SIAM J. SCI. COMPUT.
Vol. 30, No. 1, pp. 189–204

A POSTERIORI ERROR ESTIMATION FOR LOWEST ORDER RAVIART–THOMAS MIXED FINITE ELEMENTS*

MARK AINSWORTH†

Abstract. For the lowest order Raviart–Thomas mixed finite element, we derive an a posteriori error estimator that provides actual, guaranteed computable upper bounds on the error in the flux variable regardless of jumps in the material coefficients across interfaces. Moreover, the estimator is efficient in that it provides a local lower bound on the error up to a constant that is independent of the solution and the local mesh-size. The estimator may be evaluated at virtually no additional cost compared to the evaluation of the finite element approximation itself.

Key words. a posteriori error estimation, mixed finite element method, computable error bounds

AMS subject classifications. 65F10, 65N30, 65N50, 65N55

DOI. 10.1137/06067331X

1. Introduction.

Many mathematical models arising in continuum mechanics and physics are based on conservation of a physical quantity, and are naturally expressed in the form $-\mathbf{div}\,\boldsymbol{\sigma} = f$, where $\boldsymbol{\sigma}$ is a vector field representing the *flux* of the quantity and f is the corresponding source term measuring the generation of the quantity. Typically, the conservation law is coupled with a constitutive equation of the form $\boldsymbol{\sigma} = -A\,\mathbf{grad}\,u$, where A is a tensor depending on the particular medium and u is a scalar potential. Frequently, the flux is the primary quantity of interest, while the potential plays a secondary role and is often of little or even no interest. In light of such considerations, it is very natural to construct *mixed* numerical approximations of the problem that involve the flux as a primary unknown along with the potential, as opposed to, say, approximating the potential and calculating the flux via the constitutive equation.

Mixed approximations have been extensively studied and analyzed [8] and are often the method of choice in application areas such as simulation of flow in porous media, electrostatics, and magnetostatics. Such applications typically involve complex geometries and materials with widely differing material properties, corresponding to large jumps in the value of A across material interfaces. Each of these effects may give rise to local singularities in both the flux and potential that can severely degrade the accuracy of a numerical scheme. Adaptive finite element algorithms seek to improve the accuracy and convergence of numerical methods through performing local mesh refinement in the neighborhood of the points where the true solution is nonsmooth. Adaptive algorithms usually rely on the availability of a local error indicator that is used to identify elements where the accuracy is poor for refinement. Such procedures frequently produce sequences of meshes where certain elements may never be flagged for refinement by the local error indicator. If one is to have confidence in the numerical results obtained in such situations, it is imperative that the local error indicator be reliable. Moreover, the error indicator is often used as a stopping criterion for the

*Received by the editors October 25, 2006; accepted for publication (in revised form) April 25, 2007; published electronically December 7, 2007. This work was partially supported by the Engineering and Physical Sciences Research Council of Great Britain under grant GR/S35103.
http://www.siam.org/journals/sisc/30-1/67331.html
†Mathematics Department, Strathclyde University, 26 Richmond Street, Glasgow G1 1XH, Scotland (M.Ainsworth@strath.ac.uk).

adaptive feedback procedure at a level when the overall error levels are acceptable but by no means zero, or even necessarily small. Indeed, in many application areas the complexity of the problems means that it is simply not feasible to carry out *any* mesh refinement or convergence studies, and numerical results are available on just a single mesh. In such a setting, it is of considerable importance that the error estimator provides *guaranteed* and *quantitative bounds* on the error if one is to have confidence in the results of the simulation. It is of equal importance that the quality of the bounds is not adversely affected by jumps in the material properties across interfaces.

A number of a posteriori error estimators and indicators have been proposed for use with mixed finite element approximations. Residual-type estimators, based on norms of residuals and jumps in the finite element approximation weighted in terms of the local mesh-size, were obtained by Baranger and El Amri [5]. Brandts [7] derived a posteriori error estimators for flux variable based on superconvergence properties of the solution. Subsequently, Alonso [3] derived residual-type estimators and estimators based on the solution of a local BVP with the residual as data and showed both estimators were equivalent to the true error. A different estimator for the error in the flux variable using a saturation assumption was analyzed in [6] but was later shown to be suboptimal [9, 12]. An estimator for the flux error in the $H(\mathbf{div})$-norm was obtained by Carstensen [9] for the case of convex domain. Hierarchic basis-type estimators were proposed by Achchab et al. [1]. More recently, Lovadina and Stenberg [12] obtained a posteriori estimators for the error measured in a mesh-dependent norm using a postprocessing procedure, while estimators for anisotropic elements were considered in [15]. A useful overview and summary of the relationships between these estimators will be found in the work of [20]. More specifically, it is shown that many of the estimators are, under appropriate assumptions, equivalent to one another at the element level, and some of the estimators are shown to be equivalent to the actual global error up to generic, unknown constants depending on the jumps in the material properties.

Our objective in the present work is to derive an estimator that provides actual guaranteed computable upper bounds on the error in the flux variable regardless of jumps in the material coefficients. Moreover, we shall show that the estimator is efficient in that it provides a local lower bound on the error up to a constant that is independent of the solution and the local mesh-size. The local lower bound depends on the distribution of the material properties in the neighborhood of a particular element and could conceivably degenerate as jumps in the materials grow. However, the dependence is rather mild and may manifest itself only in a small number of elements located near cross points involving four or more different materials where one would expect the mesh to require local refinement in any case, which would in turn suppress any ill effects of the estimator at the global level.

The estimator itself is rather simple and may be evaluated at virtually no additional cost compared to the evaluation of the finite element approximation itself. Specifically, we use a local postprocessing procedure applied to the mixed finite element approximation at the element level, similar to those analyzed in [4, 13, 18, 19], followed by a local smoothing at element vertices weighted in terms of the material coefficients, to obtain an approximation $u^\circ_{\mathcal{P}}$ to the potential u. Remarkably, the norm of the difference $\boldsymbol{\sigma}^{RT}_{\mathcal{P}} - A\,\mathbf{grad}\,u^\circ_{\mathcal{P}}$, where $\boldsymbol{\sigma}^{RT}_{\mathcal{P}}$ is the mixed approximation of the flux variable, is shown to provide a guaranteed upper bound on the error in the case where f is piecewise constant. For more general data, we include an additional term, again completely computable, measuring the oscillation of f relative to the mesh to obtain a guaranteed upper bound on the error. The main results are summarized in Corollary 1.

The utility of the estimator is illustrated by applying it to the finite element analysis of a two-pole electric motor where the permittivity of the components varies by several orders of magnitude and the geometry gives rise to multiple internal singularities. Applying an adaptive procedure we find that the estimator does indeed provide an upper bound and overestimates the actual error by a factor of between 1.3–1.6.

2. Preliminaries. Consider the following model problem:

$$\left.\begin{aligned} \boldsymbol{\sigma} - A\,\mathbf{grad}\,u &= 0 \\ \mathbf{div}\,\boldsymbol{\sigma} + f &= 0 \end{aligned}\right\} \text{ in } \Omega \tag{1}$$

subject to $u = 0$ on $\partial\Omega$, where Ω is a polygonal domain and $f \in L_2(\Omega)$ is given. The matrix-valued function $A \in L_\infty(\Omega; \mathbb{R}^{2\times 2})$ is assumed to be symmetric positive definite and, for ease of exposition, is assumed piecewise constant on subdomains of Ω. The following theory is readily extended to include the case of mixed boundary conditions.

The mixed variational formulation of the problem consists of seeking a pair $(u, \boldsymbol{\sigma}) \in L_2(\Omega) \times H(\mathbf{div}; \Omega)$ satisfying

$$\begin{aligned} (A^{-1}\boldsymbol{\sigma}, \boldsymbol{\tau}) + (u, \mathbf{div}\,\tau) &= 0, \\ (\mathbf{div}\,\boldsymbol{\sigma}, v) + (f, v) &= 0 \end{aligned} \tag{2}$$

for all $(v, \boldsymbol{\tau}) \in L_2(\Omega) \times H(\mathbf{div}; \Omega)$. Here, $H(\mathbf{div}; \Omega)$ denotes the space of vector fields $\{\boldsymbol{\tau} \in L_2(\Omega)^2 : \mathbf{div}\,\boldsymbol{\tau} \in L_2(\Omega)\}$, which we equip with the weighted inner product defined by

$$(\boldsymbol{\sigma}, \boldsymbol{\tau})_{H(\mathbf{div};\Omega)} = (\boldsymbol{\sigma}, \boldsymbol{\tau})_{A^{-1}} + (\mathbf{div}\,\boldsymbol{\sigma}, \mathbf{div}\,\boldsymbol{\tau}), \tag{3}$$

where $(\boldsymbol{\sigma}, \boldsymbol{\tau})_{A^{-1}} = (A^{-1}\boldsymbol{\sigma}, \boldsymbol{\tau})$, with the associated norms denoted by $\|\cdot\|_{H(\mathbf{div};\Omega)}$ and $\|\cdot\|_{A^{-1}}$. When $\omega \subset \Omega$ we shall use the notation $(\cdot,\cdot)_\omega$ to denote the integral inner product over ω and omit the subscript in the case where ω is the physical domain Ω.

We consider a family of partitions $\{\mathcal{P}\}$ of the domain Ω into the union of nonoverlapping, triangular elements such that the nonempty 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 neighboring elements is uniformly bounded above and below over the whole family.

In addition, whenever possible, the partitioning is chosen to reflect the structure of the permeability matrix in the sense that individual elements do not straddle a subdomain boundary where the value of A undergoes a large jump. This requirement is reflected in the assumption that, for every element $K \in \mathcal{P}$, there exist positive constants λ_K and Λ_K satisfying

$$\lambda_K\,|\vec{p}|^2 \le \vec{p}^{\,\top} A_K \vec{p} \le \Lambda_K\,|\vec{p}|^2, \quad \vec{p} \in \mathbb{R}^2, \tag{4}$$

such that the ratio $\Upsilon_K = \Lambda_K / \lambda_K$ is uniformly bounded over the whole family of partitions. It will be important to develop a posteriori error estimators whose reliability and efficiency are insensitive to the magnitude of the jumps in permeability between differing regions.

We shall be at pains to extract the explicit dependence of the various estimates on the values of λ_K and Λ_K, and throughout we shall utilize c and C to denote generic constants whose values may be different in any two places but which are independent

of any mesh-size, the data A, and the solution of the problem. In seeking to exhibit the dependence of the constants in various estimates on the data, it will be useful to introduce the notion of *relative path permeability* [2].

Let Ω_n denote the patch composed from those elements with a vertex located at $\boldsymbol{x}_n$, and let $K,\ K' \subset \Omega_n$ be distinct elements. We can always find at least one connected path $\wp(K,K') \subset \mathcal{P}$ passing between the elements K to K' through adjacent elements belonging to the patch Ω_n. The smallest value of the permeability on all elements along the path $\wp(K,K')$ is given by $\min\{\lambda_M : M \in \wp(K,K')\}$. If $\boldsymbol{x}_n$ is an interior vertex, then there are two such paths, and we take $\wp(K,K')$ to be the path $\wp^*(K,K')$ which maximizes the value of this quantity and define

$$\lambda_{KK'} = \min\{\lambda_M : M \in \wp^*(K,K')\}. \tag{5}$$

The *relative path permeability* is then defined by

$$\Upsilon_{KK'} = \frac{\min(\Lambda_K, \Lambda_{K'})}{\lambda_{KK'}} \tag{6}$$

and measures the ratio of the path permeability to the least permeable of the two elements K and K' at the endpoints of the path. This notion was first used in [2], where the reader will find further information and examples.

The set of all edges of the elements is denoted by $\mathcal{E}$, while the edges of an element $K \in \mathcal{P}$ are denoted by $\mathcal{E}(K)$. The set of element nodes is denoted by $\mathcal{N}$, while the nodes on a particular element K or edge γ are denoted by $\mathcal{N}(K)$ or $\mathcal{N}(\gamma)$, respectively. For each edge $\gamma \in \mathcal{E}$, the set $\widetilde{\gamma}$ consists of those elements for which γ is an edge,

$$\widetilde{\gamma} = \{K' \in \mathcal{P} : \gamma \in \mathcal{E}(K')\}, \tag{7}$$

while for each element $K \in \mathcal{P}$, the set $\widetilde{K}$ consists of those elements having a vertex in common with K,

$$\widetilde{K} = \{K' \in \mathcal{P} : \mathcal{N}(K) \cap \mathcal{N}(K') \text{ is nonempty}\}. \tag{8}$$

3. Mixed finite element approximation. We discretize problem (2) by introducing finite dimensional subspaces $M_{\mathcal{P}} \subset L_2(\Omega)$ and $Q_{\mathcal{P}} \subset H(\mathbf{div};\Omega)$ relative to the partition $\mathcal{P}$ as follows:

$$\begin{aligned} M_{\mathcal{P}} &= \{v \in L_2(\Omega) : v_{|K} \in \mathbb{R} \quad \forall K \in \mathcal{P}\}, \\ Q_{\mathcal{P}} &= \{\boldsymbol{\tau} \in H(\mathbf{div};\Omega) : \boldsymbol{\tau}_{|K} \in \mathbb{R}^2 + \boldsymbol{x}\cdot\mathbb{R} \quad \forall K \in \mathcal{P}\}, \end{aligned} \tag{9}$$

and we seek a pair $(u_{\mathcal{P}}^{RT}, \boldsymbol{\sigma}_{\mathcal{P}}^{RT}) \in M_{\mathcal{P}} \times Q_{\mathcal{P}}$ satisfying

$$\begin{aligned} (A^{-1}\boldsymbol{\sigma}_{\mathcal{P}}^{RT}, \boldsymbol{\tau}) + (u_{\mathcal{P}}^{RT}, \mathbf{div}\,\boldsymbol{\tau}) &= 0, \\ (\mathbf{div}\,\boldsymbol{\sigma}_{\mathcal{P}}^{RT}, v) + (f, v) &= 0 \end{aligned} \tag{10}$$

for all $(v, \boldsymbol{\tau}) \in M_{\mathcal{P}} \times Q_{\mathcal{P}}$. The space $Q_{\mathcal{P}}$ corresponds to the lowest order Raviart–Thomas finite element [17]. The pair of spaces $M_{\mathcal{P}}$ and $Q_{\mathcal{P}}$ provides a stable approximation of the mixed problem (2). The interested reader is referred to [8] for more information and a priori error analysis. Here, we shall concern ourselves with the a posteriori error analysis of the method.

The following simple observation will prove itself useful when we come to the a posteriori error analysis. Let $\Pi_{\mathcal{P}} : L_2(\Omega) \to M_{\mathcal{P}}$ denote orthogonal projection with respect to the $L_2(\Omega)$ inner product; i.e., $\Pi_{\mathcal{P}} f \in M_{\mathcal{P}}$ satisfies

$$(\Pi_{\mathcal{P}} f, v) = (f, v) \quad \forall v \in M_{\mathcal{P}}.$$

It will be useful to consider the solution of the original model problem (2) with the datum f replaced by $\Pi_{\mathcal{P}} f$, i.e., $(u^\star, \boldsymbol{\sigma}^\star) \in L_2(\Omega) \times H(\mathbf{div}; \Omega)$:

$$\begin{aligned} (A^{-1}\boldsymbol{\sigma}^\star, \boldsymbol{\tau}) + (u^\star, \mathbf{div}\, \boldsymbol{\tau}) &= 0, \\ (\mathbf{div}\, \boldsymbol{\sigma}^\star, v) + (\Pi_{\mathcal{P}} f, v) &= 0 \end{aligned} \tag{11}$$

for all $(v, \boldsymbol{\tau}) \in L_2(\Omega) \times H(\mathbf{div}; \Omega)$. In particular, thanks to the definition of $\Pi_{\mathcal{P}} f$, the mixed finite element approximations of both problems are *identical*, given by $(u_{\mathcal{P}}^{RT}, \boldsymbol{\sigma}_{\mathcal{P}}^{RT})$. Consequently, one may choose to view $(u_{\mathcal{P}}^{RT}, \boldsymbol{\sigma}_{\mathcal{P}}^{RT})$ as an approximation of the solution of the problem (11).

Marini [13] observed that the Raviart–Thomas mixed finite element approximation (10) is related to the nonconforming Crouzeix–Raviart finite element approximation $u_{\mathcal{P}}^{CR}$ [10] of the problem

$$-\mathbf{div}(A\, \mathbf{grad}\, u) = \Pi_{\mathcal{P}} f \text{ in } \Omega; \quad u = 0 \text{ on } \partial\Omega$$

through the expressions

$$\left.\begin{aligned} \boldsymbol{\sigma}_{\mathcal{P}}^{RT}{}_{|K} &= A\, \mathbf{grad}\, u_{\mathcal{P}}^{CR} - \frac{1}{2} f_K (\boldsymbol{x} - \boldsymbol{x}_K) \\ u_{\mathcal{P}}^{RT}{}_{|K} &= \frac{1}{|K|} \int_K \left[u_{\mathcal{P}}^{CR} + \frac{1}{4} f (\boldsymbol{x} - \boldsymbol{x}_K)^\top A^{-1} (\boldsymbol{x} - \boldsymbol{x}_K) \right] \mathrm{d}\boldsymbol{x} \end{aligned}\right\}, \tag{12}$$

where f_K denotes the average value of f on the element K and $\boldsymbol{x}_K$ denotes the centroid.

The computation of the Crouzeix–Raviart approximation enjoys certain advantages over the computation of the Raviart–Thomas approximation. For instance, the Crouzeix–Raviart approximation involves fewer unknowns and is determined by a symmetric, positive definite stiffness matrix, as opposed to an indefinite saddle point problem in the case of the mixed Raviart–Thomas approximation. These advantages have prompted some practitioners to advocate the computation of the Raviart–Thomas approximation by first evaluating the Crouzeix–Raviart approximation and then applying the relations (12) to obtain the mixed approximation.

A crucial role in the subsequent a posteriori error analysis will be played by the piecewise function $u_{\mathcal{P}}^\circ$ defined on the partition $\mathcal{P}$ in terms of the Crouzeix–Raviart approximation as follows:

$$u_{\mathcal{P}|K}^\circ = u_{\mathcal{P}}^{CR} - \frac{1}{2} f_K \psi_K(\boldsymbol{x}), \quad K \in \mathcal{P}, \tag{13}$$

where ψ_K is the piecewise quadratic function defined by

$$\psi_K(\boldsymbol{x}) = \frac{1}{2} (\boldsymbol{x} - \boldsymbol{x}_K)^\top A^{-1} (\boldsymbol{x} - \boldsymbol{x}_K) - \frac{1}{|K|} \int_K (\boldsymbol{x}' - \boldsymbol{x}_K)^\top A^{-1} (\boldsymbol{x}' - \boldsymbol{x}_K)\, \mathrm{d}\boldsymbol{x}'.$$

The relations (12) may be expressed directly in terms of $u_{\mathcal{P}}^\circ$:

$$\left.\begin{aligned} \boldsymbol{\sigma}_{\mathcal{P}}^{RT} &= A\, \mathbf{grad}\, u_{\mathcal{P}}^\circ \\ u_{\mathcal{P}}^{RT}{}_{|K} &= \frac{1}{|K|} \int_K u_{\mathcal{P}}^\circ \, \mathrm{d}\boldsymbol{x} \end{aligned}\right\}. \tag{14}$$

As a matter of fact, the function $u^\circ_{\mathcal{P}}$ may be obtained directly from the Raviart–Thomas approximation without any mention of the Crouzeix–Raviart approximation at all via a local postprocessing procedure on each element $K \in \mathcal{P}$ as follows: seek $u^\circ_{\mathcal{P}} \in \mathbb{P}_2(K)$ such that

$$\left.\begin{aligned} (A\,\mathbf{grad}\,u^\circ_{\mathcal{P}}, \mathbf{grad}\,v)_K &= (\boldsymbol{\sigma}^{RT}_{\mathcal{P}}, \mathbf{grad}\,v)_K \quad \forall v \in \mathbb{P}_2(K) \\ \frac{1}{|K|}\int_K u^\circ_{\mathcal{P}}\,\mathrm{d}\boldsymbol{x} &= u^{RT}_{\mathcal{P}}(\boldsymbol{x}_K) \end{aligned}\right\}. \tag{15}$$

Let us show that the solution of (15) is identical with the function defined in (13). It is easy to see that the finite dimensional approximation $u^\circ_{\mathcal{P}}$ is well defined by (15); e.g., if the data $(u^{RT}_{\mathcal{P}}, \boldsymbol{\sigma}^{RT}_{\mathcal{P}})$ vanishes, then both $\mathbf{grad}\,u^\circ_{\mathcal{P}}$ and the average value of $u^\circ_{\mathcal{P}}$ must vanish, and hence $u^\circ_{\mathcal{P}} = 0$, and the problem is uniquely solvable. We may now go further and note that, since $\boldsymbol{\sigma}^{RT}_{\mathcal{P}} \in \mathbf{grad}(\mathbb{P}_2)$, the first condition implies that $\boldsymbol{\sigma}^{RT}_{\mathcal{P}}$ must coincide with $A\,\mathbf{grad}\,u^\circ_{\mathcal{P}}$, and hence the solution $u^\circ_{\mathcal{P}}$ satisfies (14). Similar postprocessing procedures have been analyzed from the a priori point of view in [4, 18, 19].

4. A posteriori error estimation. We shall concern ourselves with the development of computable a posteriori bounds for the error $(u - u^{RT}_{\mathcal{P}}, \boldsymbol{\sigma} - \boldsymbol{\sigma}^{RT}_{\mathcal{P}})$ in the approximation. For this purpose, it is useful to regard the error in the flux as comprising of two distinct parts:

$$\boldsymbol{\sigma} - \boldsymbol{\sigma}^{RT}_{\mathcal{P}} = (\boldsymbol{\sigma} - \boldsymbol{\sigma}^\star) + (\boldsymbol{\sigma}^\star - \boldsymbol{\sigma}^{RT}_{\mathcal{P}}). \tag{16}$$

The first component $\boldsymbol{\sigma} - \boldsymbol{\sigma}^\star$ represents the error introduced through data approximation $f \approx \Pi_{\mathcal{P}} f$, which depends explicitly on the choice of partitioning $\mathcal{P}$ but only indirectly on the discretization scheme. Conversely, the second part $\boldsymbol{\sigma}^\star - \boldsymbol{\sigma}^{RT}_{\mathcal{P}}$ depends on the choice of finite element discretization scheme but involves no data approximation, since $\boldsymbol{\sigma}^\star$ is computed from the same data $\Pi_{\mathcal{P}} f$ used in the computation of the approximation $\boldsymbol{\sigma}^{RT}_{\mathcal{P}}$. Based on these observations, we shall refer to $\boldsymbol{\sigma} - \boldsymbol{\sigma}^\star$ as the *data error*, while $\boldsymbol{\sigma}^\star - \boldsymbol{\sigma}^{RT}_{\mathcal{P}}$ is referred to as the *discretization error*.

At first sight, this decomposition may appear somewhat ad hoc. However, it has the property of being simultaneously orthogonal in two different inner products.

THEOREM 1. *The splitting* (16) *is orthogonal in the sense that both*

$$(\boldsymbol{\sigma} - \boldsymbol{\sigma}^\star, \boldsymbol{\sigma}^\star - \boldsymbol{\sigma}^{RT}_{\mathcal{P}})_{A^{-1}} = 0 \tag{17}$$

and

$$(\boldsymbol{\sigma} - \boldsymbol{\sigma}^\star, \boldsymbol{\sigma}^\star - \boldsymbol{\sigma}^{RT}_{\mathcal{P}})_{H(\mathbf{div};\Omega)} = 0. \tag{18}$$

Consequently,

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}^{RT}_{\mathcal{P}}\|^2_{A^{-1}} = \|\boldsymbol{\sigma} - \boldsymbol{\sigma}^\star\|^2_{A^{-1}} + \|\boldsymbol{\sigma}^\star - \boldsymbol{\sigma}^{RT}_{\mathcal{P}}\|^2_{A^{-1}} \tag{19}$$

and

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}^{RT}_{\mathcal{P}}\|^2_{H(\mathbf{div};\Omega)} = \|\boldsymbol{\sigma} - \boldsymbol{\sigma}^\star\|^2_{H(\mathbf{div};\Omega)} + \|\boldsymbol{\sigma}^\star - \boldsymbol{\sigma}^{RT}_{\mathcal{P}}\|^2_{H(\mathbf{div};\Omega)}. \tag{20}$$

Proof. Observe that $\Pi_{\mathcal{P}} f$ and $\mathbf{div}\,\boldsymbol{\sigma}^{RT}_{\mathcal{P}} \in M_{\mathcal{P}}$ and, for $v \in M_{\mathcal{P}}$, that the second condition of (10) and the definition of $\Pi_{\mathcal{P}} f$ give

$$(\mathbf{div}\,\boldsymbol{\sigma}^{RT}_{\mathcal{P}} + \Pi_{\mathcal{P}} f, v) = (\mathbf{div}\,\boldsymbol{\sigma}^{RT}_{\mathcal{P}} + f, v) = 0,$$

whence $-\mathbf{div}\,\boldsymbol{\sigma}_{\mathcal{P}}^{RT} = \Pi_{\mathcal{P}} f$. Furthermore, using the second part of (11) gives $-\mathbf{div}\,\boldsymbol{\sigma}^{\star} = \Pi_{\mathcal{P}} f$, and therefore

$$-\mathbf{div}(\boldsymbol{\sigma}^{\star} - \boldsymbol{\sigma}_{\mathcal{P}}^{RT}) = 0. \tag{21}$$

The first parts of (2) and (11) mean that for all $\boldsymbol{\tau} \in H(\mathbf{div}; \Omega)$,

$$(\boldsymbol{\sigma} - \boldsymbol{\sigma}^{\star}, \boldsymbol{\tau})_{A^{-1}} = (A^{-1}(\boldsymbol{\sigma} - \boldsymbol{\sigma}^{\star}), \boldsymbol{\tau}) = -(u - u^{\star}, \mathbf{div}\,\boldsymbol{\tau}), \tag{22}$$

and then choosing $\boldsymbol{\tau} = \boldsymbol{\sigma}^{\star} - \boldsymbol{\sigma}_{\mathcal{P}}^{RT}$ and using the result (21) gives (17). The orthogonality (18) is then an immediate consequence of (17) and (21), and the Pythagorean identities (19) and (20) follow at once. ∎

Theorem 1 means that the derivation of a posteriori error bounds for the flux error $\boldsymbol{\sigma} - \boldsymbol{\sigma}_{\mathcal{P}}^{RT}$ may be separated into two distinct stages—the derivation of bounds for the *data error* and bounds for the *approximation error*. The Pythagorean identities (19) and (20) may then be used to obtain bounds for the total error. Of course, one could obtain upper bounds by utilizing a triangle inequality but would give rise to a larger multiplicative coefficient.

4.1. Estimation of data error. As usual [11], we shall define the *oscillation* of a given datum $f \in L_2(\Omega)$ over an element $K \in \mathcal{P}$ by the rule

$$\operatorname{osc}(f, K) = \inf_{c \in \mathbb{R}} \|f - c\|_K. \tag{23}$$

The choice of c that minimizes the above expression is given explicitly by f_K, the average value of f over the element K. The next result gives an easily computable upper bound for the data error in terms of the oscillation, regardless of the smoothness or otherwise of the datum f.

THEOREM 2. *Let $K \in \mathcal{P}$. Then,*

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}^{\star}\|_{A^{-1}}^2 \leq \frac{1}{\pi^2} \sum_{K \in \mathcal{P}} \frac{\operatorname{diam}(K)^2}{\lambda_K} \operatorname{osc}(f, K)^2. \tag{24}$$

Moreover, if $f \in H^1(K)$, then

$$\operatorname{osc}(f, K) \leq \frac{1}{\pi} \operatorname{diam}(K) \|\mathbf{grad}\, f\|_K. \tag{25}$$

4.2. Estimation of approximation error. The estimation of the approximation error depends on the auxiliary function $u_{\mathcal{P}}^{\circ}$ constructed in (13). In general, the piecewise quadratic function $u_{\mathcal{P}}^{\circ}$ is discontinuous across element interfaces. We construct a continuous piecewise quadratic function $\widetilde{u}_{\mathcal{P}}^{\circ} \in H_0^1(\Omega)$ on the partition $\mathcal{P}$ whose values at element vertices and midpoints of edges are given by a weighted averaging of $u_{\mathcal{P}}^{\circ}$ as follows:

$$\widetilde{u}_{\mathcal{P}}^{\circ}(\boldsymbol{x}_n) = \begin{cases} 0 \quad \text{if} \boldsymbol{x}_n \in \partial\Omega, \\ \displaystyle\sum_{K \subset \Omega_n} \omega_K^{(n)} u_{\mathcal{P}|K}^{\circ}(\boldsymbol{x}_n) \quad \text{otherwise,} \end{cases} \tag{26}$$

where

$$\omega_K^{(n)} = \frac{\Lambda_K^{1/2}}{\sum_{K' \subset \Omega_n} \Lambda_{K'}^{1/2}} \tag{27}$$

and $\Omega_n = \{K \in \mathcal{P} : \boldsymbol{x}_n \in K\}$ is the patch of elements containing the node $\boldsymbol{x}_n$.

The function $\widetilde{u}^\circ_{\mathcal{P}}$ may be easily obtained from $u^\circ_{\mathcal{P}}$, which is in turn easily computed from the Crouzeix–Raviart approximation using (13) or, alternatively, directly from the Raviart–Thomas approximation using (15). The next result shows how $\widetilde{u}^\circ_{\mathcal{P}}$ may be used to obtain a readily computable upper bound for the approximation error and, moreover, that the bound is efficient up to a factor depending on the local relative path permeability (6).

THEOREM 3. *Let $\widetilde{u}^\circ_{\mathcal{P}}$ be defined in* (26)*. Then,*

$$\|\boldsymbol{\sigma}^\star - \boldsymbol{\sigma}^{RT}_{\mathcal{P}}\|_{A^{-1}} \le \|A\,\mathbf{grad}\,\widetilde{u}^\circ_{\mathcal{P}} - \boldsymbol{\sigma}^{RT}_{\mathcal{P}}\|_{A^{-1}} \tag{28}$$

and

$$\|A\,\mathbf{grad}\,\widetilde{u}^\circ_{\mathcal{P}} - \boldsymbol{\sigma}^{RT}_{\mathcal{P}}\|_{A^{-1},K} \le c_K \|\boldsymbol{\sigma}^\star - \boldsymbol{\sigma}^{RT}_{\mathcal{P}}\|_{A^{-1},\widetilde{K}}, \tag{29}$$

where $c_K = c\sum_{K'\subset\widetilde{K}} \Upsilon_{KK'}$, $\Upsilon_{KK'}$ is the relative path permeability defined in (6)*, and c is a positive constant independent of any mesh-size.*

Theorem 3 is analogous to Theorem 3.2 of [2]. However, there are some important differences. For instance, in [2] the smoothed function $\widetilde{u}^\circ_{\mathcal{P}}$ was taken to be the piecewise affine interpolant of similar weighted averages of the Crouzeix–Raviart approximation $u^{CR}_{\mathcal{P}}$ itself, as opposed to average of the augmented Crouzeix–Raviart approximation $u^\circ_{\mathcal{P}}$ defined by (13). If the smoothed approximation used in [2] were to be used in place of $\widetilde{u}^\circ_{\mathcal{P}}$ in the present situation, then we would still obtain an upper bound. However, the estimator would fail to be efficient in the sense that the local lower bound (29) would *not* hold, corresponding to the fact that the upper bounds would be increasingly pessimistic as mesh is refined.

4.3. Bounds for total error. The foregoing results give computable bounds on the total error.

COROLLARY 1. *The following computable upper bounds hold*:

$$\begin{aligned}\|\boldsymbol{\sigma} - \boldsymbol{\sigma}^{RT}_{\mathcal{P}}\|^2_{A^{-1}} &\\ \le \|A\,\mathbf{grad}\,\widetilde{u}^\circ_{\mathcal{P}} - \boldsymbol{\sigma}^{RT}_{\mathcal{P}}\|^2_{A^{-1}} &+ \frac{1}{\pi^2}\sum_{K\in\mathcal{P}} \frac{\operatorname{diam}(K)^2}{\lambda_K}\operatorname{osc}(f,K)^2\end{aligned} \tag{30}$$

and

$$\begin{aligned}\|\boldsymbol{\sigma} - \boldsymbol{\sigma}^{RT}_{\mathcal{P}}\|^2_{H(\mathbf{div};\Omega)} &\\ \le \|A\,\mathbf{grad}\,\widetilde{u}^\circ_{\mathcal{P}} - \boldsymbol{\sigma}^{RT}_{\mathcal{P}}\|^2_{A^{-1}} &+ \sum_{K\in\mathcal{P}}\left[1 + \frac{\operatorname{diam}(K)^2}{\pi^2\,\lambda_K}\right]\operatorname{osc}(f,K)^2.\end{aligned} \tag{31}$$

Moreover, the bounds are efficient in the sense that

$$c\|A\,\mathbf{grad}\,\widetilde{u}^\circ_{\mathcal{P}} - \boldsymbol{\sigma}^{RT}_{\mathcal{P}}\|^2_{A^{-1}} \le \|\boldsymbol{\sigma} - \boldsymbol{\sigma}^{RT}_{\mathcal{P}}\|^2_{A^{-1}} \tag{32}$$

and

$$c\|A\,\mathbf{grad}\,\widetilde{u}^\circ_{\mathcal{P}} - \boldsymbol{\sigma}^{RT}_{\mathcal{P}}\|^2_{A^{-1}} + \sum_{K\in\mathcal{P}}\operatorname{osc}(f,K)^2 \le \|\boldsymbol{\sigma} - \boldsymbol{\sigma}^{RT}_{\mathcal{P}}\|^2_{H(\mathbf{div};\Omega)}, \tag{33}$$

where $c = \min_{K\in\mathcal{P}} c_K$ with c_K defined in (29)*.*

Proof. The results follow at once by inserting the bounds given in Theorems 2 and 3 into the Pythagorean identities (19) and (20) and recalling that $-\mathbf{div}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_{\mathcal{P}}^{RT}) = f - \Pi_{\mathcal{P}} f$. ∎

The terms appearing in the lower bounds given in Corollary 1 differ from those in the upper bounds in that the term

$$\frac{1}{\pi^2} \sum_{K \in \mathcal{P}} \frac{\operatorname{diam}(K)^2}{\lambda_K} \operatorname{osc}(f, K)^2 \tag{34}$$

is absent. As shown in Theorem 2, for piecewise smooth data $f \in H^1(K)$, each contribution is of the order $\mathcal{O}(h_K^4 \|\mathbf{grad}\, f\|_K^2)$. Therefore, in view of the fact that we expect $\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_{\mathcal{P}}^{RT}\|_{A^{-1},K}^2 = \mathcal{O}(h_K^4)$, the additional terms are seen to be of higher order than the remaining terms appearing in the estimator, and one can expect the estimator to be efficient. Of course, it is always possible to construct pathological examples where this heuristic reasoning breaks down, but we reiterate that *the upper bounds are guaranteed regardless of the regularity or otherwise of the data* and are fully computable given in terms of the oscillation of the data.

5. Numerical examples.

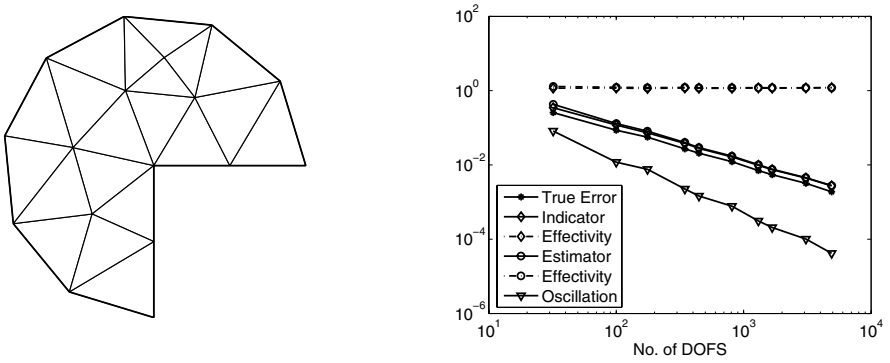
5.1. Singular Poisson problem. We consider the problem of finding $u \in H_0^1(\Omega)$ such that $-\Delta u = f$ on the domain $\Omega = \{(r, \theta) : 0 \le r < 1, 0 < \theta < 3\pi/2\}$, with the data chosen so that the true solution is given by $u(r, \theta) = (r^{2/3} - r^3)\sin(2\theta/3)$. The problem is approximated using an adaptive refinement algorithm with initial mesh shown in Figure 1(a) and successively refining all elements whose contribution to the error estimator exceeds 30% of the largest local estimator. Table 1 shows the values of the true error in the fluxes $\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_{\mathcal{P}}^{RT}\|_{A^{-1}}$, the error indicator $\|A\,\mathbf{grad}\,\widetilde{u}_{\mathcal{P}}^{\circ} - \boldsymbol{\sigma}_{\mathcal{P}}^{RT}\|_{A^{-1}}$, the oscillation term (34), and the error estimator (the square root of the sum of the indicator and oscillation terms), along with the effectivity indices given by ratios of the indicated and estimated errors to the true error. The same information is shown graphically in Figure 1(b). We observe that the estimator indeed provides a guaranteed upper bound on the true error. Moreover, the oscillation term represents a higher order contribution, as predicted by the theory, and in this particular example is sufficiently small so that the error indicator also provides an upper bound on the true error.

5.2. Two-pole electric motor. We illustrate the quality and performance of the estimator in the case of a two-pole electric motor shown in Figure 2(a) as defined in [14]. The problem is modelled using the equations of magnetostatics

$$\begin{aligned} \operatorname{curl} \boldsymbol{H} &= J, \\ \mathbf{div}\, \boldsymbol{B} &= 0, \end{aligned} \tag{35}$$

where J is the prescribed current density, where $\boldsymbol{B}$ is the unknown magnetic flux density related to the magnetic field $\boldsymbol{H}$ by the constitutive relation $\boldsymbol{B} = \mu \boldsymbol{H}$, and where μ is the magnetic permeability of the material. The current density J is nonzero only in the coils where it takes the value unity. The magnetic permeability is given by $\mu = 5200$ in the stator and rotor and by $\mu = 1$ in the coils and the air gap. We seek a solution varying only in the plane by using the second equation in (35) to write $\boldsymbol{B} = \mathbf{curl}\, u$, where u is a scalar magnetic potential, leading to the equations

$$\begin{aligned} \boldsymbol{H} - \mu^{-1}\, \mathbf{curl}\, u &= \mathbf{0}, \\ -\operatorname{curl} \boldsymbol{H} + J &= 0. \end{aligned}$$



(a) (b)

FIG. 1. (a) *Initial mesh and* (b) *performance of the estimator and indicator for the singular Poisson problem.*

TABLE 1
Performance of the estimator and indicator for the singular Poisson problem (with const. = true error $*\sqrt{Ndof}$*).*

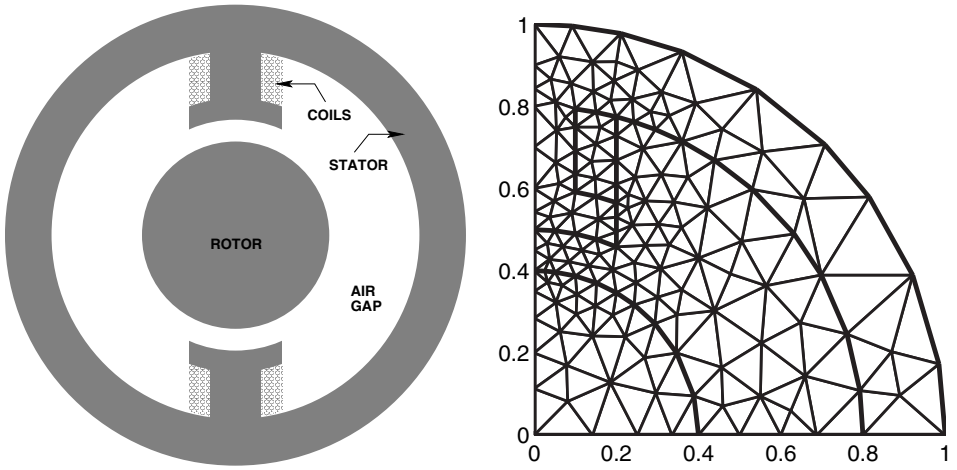
Ndof	True error	Indic.	Effect.	Estim.	Effect.	Oscill.	Const.
32	5.07(−1)	6.41(−1)	1.27	7.02(−1)	1.39	2.86(−1)	2.87
98	3.08(−1)	3.94(−1)	1.28	4.09(−1)	1.33	1.10(−1)	3.04
122	2.83(−1)	3.58(−1)	1.27	3.73(−1)	1.32	1.04(−1)	3.13
198	2.33(−1)	2.82(−1)	1.21	2.96(−1)	1.27	8.71(−2)	3.28
348	1.71(−1)	2.08(−1)	1.22	2.15(−1)	1.26	5.40(−2)	3.19
462	1.46(−1)	1.75(−1)	1.20	1.79(−1)	1.23	4.09(−2)	3.13
672	1.22(−1)	1.45(−1)	1.19	1.49(−1)	1.21	3.13(−2)	3.17
1170	9.22(−2)	1.11(−1)	1.20	1.13(−1)	1.22	2.07(−2)	3.15
1748	7.54(−2)	9.00(−2)	1.19	9.12(−2)	1.21	1.48(−2)	3.15
2618	6.16(−2)	7.34(−2)	1.19	7.42(−2)	1.20	1.09(−2)	3.15
4242	4.82(−2)	5.87(−2)	1.22	5.92(−2)	1.23	7.60(−3)	3.14

In turn, this system may be reduced to the form (1) by setting $\mu = A^{-1}$, $J = f$, and $\boldsymbol{H} = \boldsymbol{\sigma}^{\perp}$. Consequently, we may use the mixed method to approximate the problem. Thanks to symmetry considerations, the computations may be carried out on a quarter domain as shown in Figure 2(b) with homogeneous Neumann data on the horizontal edge and homogeneous Dirichlet data on the remaining edges. The complex geometry and variation of the material mean that true solution of this problem is expected to have corner singularities in the neighborhoods of the material cross points but is unknown analytically.

The data A and f are piecewise constant, meaning that the oscillation terms and the approximation error vanish, with the result that the quantity $\eta_{\mathcal{P}} = \|A\,\mathbf{grad}\,\widetilde{u}^{\circ}_{\mathcal{P}} - \boldsymbol{\sigma}^{RT}_{\mathcal{P}}\|_{A^{-1}}$ provides a guaranteed upper bound on the error in this case. Moreover, the relative path permeability for this distribution of material properties is unity, meaning that the upper bound is robust against variation of the size of μ in the stator and rotor. However, the lack of a closed form solution means that we are unable to compute the effectivity index of the estimator explicitly. Nevertheless, elementary manipulation using (2) and (10) coupled with the fact that $\Pi_{\mathcal{P}} f = f$ shows that

$$0 \leq \mathcal{E}(\boldsymbol{\sigma}, u) - \mathcal{E}(\boldsymbol{\sigma}^{RT}_{\mathcal{P}}, u^{RT}_{\mathcal{P}}) = \|\boldsymbol{\sigma} - \boldsymbol{\sigma}^{RT}_{\mathcal{P}}\|^2_{A^{-1}}, \tag{36}$$

i.e., the familiar result showing that mixed methods underestimate the true energy,



(a) (b)

FIG. 2. (a) *Geometry of a two-pole electric motor and* (b) *initial mesh used to discretize a quarter domain.*

where $\mathcal{E}$ is the potential energy functional defined by

$$\mathcal{E}(\boldsymbol{\tau}, v) = \|\boldsymbol{\tau}\|^2_{A^{-1}} - 2(f, v). \tag{37}$$

The identity (36) means that if the potential energy of the true solution were available, then the true error could be calculated and in turn the effectivity index of the estimator could be obtained. Unfortunately, the true potential energy is not known explicitly either. However, exploiting the fact that the estimator $\eta_{\mathcal{P}}$ is a guaranteed upper bound on the true error, we obtain a guaranteed enclosure on the value of the true potential energy in terms of the approximate potential and the estimator:

$$\mathcal{E}(\boldsymbol{\sigma}^{RT}_{\mathcal{P}}, u^{RT}_{\mathcal{P}}) \le \mathcal{E}(\boldsymbol{\sigma}, u) \le \mathcal{E}(\boldsymbol{\sigma}^{RT}_{\mathcal{P}}, u^{RT}_{\mathcal{P}}) + \eta^2_{\mathcal{P}}. \tag{38}$$

By performing a sequence of adaptive refinements and using this result, we are able to calculate guaranteed bounds $\mathcal{E}_L$ and $\mathcal{E}_U$ on the value of the true potential.

Figure 3 shows the enclosures for the true value of the potential energy using the same kind of adaptive refinement procedure that we applied to the singular Poisson problem. The adaptive procedure with up to 120,000 degrees of freedom provides reasonably tight bounds $\mathcal{E}_L$ and $\mathcal{E}_U$ on the value of the true potential.

At the end of the adaptive refinement procedure, we may utilize the bounds $\mathcal{E}_L$ and $\mathcal{E}_U$ to obtain guaranteed enclosures on the effectivity index $\theta_{\mathcal{P}}$ of the estimator obtained during the computation:

$$\frac{\eta_{\mathcal{P}}}{\sqrt{\mathcal{E}_U - \mathcal{E}(\boldsymbol{\sigma}^{RT}_{\mathcal{P}}, u^{RT}_{\mathcal{P}})}} \le \theta_{\mathcal{P}} \le \frac{\eta_{\mathcal{P}}}{\sqrt{\mathcal{E}_L - \mathcal{E}(\boldsymbol{\sigma}^{RT}_{\mathcal{P}}, u^{RT}_{\mathcal{P}})}}.$$

A simpler approach would be to calculate an approximation to the error by comparing with a finite element approximation obtained using a highly refined mesh. However, this kind of approach would not enable us to deduce guaranteed bounds on the value of the effectivity index that we obtain using the method described above.

The variation of these enclosures and the "true" value, which we take to be the average of the upper and lower bounds, during the adaptive computation is shown

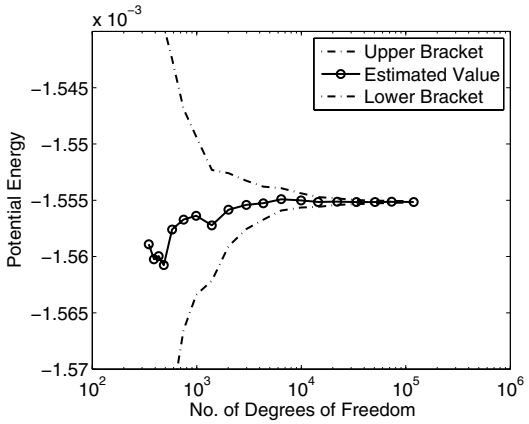


FIG. 3. *Enclosures on the true potential energy obtained by solving a two-pole electric motor using an adaptive algorithm.*

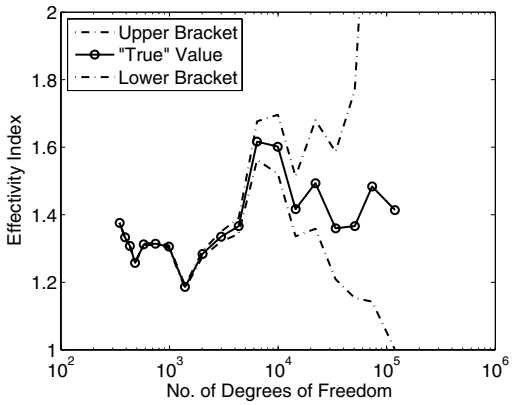


FIG. 4. *Estimated effectivity index and enclosures on actual value obtained by solving a two-pole electric motor using an adaptive algorithm.*

in Figure 4. It is observed that the enclosures on the effectivity index are rather tight for the coarser meshes but begin to open as the accuracy of the discretization error approaches that of the upper and lower bounds on the true potential energy. This is to be expected and merely reflects inaccurate knowledge of the true potential rather than any possible deficiency of the estimator. Nonetheless, the results obtained indicate that the true effectivity index remains bounded of the same order as the results obtained in the singular Poisson problem, despite the large variation of material coefficients, multiple singularities in the solution, and complex geometry.

Finally, in Figure 5, we illustrate the convergence rate of the adaptive procedure by showing the values of the estimator obtained during the adaptive refinement procedure. Ideally, we would also like to show that actual error for purposes of comparison, but of course it is unavailable. Instead, we show the values of $\mathcal{E}_U - \mathcal{E}(\boldsymbol{\sigma}_{\mathcal{P}}^{RT}, u_{\mathcal{P}}^{RT})$, $\mathcal{E}_L - \mathcal{E}(\boldsymbol{\sigma}_{\mathcal{P}}^{RT}, u_{\mathcal{P}}^{RT})$ and their average, dubbed the "true" value, during the refinement. Once again, we see that the estimator provides very satisfactory guaranteed bounds on the true error.

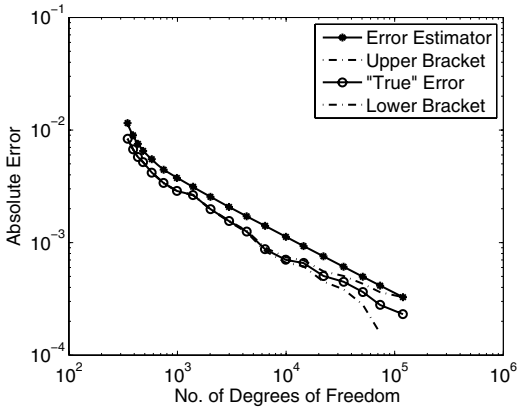


FIG. 5. *Convergence of an adaptive algorithm for a two-pole electric motor.*

6. Proof of the results. First, we give the proof of Theorem 2.

Proof. Equation (22) gives

$$(\boldsymbol{\sigma} - \boldsymbol{\sigma}^\star, \mathbf{curl}\, w)_{A^{-1}} = 0 \quad \forall w \in H^1(\Omega),$$

and hence there exists $\varphi \in H_0^1(\Omega)$ such that $A^{-1}(\boldsymbol{\sigma} - \boldsymbol{\sigma}^\star) = \mathbf{grad}\, \varphi$. The second part of (11) implies that $-\mathbf{div}\, \boldsymbol{\sigma}^\star = \Pi_{\mathcal{P}} f$, so that

$$(\boldsymbol{\sigma} - \boldsymbol{\sigma}^\star, \mathbf{grad}\, v) = (f - \Pi_{\mathcal{P}} f, v) \quad \forall v \in H_0^1(\Omega) \tag{39}$$

and, in particular,

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}^\star\|_{A^{-1}}^2 = (\boldsymbol{\sigma} - \boldsymbol{\sigma}^\star, \mathbf{grad}\, \varphi) = (f - \Pi_{\mathcal{P}} f, \varphi) = \sum_{K \in \mathcal{P}} (f - f_K, \varphi)_K.$$

This terms appearing under the summation may be bounded with the aid of the Poincaré inequality [16] as follows:

$$\begin{aligned}(f - f_K, \varphi)_K &= \inf_{c \in \mathbb{R}} (f - f_K, \varphi - c)_K \le \|f - f_K\|_K \inf_{c \in \mathbb{R}} \|\varphi - c\|_K \\ &\le \pi^{-1} \operatorname{diam}(K) \|f - f_K\|_K \|\mathbf{grad}\, \varphi\|_K \\ &\le \pi^{-1} \operatorname{diam}(K)\, \lambda_K^{-1/2} \operatorname{osc}(f, K) \|\mathbf{grad}\, \varphi\|_{A,K}.\end{aligned}$$

The upper bound then follows after using a Cauchy inequality and cancelling a factor $\|\mathbf{grad}\, \varphi\|_A = \|\boldsymbol{\sigma} - \boldsymbol{\sigma}^\star\|_{A^{-1}}$. The estimate for the oscillation follows at once from the Poincaré inequality. ∎

The proof of the upper bound in Theorem 3 relies on the following characterization.

LEMMA 1.

$$\|\boldsymbol{\sigma}^\star - \boldsymbol{\sigma}_{\mathcal{P}}^{RT}\|_{A^{-1}} = \min_{v \in H_0^1(\Omega)} \|A\, \mathbf{grad}\, v - \boldsymbol{\sigma}_{\mathcal{P}}^{RT}\|_{A^{-1}}. \tag{40}$$

Proof. Let $\boldsymbol{\tau} \in H(\mathbf{div}; \Omega)$ be divergence-free; then using the first part of (11) we obtain, for arbitrary $v \in H_0^1(\Omega)$,

$$(\boldsymbol{\sigma}^\star - \boldsymbol{\sigma}_{\mathcal{P}}^{RT}, \boldsymbol{\tau})_{A^{-1}} = -(\boldsymbol{\sigma}_{\mathcal{P}}^{RT}, \boldsymbol{\tau})_{A^{-1}} = -(v, \mathbf{div}\, \boldsymbol{\tau}) - (\boldsymbol{\sigma}_{\mathcal{P}}^{RT}, \boldsymbol{\tau})_{A^{-1}}, \tag{41}$$

and integrating the first term by parts gives

$$(\boldsymbol{\sigma}^\star - \boldsymbol{\sigma}_{\mathcal{P}}^{RT}, \boldsymbol{\tau})_{A^{-1}} = (A\,\mathbf{grad}\, v - \boldsymbol{\sigma}_{\mathcal{P}}^{RT}, \boldsymbol{\tau})_{A^{-1}}.$$

Choosing $\boldsymbol{\tau} = \boldsymbol{\sigma}^\star - \boldsymbol{\sigma}_{\mathcal{P}}^{RT}$ (recall (21)) and applying a Cauchy–Schwarz inequality, we obtain

$$\|\boldsymbol{\sigma}^\star - \boldsymbol{\sigma}_{\mathcal{P}}^{RT}\|_{A^{-1}} \le \|A\,\mathbf{grad}\, v - \boldsymbol{\sigma}_{\mathcal{P}}^{RT}\|_{A^{-1}}.$$

The result then follows on observing that equality holds when $v = u^\star$. ∎

The proof of the lower bound in Theorem 3 will require some preparation. First, we show that the discontinuous function $u_{\mathcal{P}}^\circ$ does possess some interelement continuity properties, as shown by the following lemma.

LEMMA 2. *Let $u_{\mathcal{P}}^\circ$ be defined in* (13)*. Then, the average value of $u_{\mathcal{P}}^\circ$ is continuous on each element edge, i.e.*,

$$\int_\gamma [u_{\mathcal{P}}^\circ]\,\mathrm{d}s = 0 \quad \forall \gamma \in \mathcal{E}, \tag{42}$$

where $[u_{\mathcal{P}}^\circ]$ denotes the jump in value across the edge γ.

Proof. By construction, the jump $\left[u_{\mathcal{P}}^{CR}\right]$ in the piecewise affine Crouzeix–Raviart approximation vanishes at the midpoint of an edge γ, and so $\int_\gamma \left[u_{\mathcal{P}}^{CR}\right]\,\mathrm{d}s = 0$. It therefore suffices to show that $\int_\gamma [f\psi]\,\mathrm{d}s$ vanishes on each edge. At first sight, the validity of such a result may appear somewhat surprising since the data f and A are in general discontinuous across element interfaces. Nevertheless, to verify the claim, we introduce the function $\boldsymbol{\phi} \in Q_{\mathcal{P}}$ characterized by the conditions that the normal component $\boldsymbol{\nu}_K \cdot \boldsymbol{\phi} = |K|/|\gamma|$ on γ and vanishes on remaining edges. Since $\mathbf{div}\,\boldsymbol{\phi}$ is constant, we see that

$$\mathbf{div}\,\boldsymbol{\phi} = \frac{1}{|K|}\int_K \mathbf{div}\,\boldsymbol{\phi}\,\mathrm{d}\boldsymbol{x} = \frac{1}{|K|}\int_\gamma \boldsymbol{\nu}_K \cdot \boldsymbol{\phi}\,\mathrm{d}s = 1,$$

and hence $\boldsymbol{\phi}$ is of the form $\boldsymbol{\phi} = \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_K) + \boldsymbol{c}_K$ for a suitable constant vector $\boldsymbol{c}_K$. Now,

$$\frac{|K|}{|\gamma|}\int_\gamma \psi_K\,\mathrm{d}\boldsymbol{x} = \int_{\partial K} \psi_K \boldsymbol{\nu}_K \cdot \boldsymbol{\phi}\,\mathrm{d}\boldsymbol{x} = \int_K \psi_K\,\mathbf{div}\,\boldsymbol{\phi}\,\mathrm{d}\boldsymbol{x} + \int_K \mathbf{grad}\,\psi_K \cdot \boldsymbol{\phi}\,\mathrm{d}\boldsymbol{x},$$

and inserting the value of $\mathbf{div}\,\boldsymbol{\phi}$ shows that the first term may be rewritten as

$$\int_K \psi_K\,\mathrm{d}\boldsymbol{x} = -\frac{1}{2}\int_K (\boldsymbol{x} - \boldsymbol{x}_K)^\top A^{-1}(\boldsymbol{x} - \boldsymbol{x}_K)\,\mathrm{d}\boldsymbol{x},$$

while inserting the value of $\mathbf{grad}\,\psi_K$ shows that the second term simplifies to

$$\int_K (\boldsymbol{x} - \boldsymbol{x}_K)^\top A^{-1}\boldsymbol{\phi}\,\mathrm{d}\boldsymbol{x} = \frac{1}{2}\int_K (\boldsymbol{x} - \boldsymbol{x}_K)^\top A^{-1}(\boldsymbol{x} - \boldsymbol{x}_K)\,\mathrm{d}\boldsymbol{x},$$

since $\boldsymbol{x} - \boldsymbol{x}_K$ is orthogonal to the constant vector $A^{-1}\boldsymbol{c}_K$, and hence $\int_\gamma \psi_K\,\mathrm{d}s$ vanishes on each edge $\gamma \subset \partial K$. The claim now follows at once. ∎

The next result bounds the maximum value of the jump in terms of the approximation error.

LEMMA 3. *Let $\gamma \in \mathcal{E}$ be any edge. Then,*

$$\|[u_{\mathcal{P}}^\circ]\|_{L_\infty(\gamma)} \le C\,\lambda_\gamma^{-1/2}\|\boldsymbol{\sigma}^\star - \boldsymbol{\sigma}_{\mathcal{P}}^{RT}\|_{A^{-1},\tilde{\gamma}}, \tag{43}$$

where $\lambda_\gamma = \min_{K \subset \tilde{\gamma}} \lambda_K$ and C is a positive constant independent of any mesh-size.

Proof. The jump in value $[u^\circ_{\mathcal{P}}]$ across the edge γ is a quadratic polynomial with vanishing average. By a scaling argument, we deduce that there exists a constant C, independent of any mesh size, such that

$$\|[u^\circ_{\mathcal{P}}]\|_{L_\infty(\gamma)} \leq C|\gamma|^{1/2}\|\partial\,[u^\circ_{\mathcal{P}}]\,/\partial s\|_\gamma,$$

where s denotes the arc length on γ. Using (14) gives

$$\frac{\partial}{\partial s}\,[u^\circ_{\mathcal{P}}] = [\boldsymbol{\tau}_\gamma \cdot \mathbf{grad}\, u^\circ_{\mathcal{P}}] = \left[\boldsymbol{\tau}_\gamma \cdot A^{-1}\boldsymbol{\sigma}^{RT}_{\mathcal{P}}\right],$$

where $\boldsymbol{\tau}_\gamma$ denotes the unit tangent vector on γ, so that

$$\|[u^\circ_{\mathcal{P}}]\|_{L_\infty(\gamma)} \leq C|\gamma|^{1/2}\left\|\left[\boldsymbol{\tau}_\gamma \cdot A^{-1}\boldsymbol{\sigma}^{RT}_{\mathcal{P}}\right]\right\|\gamma.$$

The result then follows from Lemma 4. ∎

LEMMA 4. *Let γ denote any edge. Then,*

$$|\gamma|^{1/2}\left\|\left[\boldsymbol{\tau}_\gamma \cdot A^{-1}\boldsymbol{\sigma}^{RT}_{\mathcal{P}}\right]\right\|\gamma \leq C\,\lambda_\gamma^{-1/2}\|\boldsymbol{\sigma}^\star - \boldsymbol{\sigma}^{RT}_{\mathcal{P}}\|_{A^{-1},\widetilde{\gamma}}. \tag{44}$$

Proof. Estimates of this kind are rather standard, but for the sake of completeness we give an indication of the proof. Identity (41) shows that for all divergence-free vector fields $\boldsymbol{\tau}$ there holds

$$(\boldsymbol{\sigma}^{RT}_{\mathcal{P}}, \boldsymbol{\tau})_{A^{-1}} = -(\boldsymbol{\sigma}^\star - \boldsymbol{\sigma}^{RT}_{\mathcal{P}}, \boldsymbol{\tau})_{A^{-1}}.$$

Let β_K denote the usual cubic bubble function supported on element K, and pick $\boldsymbol{\tau} = \mathbf{curl}\,\beta_K$ in the above identity to obtain

$$\operatorname{curl}(A^{-1}\boldsymbol{\sigma}^{RT}_{\mathcal{P}})_{|K}\int_K \beta_K\,\mathrm{d}\boldsymbol{x} = -(\boldsymbol{\sigma}^\star - \boldsymbol{\sigma}^{RT}_{\mathcal{P}}, \mathbf{curl}\,\beta_K)_K.$$

Exploiting the fact that $\int_K \beta_K \approx |K|$ and $\|\mathbf{curl}\,\beta_K\|_K \approx 1$, we obtain

$$\left|\operatorname{curl}(A^{-1}\boldsymbol{\sigma}^{RT}_{\mathcal{P}})_{|K}\right| \leq C|K|^{-1}\,\lambda_K^{-1/2}\|\boldsymbol{\sigma}^\star - \boldsymbol{\sigma}^{RT}_{\mathcal{P}}\|_{A^{-1},K}. \tag{45}$$

Now let $\beta^\pm_\gamma = \lambda_\ell\lambda_r(\lambda_\ell \pm \lambda_r)$, where λ_ℓ and λ_r denote standard piecewise linear hat functions associated with the endpoints of an edge γ. Choosing $\boldsymbol{\tau} = \mathbf{curl}\,\beta^\pm_\gamma$ in the above identity gives

$$\begin{aligned}
&\int_\gamma \beta^\pm_\gamma\left[\boldsymbol{\tau}_\gamma \cdot A^{-1}\boldsymbol{\sigma}^{RT}_{\mathcal{P}}\right]\mathrm{d}s \\
&\quad= (\boldsymbol{\sigma}^\star - \boldsymbol{\sigma}^{RT}_{\mathcal{P}}, \operatorname{curl}\beta^\pm_\gamma)_{A^{-1}} + \sum_{K\subset\widetilde{\gamma}}\int_K \beta^\pm_\gamma\operatorname{curl}(A^{-1}\boldsymbol{\sigma}^{RT}_{\mathcal{P}})\,\mathrm{d}\boldsymbol{x} \\
&\quad\leq C\,\lambda_{\widetilde{\gamma}}^{-1/2}\|\boldsymbol{\sigma}^\star - \boldsymbol{\sigma}^{RT}_{\mathcal{P}}\|_{A^{-1},\widetilde{\gamma}} + C\sum_{K\subset\widetilde{\gamma}}|K|\left|\operatorname{curl}(A^{-1}\boldsymbol{\sigma}^{RT}_{\mathcal{P}})_{|K}\right|.
\end{aligned}$$

Hence, using estimate (45), selecting the appropriate sign in the definition of β_γ, and noting that $\left[\boldsymbol{\tau}_\gamma \cdot A^{-1}\boldsymbol{\sigma}^{RT}_{\mathcal{P}}\right]$ is linear on each edge, we obtain estimates for the linear and constant part of the jump on the edge. Combining these estimates gives the result claimed. ∎

The next result is proved using Lemma 3 using an argument similar to that used in the proof of Lemma 6.3 in [2].

LEMMA 5. *Let $K \in \mathcal{P}$ and $\boldsymbol{x}_n$ denote any node of K; then*

$$\left|\widetilde{u}^{\circ}_{\mathcal{P}}(\boldsymbol{x}_n) - u^{\circ}_{\mathcal{P}|K}(\boldsymbol{x}_n)\right| \leq C\|\boldsymbol{\sigma}^{\star} - \boldsymbol{\sigma}^{RT}_{\mathcal{P}}\|_{A^{-1},\Omega_n} \sum_{K' \subset \Omega_n} \omega^{(n)}_{K'}\, \lambda^{-1/2}_{KK'}, \tag{46}$$

where $\lambda_{KK'}$ denotes the path permeability defined in (5).

Finally, we come to the proof of Theorem 3.

Proof. The upper bound follows at once from Lemma 1 by choosing $v = \widetilde{u}^{\circ}_{\mathcal{P}}$. The lower bound follows from Lemma 5 in the same way that Theorem 6.4 in [2] follows from Lemma 6.3 in [2]. ∎

REFERENCES

[1] B. ACHCHAB, A. AGOUZAL, J. BARANGER, AND J. F. MAITRE, *Estimateur d'erreur a posteriori hiérarchique. Application aux éléments finis mixtes*, Numer. Math., 80 (1998), pp. 159–179.

[2] M. AINSWORTH, *Robust a posteriori error estimation for nonconforming finite element approximation*, SIAM J. Numer. Anal., 42 (2005), pp. 2320–2341.

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

[4] D. N ARNOLD AND F. BREZZI, *Mixed and non-conforming finite element methods: Implementation post-processing and error estimates*, RAIRO Modél. Math. Anal. Numér., 19 (1985), pp. 7–32.

[5] J. BARANGER AND H. EL AMRI, *A posteriori error estimators for mixed finite element approximation of some quasi-Newtonian flows*, Mat. Apl. Comput., 10 (1991), pp. 89–102.

[6] B. BRAESS AND R. VERFÜRTH, *A posteriori error estimators for the Raviart–Thomas element*, SIAM J. Numer. Anal., 33 (1996), pp. 2431–2444.

[7] J. H. BRANDTS, *Superconvergence and a posteriori error estimation for triangular mixed finite elements*, Numer. Math., 68 (1994), pp. 311–324.

[8] F. BREZZI AND M. FORTIN, *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, Berlin, 1991.

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

[10] M. CROUZEIX AND P. A. RAVIART, *Conforming and nonconforming finite element methods for solving the stationary Stokes equations*, Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge, 7 (1973), pp. 33–75.

[11] W. DÖRFLER AND R. H. NOCHETTO, *Small data oscillation implies the saturation assumption*, Numer. Math., 91 (2002), pp. 1–12.

[12] C. LOVADINA AND R. STENBERG, *Energy norm a posteriori error estimates for mixed finite element methods*, Math. Comp., 75 (2006), pp. 1659–1674.

[13] L. D. MARINI, *An inexpensive method for the evaluation of the solution of the lowest order Raviart–Thomas mixed method*, SIAM J. Numer. Anal., 22 (1985), pp. 493–496.

[14] MATLAB *PDE Toolbox*, The MathWorks, Inc., Natick, MA, 1996.

[15] S. NICAISE AND E. CREUSÉ, *Isotropic and anisotropic a posteriori error estimation of the mixed finite element method for second order operators in divergence form*, Electron. Trans. Numer. Anal., 23 (2006), pp. 38–62.

[16] L. E. PAYNE AND H. F. WEINBERGER, *An optimal Poincaré inequality for convex domains*, Arch. Rational Mech. Anal., 5 (1960), pp. 286–292.

[17] P. A. RAVIART AND J. M. THOMAS, *Primal hybrid finite element methods for 2nd order elliptic equations*, Math. Comp., 31 (1977), pp. 391–413.

[18] R. STENBERG, *Some new families of finite elements for the Stokes equations*, Numer. Math., 56 (1990), pp. 827–838.

[19] R. STENBERG, *Postprocessing schemes for some mixed finite elements*, RAIRO Modél. Math. Anal. Numér., 25 (1991), pp. 151–167.

[20] B. I. WOHLMUTH AND R. H. W. HOPPE, *A comparison of a posteriori error estimators for mixed finite element discretizations by Raviart-Thomas elements*, Math. Comp., 68 (1999), pp. 1347–1378.