,\phi, \mathbf{curl}\,w) = 0 \quad \forall w \in \mathcal{H}. \tag{3.7}$$

The astute reader will realise that in the case where the coefficient a is constant the splitting (3.5) is nothing more than the classical Helmholtz decomposition. Indeed, the usual approach to the derivation of the following result starts from the Helmholtz decomposition:

THEOREM 1. *Let $\mathcal{H}$ be defined as above. The error e may be decomposed in the form* (3.5) *where $\phi \in H^1_E(\Omega)$ satisfies* (3.1) *and $\psi \in \mathcal{H}$ satisfies* (3.6)*. Moreover, the splitting is orthogonal in the sense that the Pythagorean identity is valid:*

$$(a^{-1}\boldsymbol{\sigma}_{\mathcal{P}}(e), \boldsymbol{\sigma}_{\mathcal{P}}(e)) = (a^{-1}\boldsymbol{\sigma}(\phi), \boldsymbol{\sigma}(\phi)) + (a^{-1}\,\mathbf{curl}\,\psi, \mathbf{curl}\,\psi). \tag{3.8}$$

The proof of the fact that the decomposition (3.5) is orthogonal is seen by rewriting the cross-term as follows

$$(a^{-1}\boldsymbol{\sigma}(\phi), \mathbf{curl}\,\psi) = (\mathbf{grad}\,\phi, \boldsymbol{\sigma}_{\mathcal{P}}(\rho)) = (a\,\mathbf{grad}_{\mathcal{P}}\,\rho, \mathbf{grad}\,\phi)$$

and then observing that this vanishes thanks to (3.3).

3.1. Estimation of Non-conforming Error. The technique for estimating the non-conforming part of the error is based on the following identity:

$$(a^{-1}\,\mathbf{curl}\,\psi, \mathbf{curl}\,\psi) = \min_{u^* \in H^1_E(\Omega)} (a\,\mathbf{grad}_{\mathcal{P}}(u^* - U_{\mathcal{P}}), \mathbf{grad}_{\mathcal{P}}(u^* - U_{\mathcal{P}})). \tag{3.9}$$

The intuitive interpretation of this result lies in the notion that the non-conforming part of the error should, roughly speaking, be related to how 'close' the approximation $U_{\mathcal{P}}$ is to being a conforming approximation. The above estimate confirms that this is indeed the correct viewpoint since the term appearing in the minimisation measures the distance between $U_{\mathcal{P}}$ and the conforming subspace $H^1_E(\Omega)$.

The proof of the above estimate is simple. Let $u^* \in H^1_E(\Omega)$ be taken arbitrarily and use (3.7) to obtain

$$(a^{-1}\,\mathbf{curl}\,\psi, \mathbf{curl}\,w) = (\mathbf{grad}_{\mathcal{P}}(u^* - U_{\mathcal{P}}), \mathbf{curl}\,w).$$

Then, applying the Cauchy-Schwarz inequality we readily conclude that

$$(a^{-1}\,\mathbf{curl}\,\psi, \mathbf{curl}\,\psi) \le (a\,\mathbf{grad}_{\mathcal{P}}(u^* - U_{\mathcal{P}}), \mathbf{grad}_{\mathcal{P}}(u^* - U_{\mathcal{P}})).$$

The proof is completed by noting that equality holds when we choose $u^* = u - \phi \in H^1_E(\Omega)$.

The application of the above bound requires a suitable choice for the function $u^* \in H^1_E(\Omega)$. We construct u^* by smoothing the finite element approximation $U_{\mathcal{P}}$. More precisely, we take u^* to be a piecewise affine function on the partitioning $\mathcal{P}$

with values at the vertices of the grid obtained by averaging the approximation $U_{\mathcal{P}}$: That is to say, given a vertex $\boldsymbol{x}_n \notin \Gamma_D$, we take

$$u^*(\boldsymbol{x}_n) = \frac{1}{|\Omega_n|} \sum_{K \subset \Omega_n} U_{\mathcal{P}|K}(\boldsymbol{x}_n) \tag{3.10}$$

where Ω_n is the set of elements that have a vertex at the grid point $\boldsymbol{x}_n$. Of course, the value at a node $\boldsymbol{x}_n \in \Gamma_D$ must be taken as zero to ensure that u^* belongs to the space $H^1_E(\Omega)$. This process gives a computable upper bound on the non-conforming part of the error of the form

$$(a^{-1}\,\mathbf{curl}\,\psi, \mathbf{curl}\,\psi) \leq \sum_{K \in \mathcal{P}} \eta^2_{\mathsf{NC},K} \tag{3.11}$$

where

$$\eta^2_{\mathsf{NC},K} = (a\,\mathbf{grad}_{\mathcal{P}}(u^* - U_{\mathcal{P}}), \mathbf{grad}_{\mathcal{P}}(u^* - U_{\mathcal{P}}))_K. \tag{3.12}$$

Moreover, it is possible to show that this bound is efficient in the sense that there exists a positive constant C, independent of any mesh-size, such that the local lower bound holds

$$\eta^2_{\mathsf{NC},K} \leq C(a^{-1}\,\mathbf{curl}\,\psi, \mathbf{curl}\,\psi)_{\widetilde{K}}$$

where $\widetilde{K}$ is the patch of element composed of the element K and its neighbours.

It will be noted that throughout the above considerations no mention has been made as to exactly which of the finite element schemes and spaces we had in mind. The reason for this is that the *same process is valid in all cases*. Thus, in the case of the conforming space $U_{\mathcal{P}} \in X_{\mathsf{CG}} \subset H^1_E(\Omega)$, we find that u^* and $U_{\mathcal{P}}$ coincide and the estimator vanishes. This is to be expected since in this case, $e \in H^1_E(\Omega)$ and the data appearing in (3.6) vanishes thanks to (3.7), in turn meaning that $\mathbf{curl}\,\psi$ vanishes. It is only when one comes to the proof of the local lower bound that the precise nature of the finite element scheme plays a role and for details we refer to [**3**] in the case of the Crouzeix-Raviart scheme and [**1**] in the case of the discontinuous Galerkin finite element method.

3.2. Estimation of Conforming Error. In order to handle the conforming part of the error, we introduce a set of *flux functions* $\{g_K : K \in \mathcal{P}\}$ defined on the element boundaries $g_K : \partial K \to \mathbb{R}$. At this stage, the functions have yet to be determined but we shall require that they be (i) discontinuous piecewise polynomials on element edges and (ii) satisfy the identity

$$\sum_{K \in \mathcal{P}} \int_{\partial K} g_K v \,\mathrm{d}s = \int_{\Gamma_N} g v \,\mathrm{d}s \quad \forall v \in H^1_E(\Omega). \tag{3.13}$$

This condition is equivalent to the following conditions on the functions on individual edges $\gamma \in \mathcal{E}_I \cup \mathcal{E}_D$:

$$\begin{aligned} g_K + g_{K'} &= 0 \quad \text{on} \quad \gamma = \partial K \cap \partial K' \\ g_K &= g \quad \text{on} \quad \gamma = \partial K \cap \Gamma_N. \end{aligned} \tag{3.14}$$

There are no requirements on the values of g on the Dirichlet portion of the boundary since condition (3.13) is posed over the subspace $H^1_E(\Omega)$ consisting of functions with vanishing traces on Γ_D.

Now, inserting the condition (3.13) into the data of (3.2) and writing as a sum of contributions from individual elements, we obtain

$$(a\,\mathbf{grad}\,\phi, \mathbf{grad}\,v) = \sum_{K\in\mathcal{P}} \left\{ (f,v)_K + \int_{\partial K} g_K v \,\mathrm{d}s - (a\,\mathbf{grad}_{\mathcal{P}}\,U_{\mathcal{P}}, \mathbf{grad}\,v)_K \right\} \quad \forall v \in H^1_E(\Omega). \tag{3.15}$$

We would like to arrange that the contribution to this sum from a given element $K \in \mathcal{P}$ vanishes when the restriction of the test function v to the element is constant. This requirement imposes the following condition (iii) on the flux functions:

$$\int_{\partial K} g_K \chi_K \,\mathrm{d}s = (a\,\mathbf{grad}_{\mathcal{P}}\,U_{\mathcal{P}}, \mathbf{grad}\,\chi_K) - (f, \chi_K) \tag{3.16}$$

where χ_K is the characteristic function on element K. Of course, the first term appearing on the right hand side vanishes automatically but it will helpful for later purposes to retain the term.

The precise details of the construction of the flux functions differ depending on the actual scheme owing to the differing levels of inter-element continuity conditions on the test functions in the various formulations.

3.2.1. *Discontinuous Galerkin FEM.* The absence of any formal continuity requirements on the test functions for this method means that a simple closed form expression can be given for the flux function

$$g_{K|\gamma} = \begin{cases} \mu_K \left(\langle \sigma_\nu(U_{\mathcal{P}}) \rangle - \kappa h_\gamma^{-1} [U_{\mathcal{P}}] \right) & \text{on } \gamma \in \mathcal{E}_I \cup \mathcal{E}_D \\ g & \text{on } \gamma \in \mathcal{E}_N. \end{cases} \tag{3.17}$$

Conditions (i) and (ii) are automatically satisfied in view of the definition of the sign function μ_K and the equivalence of (ii) with (3.14). To verify that condition (iii) is satisfied, we proceed directly using the definitions as follows:

$$\begin{aligned}
&\int_{\partial K} g_K \chi_K \,\mathrm{d}s \\
&\quad= \sum_{\gamma \subset \partial K \cap \Gamma_N} \int_\gamma g_K \chi_K \,\mathrm{d}s + \sum_{\gamma \subset \partial K \setminus \Gamma_N} \int_\gamma \left(\langle \sigma_\nu(U_{\mathcal{P}}) \rangle - \kappa h_\gamma^{-1} [U_{\mathcal{P}}] \right) \mu_K \chi_K \\
&\quad= \int_{\partial K \cap \Gamma_N} g \chi_K \,\mathrm{d}s + \sum_{\gamma \subset \partial K \setminus \Gamma_N} \int_\gamma \left(\langle \sigma_\nu(U_{\mathcal{P}}) \rangle - \kappa h_\gamma^{-1} [U_{\mathcal{P}}] \right) [\chi_K] \\
&\quad= \mathcal{L}_{\mathcal{P}}(\chi_K) - (f, \chi_K) - \mathcal{B}_{\mathcal{P}}(U_{\mathcal{P}}, \chi_K) + (a\,\mathbf{grad}_{\mathcal{P}}\,U_{\mathcal{P}}, \mathbf{grad}_{\mathcal{P}}\,\chi_K) \\
&\quad= (a\,\mathbf{grad}_{\mathcal{P}}\,U_{\mathcal{P}}, \mathbf{grad}_{\mathcal{P}}\,\chi_K) - (f, \chi_K)
\end{aligned}$$

where we have used the fact that $[\chi_K] = \mu_K \chi_K$ and $\langle \sigma_\nu(\chi_K) \rangle = 0$ along with fact that $U_{\mathcal{P}}$ satisfies the variational formulation with (2.13). The key point here that allowed us to give a simple closed form for g_K was the fact that the characteristic function $\chi_K \in X_{\mathsf{DG}}$ is an admissible test function for the discontinuous Galerkin finite element method.

3.2.2. *Crouzeix-Raviart FEM.* In contrast to the discontinuous Galerkin scheme, the Crouzeix-Raviart scheme imposes a weak continuity condition on the test functions between neighbouring elements. This impacts the construction of the flux function since one no longer has the freedom to choose a test function supported on a single element. Nevertheless, the continuity conditions are sufficiently weak that it remains possible to give an explicit construction of the fluxes.

In order to give the expression for the fluxes, we shall need to introduce some notation. For a given edge $\gamma \in \mathcal{E}$, let $\theta_\gamma \in X_{\mathsf{CR}}$ denote the function whose value is unity at the midpoint of edge γ, but vanishes at all remaining edge midpoints. These conditions mean that θ_γ is well-defined and the set of all such functions is typically used in the actual finite element implementation. It is readily verified that the support, $\operatorname{supp}(\theta_\gamma)$, of the function θ_γ consists of the (pair of) element(s) of which γ is an edge and there holds

$$[\theta_\gamma] = 0 \text{ on } \gamma; \quad \theta_\gamma = 1 \text{ on } \gamma; \quad \sum_{\gamma \subset \partial K} \theta_\gamma = 1 \text{ on } K. \tag{3.18}$$

The flux $g_{K|\gamma}$ on an edge $\gamma \in \mathcal{E}_I \cup \mathcal{E}_N$ is taken to be piecewise constant with the value defined by

$$h_\gamma g_{K|\gamma} = (a\, \mathbf{grad}_{\mathcal{P}}\, U_{\mathcal{P}}, \mathbf{grad}_{\mathcal{P}}\, \theta_\gamma)_K - (f, \theta_\gamma)_K + \sum_{\gamma' \subset \partial K} \frac{\kappa}{h_{\gamma'}} \int_{\gamma'} [U_{\mathcal{P}}]\,[\theta_\gamma]\; \mathrm{d}s. \tag{3.19}$$

Condition (ii) is verified for an interior edge $\gamma \in \mathcal{E}_I$ by summing the above expression over the pair of elements K, K' forming the support of θ_γ to obtain

$$\begin{aligned} h_\gamma (g_K + g_{K'})_{|K} &= (a\, \mathbf{grad}_{\mathcal{P}}\, U_{\mathcal{P}}, \mathbf{grad}_{\mathcal{P}}\, \theta_\gamma) + \sum_{\gamma' \in \mathcal{E}_I \cup \mathcal{E}_D} \frac{\kappa}{h_{\gamma'}} \int_{\gamma'} [U_{\mathcal{P}}]\,[\theta_\gamma]\; \mathrm{d}s - (f, \theta_\gamma) \\ &= \mathcal{B}_{\mathcal{P}}(U_{\mathcal{P}}, \theta_\gamma) - \mathcal{L}_{\mathcal{P}}(\theta_\gamma) = 0 \end{aligned}$$

where we have used the fact that

$$\int_\gamma \left(\langle \sigma_\nu(U_{\mathcal{P}}) \rangle\, [v] + [U_{\mathcal{P}}]\, \langle \sigma_\nu(v) \rangle \right) \mathrm{d}s = 0 \quad \forall v \in X_{\mathsf{CR}}$$

and the definition of $U_{\mathcal{P}}$. Essentially the same argument shows that $g_{K|\gamma} = g$ on an edge $\gamma \in \mathcal{E}_N$.

The verification of condition (iii) makes use of the second and third identities in (3.18) along with the fact that g_K is piecewise constant, as follows. Summing over all edges γ of element $K \in \mathcal{P}$, we have

$$\begin{aligned} &\int_{\partial K} g_K \chi_K \,\mathrm{d}s \\ &\quad= \sum_{\gamma \subset \partial K} h_\gamma g_{K|\gamma} \\ &\quad= \sum_{\gamma \subset \partial K} (a\, \mathbf{grad}_{\mathcal{P}}\, U_{\mathcal{P}}, \mathbf{grad}_{\mathcal{P}}\, \theta_\gamma)_K - (f, \theta_\gamma)_K + \sum_{\gamma' \subset \partial K} \frac{\kappa}{h_{\gamma'}} \int_{\gamma'} [U_{\mathcal{P}}]\,[\theta_\gamma]\; \mathrm{d}s \\ &\quad= (a\, \mathbf{grad}_{\mathcal{P}}\, U_{\mathcal{P}}, \mathbf{grad}_{\mathcal{P}}\, \chi_K)_K - (f, \chi_K)_K \end{aligned}$$

where we have used the fact that

$$\sum_{\gamma \subset \partial K} [\theta_\gamma]_{|\gamma'} = \left[\sum_{\gamma \subset \partial K} \theta_\gamma \right]_{|\gamma'} = [\theta_{\gamma'}]_{|\gamma'} = 0.$$

Thus, the fluxes defined above satisfy conditions (i)–(iii). It is worth pointing out that, as was the case with discontinuous Galerkin FEM, the defining equation (2.13) plays a key role in the arguments. In contrast with DGFEM, where the test functions could be chosen with support on *a single element*, the smallest support in the case of the Crouzeix-Raviart FEM consists of *a pair of elements*. Nonetheless, the fact that only two elements were involved permitted an explicit solution.

3.2.3. *Conforming FEM.* In view of the foregoing developments, it will not come as a surprise to find that in the case of conforming FEM one must again use the defining equation (2.13) and work over patches of elements dictated by test functions of minimal support. Here, the smallest supports are patches of elements sharing a single common vertex. Bearing in mind that the number of elements in such a patch varies depending on the topology, one may anticipate that the construction of a closed form expression for the flux functions will be rather more complicated than in the previous cases and this indeed turns out to be the case. Algorithms for the construction of equilibrated fluxes for conforming approximation have already been discussed in detail in the literature [**4**, Section 6.3] and we shall not dwell upon this further other than to point out that it is possible to construct piecewise linear flux functions satisfying conditions (ii) and (iii).

3.3. Treatment of Local Residual Problem. We have shown that it is possible to construct flux functions $\{g_K : K \in \mathcal{P}\}$ such that the functional

$$H^1_E(K) \ni v \to (f, v)_K + \int_{\partial K} g_K v \, \mathrm{d}s - (a\, \mathbf{grad}_{\mathcal{P}}\, U_{\mathcal{P}}, \mathbf{grad}\, v)_K \tag{3.20}$$

is bounded, and vanishes when v is chosen to be constant. Thus, the functional is continuous and well-defined on the quotient space $H^1_E(K)/\mathbb{R}$ and it follows from the Lax-Milgram Theorem that there exists a vector-field $\boldsymbol{\alpha}_K \in \boldsymbol{L}_2(K)$ representing the functional:

$$\begin{aligned}&(\boldsymbol{\alpha}_K, \mathbf{grad}\, v)_K = \\ &\qquad (f, v)_K + \int_{\partial K} g_K v \, \mathrm{d}s - (a\, \mathbf{grad}_{\mathcal{P}}\, U_{\mathcal{P}}, \mathbf{grad}\, v)_K \quad \forall v \in H^1_E(K).\end{aligned}$$

In fact, the vector-field $\boldsymbol{\alpha}_K$ could be taken as potential, i.e. $\boldsymbol{\alpha}_K = \mathbf{grad}\, \varepsilon_K$, where $\varepsilon_K \in H^1_E(K)/\mathbb{R}$. Hence, continuing from identity (3.15), we have

$$(a\, \mathbf{grad}\, \phi, \mathbf{grad}\, v) = \sum_{K \in \mathcal{P}} (\boldsymbol{\alpha}_K, \mathbf{grad}\, v)_K \quad \forall v \in H^1_E(\Omega).$$

Applying the Cauchy-Schwarz inequality we deduce the following upper bound for the conforming part of the error

$$(a^{-1} \boldsymbol{\sigma}(\phi), \boldsymbol{\sigma}(\phi)) \le \sum_{K \in \mathcal{P}} (a^{-1} \boldsymbol{\alpha}_K, \boldsymbol{\alpha}_K)_K. \tag{3.21}$$

The estimate shows that one may obtain a guaranteed upper bound on the conforming part of the error provided we can *compute* the norm of the representor $\boldsymbol{\alpha}_K$. The particular structure of the local problem means it is possible to compute the representor and an explicit closed form for its norm:

$$\|\boldsymbol{\alpha}_K\|^2_K = \frac{1}{12} |K| \left[\left| \sum_{\boldsymbol{x}_n \in K} \boldsymbol{\varrho}_{K,n} \right|^2 + \sum_{\boldsymbol{x}_n \in K} \left| \boldsymbol{\varrho}_{K,n} \right|^2 \right], \tag{3.22}$$

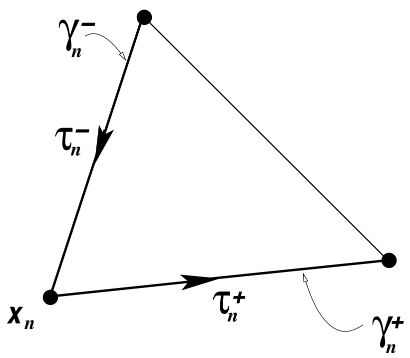

FIGURE 1. Orientation of tangent vectors $\boldsymbol{\tau}_n^{\pm}$ associated with a vertex $\boldsymbol{x}_n$.

where $|K|$ is the area of the element, $\boldsymbol{\varrho}_n$ is a vector associated with vertex $\boldsymbol{x}_n$ of element K defined by

$$\boldsymbol{\varrho}_{K,n} = \frac{1}{\boldsymbol{\tau}_n^+ \cdot \boldsymbol{R}\boldsymbol{\tau}_n^-} \left[\rho_{K,n}^+ \|\boldsymbol{\tau}_n^+\| \, \boldsymbol{\tau}_n^- - \rho_{K,n}^- \|\boldsymbol{\tau}_n^-\| \, \boldsymbol{\tau}_n^+ \right]$$

where

$$\rho_{K,n}^{\pm} = (g_K - \sigma_{\nu_K}(U_{\mathcal{P}}))\,|_{\gamma_n^{\pm}}(\boldsymbol{x}_n),$$

$\boldsymbol{\tau}_n^{\pm}$ are tangent vectors on the edges $\gamma_n^{\pm}$ of element K adjacent to vertex $\boldsymbol{x}_n$ (see Fig. 1) and $\boldsymbol{R}$ is the matrix

$$\boldsymbol{R} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

We refer the interested reader to [**1**] for further details. Once the fluxes g_K for a particular element are in hand, it is a simple matter to evaluate the vectors $\boldsymbol{\varrho}_{K,n}$ in terms of quantities local to triangle K and hence, from (3.21), we obtain a computable upper bound on the conforming part of the error:

$$(a^{-1}\boldsymbol{\sigma}(\phi), \boldsymbol{\sigma}(\phi)) \leq \sum_{K \in \mathcal{P}} \eta_{\mathsf{CF},K}^2 \tag{3.23}$$

where

$$\eta_{\mathsf{CF},K}^2 = \frac{|K|}{12 a_K} \left[\left| \sum_{\boldsymbol{x}_n \in K} \boldsymbol{\varrho}_{K,n} \right|^2 + \sum_{\boldsymbol{x}_n \in K} \left| \boldsymbol{\varrho}_{K,n} \right|^2 \right]. \tag{3.24}$$

Although the datum f does not appear explicitly in the estimator, the relationship between the data and the equilibrated fluxes means that f is taken into account implicitly. It may be shown that this estimator is efficient in the sense that there exists a local lower bound of the form

$$\eta_{\mathsf{CF},K}^2 \leq C(a^{-1}\boldsymbol{\sigma}(\phi), \boldsymbol{\sigma}(\phi))_{\widetilde{K}} \tag{3.25}$$

where C is a positive constant independent of any mesh-size.

3.4. A Posteriori Bounds on Total Error. Combining the result of Theorem 1 and the estimates (3.11) and (3.23) obtained above, we conclude that

$$(a^{-1}\boldsymbol{\sigma}_{\mathcal{P}}(e), \boldsymbol{\sigma}_{\mathcal{P}}(e)) \leq \sum_{K\in\mathcal{P}} (\eta^2_{\mathsf{CF},K} + \eta^2_{\mathsf{NC},K}) \tag{3.26}$$

along with the lower bound

$$\sum_{K\in\mathcal{P}} (\eta^2_{\mathsf{CF},K} + \eta^2_{\mathsf{NC},K}) \leq C(a^{-1}\boldsymbol{\sigma}_{\mathcal{P}}(e), \boldsymbol{\sigma}_{\mathcal{P}}(e)). \tag{3.27}$$

Each of the contributions to the bound can be computed explicitly and do not involve any additional unknown constants. As such, one obtains a computable upper bound on the error in energy. Although the constant appearing in the lower bound is not explicitly given, it is known to be independent of the solution and any mesh-size and hence the upper bound is efficient in the sense that it will not degenerate as a mesh is refined provided that one maintains shape regularity of the elements.

References

1. M. Ainsworth, *A posteriori error estimation for a discontinuous Galerkin finite element method*, In preparation.
2. ———, *A posteriori error estimation for non-conforming quadrilateral finite elements*, Accepted by Int. J. Numer. Anal. Model.
3. ———, *Robust a posteriori error estimation for non-conforming finite element approximation*, Accepted by SIAM J. Numer. Anal.
4. M. Ainsworth and J.T. Oden, *A posteriori error estimation in finite element analysis*, Pure and Applied Mathematics, Wiley-Interscience, John Wiley & Sons, New York, 2000.
5. D.N. Arnold, F. Brezzi, B. Cockburn, and L.D. Marini, *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM J. Numer. Anal. **39** (2001/02), no. 5, 1749–1779, (electronic).
6. I. Babuska and T. Strouboulis, *The finite element method and its reliability*, Numerical mathematics and scientific computation, Oxford University Press, Oxford, 2001.
7. W. Bangerth and R. Rannacher, *Adaptive finite element methods for differential equations*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2003.
8. R. Becker, P. Hansbo, and M.G. Larson, *Energy norm a posteriori error estimation for discontinuous Galerkin methods*, Comput. Methods Appl. Mech. Engrg. **192** (2003), no. 5-6, 723–733.
9. R. Becker, P. Hansbo, and R. Stenberg, *A finite element method for domain decomposition with non-matching grids*, M2AN Math. Model. Numer. Anal. **37** (2003), no. 2, 209–225.
10. C. Carstensen, S. Bartels, and S. Jansche, *A posteriori error estimates for nonconforming finite element methods*, Numer. Math. **92** (2002), no. 2, 233–256.
11. M. Crouzeix and P.A. Raviart, *Conforming and nonconforming finite element methods for solving the stationary Stokes equations*, M2AN Math. Model. Numer. Anal. **3** (1973), 33–75.
12. E. Dari, R. Duran, and C. Padra, *Error estimators for nonconforming finite-element approximations of the Stokes problem*, Math. Comp. **64** (1995), no. 211, 1017–1033.
13. E. Dari, R. Duran, C. Padra, and V. Vampa, *A posteriori error estimators for nonconforming finite element methods*, M2AN Math. Model. Numer. Anal. **30** (1996), no. 4, 385–400.
14. W. Dörfler and M. Ainsworth, *Reliable a posteriori error control for non-conforming finite element approximation of Stokes flow*, Tech. Report NI03999-CPD, Newton Institute for Mathematical Sciences, Cambridge, 2003.
15. K. Eriksson, D. Estep, P. Hansbo, and C. Johnson, *Introduction to adaptive methods for differential equations*, Acta numerica, 1995, Acta Numer., Cambridge Univ. Press, Cambridge, 1995, pp. 105–158.
16. R.H.W. Hoppe and B. Wohlmuth, *Element-oriented and edge-oriented local error estimators for nonconforming finite element methods*, M2AN Math. Model. Numer. Anal. **30** (1996), 237–263.

17. G. Kanschat and F.-T. Suttmeier, *A posteriori error estimates for nonconforming finite element schemes*, Calcolo **36** (1999), no. 3, 129–141.
18. O.A. Karakashian and F. Pascal, *A posteriori error estimation for a discontinuous Galerkin approximation of second-order elliptic problems*, SIAM J. Numer. Anal. **41** (2003), 2374–2399.
19. R. Rannacher and S. Turek, *Simple nonconforming quadrilateral Stokes element*, Numer. Meth. PDE. **8** (1992), no. 2, 97–111.
20. B. Rivière and M. F. Wheeler, *A posteriori error estimates for a discontinuous Galerkin method applied to elliptic problems*, Comput. Math. Appl. **46** (2003), no. 1, 141–163.
21. R. Verfürth, *A review of a posteriori error estimation and adaptive mesh refinement techniques*, Wiley-Teubner, 1996.

Mathematics Department, Strathclyde University, 26 Richmond Street, Glasgow G1 1XH, Scotland.

E-mail address: `M.Ainsworth@strath.ac.uk`