REMOVING THE SATURATION ASSUMPTION IN A POSTERIORI ERROR ANALYSIS*

RICARDO H. NOCHETTO†

Nota presentata dal m.e. Enrico Magenes nella seduta del 10/6/93

Sunto. Si considera un problema ellittico con coefficienti variabili definito su un dominio poligonale. Si dimostra, senza richiedere la "ipotesi di saturazione", che l'estimatore a posteriori dell'errore di Bank e Weiser per elementi finiti lineari a tratti è equivalente in norma dell'energia all'errore del metodo degli elementi finiti. Tale estimatore si ottiene dalla risoluzione di problemi di Neumann su ogni elemento. Il presente risultato viene dimostrato con una regolarità minimale dei dati, ma non richiede nessuna regolarità della soluzione ed è valido anche per coefficienti discontinui.

1. Introduction. A posteriori error estimators provide the basis for adaptive mesh refinement and quantitative error control. The ultimate goal is often to equidistribute the local discretization error, typically in the energy norm, via a proper use of information extracted from both the computed solution and data. This can be phrased in terms of optimizing the computational effort for a given accuracy. One of the most successful estimators, implemented in [2], is due to Bank and Weiser [4]. It is obtained from the solution of local Neumann problems involving both a volumetric and jump (or line) residual of the finite element solution. It is inexpensive as compared with the solution process and performed elementwise, thereby leading to a computable estimate of the error.

Even though asymptotic exactness is a desirable property, it is known to require geometric mesh constraints related to superconvergence that are rarely met in applications [6]. Global equivalence between estimators and the true error is instead a more realistic property to aim at. It guarantees *reliability* and *efficiency* of associated mesh refinement algorithms [8]. Equivalence for the Bank-Weiser's estimator, hereafter called BW, has been derived in the energy norm under the so-called *saturation assumption* [3,4]. This says that the solution can be approximated asymptotically better with quadratic than with linear finite elements. Although such a claim is true for functions smoother than those in $H^2(\Omega)$, it may not be valid for those intermediate between $H^1(\Omega)$ and $H^2(\Omega)$. This is the case of reentrant corners [7]. The existing theory therefore does not fully explain the success of BW in a number of practical situations [2]. In this paper we prove the equivalence between BW and the true error, for piecewise linear finite elements, under minimal regularity of the data and no regularity of the continuous solution. We thus conclude that the saturation assumption is superfluous. The proof extends to the energy norm the pointwise argument discussed in [9].

---

*This work was partially supported by NSF Grant DMS-9305935.

†Department of Mathematics and Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742, USA.

This paper is organized as follows. In §2 we introduce both the continuous and discrete problems along with the assumptions on the data. We also recall an auxiliary estimator, essentially due to Babuška and Miller [1], that is shown to be equivalent to BW in §§3, 4. The lower bound is derived in §3 and the most important upper bound is proved in §4. We conclude in §5 with a proof of equivalence for discontinuous coefficients.

2. Assumptions and Result. Despite the problem and results are formulated in 2D they extend to any dimensions. They also easily extend to the case of systems that are coercive, and so to linear elasticity.

We start with the assumptions on the data. Let f be a forcing term (or load), $\mathbf{A}$ be a 2×2 matrix, and $\Omega \subset \mathbf{R}^2$ be a domain satisfying

$$f \in L^2(\Omega), \tag{2.1}$$

$$\mathbf{A} \in [W^{1,\infty}(\Omega)]^{2\times 2} \quad \text{is positive definite,} \tag{2.2}$$

$$\Omega \text{ is a bounded polygon, possibly with slit boundary.} \tag{2.3}$$

In §5 we weaken (2.2) by allowing discontinuities of $\mathbf{A}$ across interelement boundaries.

For a generic domain Ξ and surface Λ we will use the notation

$$\langle \varphi, \chi \rangle_\Xi = \int_\Xi \varphi\chi, \qquad \langle\!\langle \varphi, \chi \rangle\!\rangle_\Lambda = \int_\Lambda \varphi\chi.$$

The u be the solution to the following model scalar elliptic problem with variable coefficients [7]:

$$u \in H_0^1(\Omega): \qquad \langle \mathbf{A}\cdot\nabla u, \nabla\phi\rangle_\Omega = \langle f, \phi\rangle_\Omega \qquad \forall\ \phi \in H_0^1(\Omega).$$

Let $\mathcal{T}$ be a *graded* triangulation of Ω into triangles. Let $\mathcal{T}$ satisfy [5]:

$$\mathcal{T} \text{ is shape regular,} \tag{2.4}$$

We denote by h_T the size of a generic element $T \in \mathcal{T}$ and by $h_\mathcal{T} := \max_{T\in\mathcal{T}} h_T$ the size of the partition $\mathcal{T}$. We also indicate by $\mathcal{S}$ the set of internal interelement boundaries (or sides) S, by h_S the length of $S \in \mathcal{S}$, and by $\mathcal{N} := \{\mathbf{x}_i\}_{i=1}^I$ the set of internal nodes of $\mathcal{T}$. Let $\mathcal{V}_\mathcal{T}^k$ denote the space of continuous piecewise polynomials over $\mathcal{T}$ of degree $\leq k$. Let $u_\mathcal{T}$ be the piecewise linear finite element approximation of u, that is the solution to

$$u_\mathcal{T} \in \mathcal{V}_\mathcal{T}^1: \qquad \langle \mathbf{A}\cdot\nabla u_\mathcal{T}, \nabla\chi\rangle_\Omega = \langle f, \chi\rangle_\Omega \qquad \forall\ \chi \in \mathcal{V}_\mathcal{T}^1.$$

We assume that the error $e_\mathcal{T} := u - u_\mathcal{T}$ is proper, namely that the following *nondegeneracy* property holds

$$\|\nabla e_\mathcal{T}\|_{L^2(\Omega)} \geq C h_\mathcal{T}. \tag{2.5}$$

This is guaranteed, for instance, if $|D_{ii}u(\mathbf{x})| \geq C > 0$ for all $\mathbf{x}$ in a fixed region $D \subset \Omega$ where the local meshsize $h_T \geq C h_{\mathcal{T}}$ [1]. Therefore this assumption is not very restrictive in practice.

In order to write the *error equation* we first perform an elementwise integration by parts to arrive at

$$0 = \langle \mathbf{A}\cdot\nabla e_{\mathcal{T}}, \nabla\chi\rangle_\Omega = \sum_{T\in\mathcal{T}} \langle R, \chi\rangle_T + \sum_{S\in\mathcal{S}} \langle\!\langle J, \chi\rangle\!\rangle_S = \sum_{T\in\mathcal{T}} \Big(\langle R, \chi\rangle_T + \tfrac{1}{2}\langle\!\langle J, \chi\rangle\!\rangle_{\partial T} \Big) \tag{2.6}$$

for all $\chi \in \mathcal{V}^1_{\mathcal{T}}$. Hereafter $\mathbf{div}\,\mathbf{A}$ denotes the 2-vector whose entries are the divergence of the corresponding columns of $\mathbf{A}$, R indicates the volumetric residual

$$R := f + \mathbf{div}\,\mathbf{A}\cdot\nabla u_{\mathcal{T}}, \tag{2.7}$$

and J stands for the so-called jump residual

$$J = [\![\mathbf{A}\cdot\nabla u_{\mathcal{T}}]\!]\cdot\mathbf{n}_S \qquad \forall\ S \in \mathcal{S}; \tag{2.8}$$

$\mathbf{n}_S$ is a unit normal vector to S and $[\![\cdot]\!]$ is the associated jump operator. Note that the definition of J is independent of the orientation of $\mathbf{n}_S$.

Let $\mathcal{P}^0_2(T)$ denote the set of quadratic polynomials in $T \in \mathcal{T}$ that vanish at the vertices of T. Let ω_T be the unique solution of the discrete Neumann problem [2,3,4]:

$$\omega_T \in \mathcal{P}^0_2(T): \quad \langle \mathbf{A}\cdot\nabla\omega_T, \nabla\varphi\rangle_T = \langle R, \varphi\rangle_T + \tfrac{1}{2}\langle\!\langle J, \varphi\rangle\!\rangle_{\partial T} \quad \forall\ \varphi \in \mathcal{P}^0_2(T). \tag{2.9}$$

We stress that the choice of elliptic operator in (2.9) is arbitrary, and that we could have taken $\mathbf{A} = \mathbf{I}$ at this stage. The corresponding a posteriori error estimator BW is given by

$$\mathcal{E}_1(\mathcal{T}) := \left(\sum_{T\in\mathcal{T}} \|\nabla\omega_T\|^2_{L^2(T)} \right)^{1/2}. \tag{2.10}$$

THEOREM 1. *Let* (2.1) *to* (2.5) *hold. Then there exist constants* $C_1, C_2, h^* > 0$ *depending on the data* $f, \mathbf{A}, \Omega$, *and the minimum angle of* $\mathcal{T}$ *but otherwise independent of* u *and* $\mathcal{T}$ *satisfying*

$$C_1\mathcal{E}_1(\mathcal{T}) \leq \|\nabla e_{\mathcal{T}}\|_{L^2(\Omega)} \leq C_2\mathcal{E}_1(\mathcal{T}) \qquad \forall\ h_{\mathcal{T}} \leq h^*. \tag{2.11}$$

This result entails *minimal* regularity on $f, \mathbf{A}$, and Ω, and no regularity whatsoever on u beyond H^1, which is a major improvement with respect to [3,4]. In fact the *saturation assumption* of [4], that requires the existence of $\beta_{\mathcal{T}} \to 0$ as $h_{\mathcal{T}} \to 0$ such that

$$\inf_{\chi\in\mathcal{V}^2_{\mathcal{T}}} \Big(\|\nabla(u-\chi)\|^2_{L^2(\Omega)} + \sum_{S\in\mathcal{S}} h_S \|[\![\nabla(u-\chi)]\!]\cdot\mathbf{n}_S\|^2_{L^2(S)} \Big)^{1/2} \leq \beta_{\mathcal{T}} \inf_{\chi\in\mathcal{V}^1_{\mathcal{T}}} \|\nabla(u-\chi)\|_{L^2(\Omega)},$$

is a hidden regularity assumption on u. Unless u is smoother than a function in $H^2(\Omega)$, which excludes the important case of reentrant corners, such a $\beta_{\mathcal{T}}$ may not exist. The existing theory therefore does not fully explain the efficiency of BW observed in practice.

The proof of (2.11) is indirect and based on comparison with an auxiliary estimator due essentially to Babuška and Miller [1]. Such an estimator was originally introduced and analyzed for graded rectangular meshes with irregular nodes. However both its formulation and analysis carry over to triangular meshes. Perhaps a disadvantage of our proof is that it does not provide an accurate estimate of the constants C_1 and C_2 nor of the threshold h^*. Let

$$\mathcal{E}_2(\mathcal{T}) := \left(\sum_{T\in\mathcal{T}} h_T^2 \|R\|_{L^2(T)}^2 + \sum_{S\in\mathcal{S}} h_S \|J\|_{L^2(S)}^2 \right)^{1/2} \tag{2.12}$$

be the full estimator involving both the volumetric and jump residual, R and J, and let

$$\mathcal{E}_3(\mathcal{T}) := \left(\sum_{S\in\mathcal{S}} h_S \|J\|_{L^2(S)}^2 \right)^{1/2} \tag{2.13}$$

be the simplified one. Then, assuming (2.1) to (2.5), there exist constants $C_3, C_4, h^\# > 0$ depending on $f, \mathbf{A}, \Omega$, and the minimum angle of $\mathcal{T}$ but not on u nor on $\mathcal{T}$ such that [1]:

$$C_3 \|\nabla e_{\mathcal{T}}\|_{L^2(\Omega)} \le \mathcal{E}_2(\mathcal{T}), \mathcal{E}_3(\mathcal{T}) \le C_4 \|\nabla e_{\mathcal{T}}\|_{L^2(\Omega)} \qquad \forall\ h_{\mathcal{T}} \le h^\#. \tag{2.14}$$

The fact that $h^\#$ is independent of u is a consequence of the argument in step 5 of §4. In §5 we will show how to extend (2.11) and (2.14) to the case of discontinuous coefficients.

3. Lower Bound. We first prove the simplest estimate, which turns out to be the lower bound

$$C_1 \mathcal{E}_1(\mathcal{T}) \le \|\nabla e_{\mathcal{T}}\|_{L^2(\Omega)} \qquad \forall\ h_{\mathcal{T}} \le h^*. \tag{3.1}$$

To this end, we observe that any function $\varphi \in \mathcal{P}_2^0(T)$, being quadratic and zero at the vertices of $T \in \mathcal{T}$, satisfies

$$\|\varphi\|_{L^2(T)} + h_T^{1/2} \|\varphi\|_{L^2(\partial T)} \le C h_T \|\nabla\varphi\|_{L^2(T)}. \tag{3.2}$$

This is a simple exercise in interpolation theory [5]. Select now $\varphi = \omega_T$ in (2.9) and use (2.2) together with (3.2) to arrive at

$$\begin{aligned} C\|\nabla\omega_T\|_{L^2(T)}^2 &\le \langle \mathbf{A}\cdot\nabla\omega_T, \nabla\omega_T\rangle_T = \langle R, \omega_T\rangle_T + \tfrac{1}{2}\langle\langle J, \omega_T\rangle\rangle_{\partial T} \\ &\le C\|\nabla\omega_T\|_{L^2(T)} \left(h_T\|R\|_{L^2(T)} + h_T^{1/2}\|J\|_{L^2(\partial T)} \right). \end{aligned}$$

Summation on $T \in \mathcal{T}$, in conjunction with (2.12), yields $\mathcal{E}_1(\mathcal{T}) \le C\mathcal{E}_2(\mathcal{T})$. Therefore (3.1) follows from (2.14).

4. Upper Bound. The following is the most delicate and essentially new estimate:

$$\|\nabla e_{\mathcal{T}}\|_{L^2(T)} \leq C_2 \mathcal{E}_1(\mathcal{T}) \qquad \forall\ h_{\mathcal{T}} \leq h^*.$$

A similar result but for the maximum norm was obtained in [9], provided $\mathbf{A} = \mathbf{I}$. The present proof is indeed an adaptation to the energy norm of that given in [9], and consist of several steps. In view of (2.5) and (2.14), it suffices to demonstrate

$$\sum_{S\in\mathcal{S}} h_S \|J\|^2_{L^2(S)} \leq C \sum_{T\in\mathcal{T}} \|\nabla\omega_T\|^2_{L^2(T)} + o(h^2_{\mathcal{T}}). \tag{4.1}$$

The idea is to construct an explicit piecewise quadratic function that *locally* links jump residuals J with BW indicators ω_T. This is feasible at the expense of having asymptotically small terms plus others involving R and J, that can be handled via (2.6). It turns out then to be essential the substitution of piecewise quadratic integrands by piecewise linears and viceversa, which in turn requires approximating both R and J by piecewise constants R_i and J_i.

Let us now introduce some further notation. For each $\mathbf{x}_i \in \mathcal{N}$ let Ω_i indicate the union of all elements T and Λ_i the union of all sides S sharing $\mathbf{x}_i$. The cardinal of those sets of triangles and sides is bounded independently of i and $\mathcal{T}$, owing to (2.4). Let h_i denote the size of Ω_i and ϕ_i the canonical function of $\mathcal{V}^1_{\mathcal{T}}$ verifying $\phi_i(\mathbf{x}_j) = \delta_{ij}$; thus $\Omega_i = \text{supp}\ \phi_i$. For each $S \in \mathcal{S}$ let $\varphi_S \in H^1_0(\Omega)$ be the piecewise quadratic function that attains the value 1 at the midpoint $\mathbf{x}_S$ of S and vanishes at all nodes and remaining midpoints of $\mathcal{T}$.

1. To establish a *local* version of (4.1) we resort to a technique due to Verfürth [11]. Let J_i be the piecewise constant approximation of J on Λ_i defined by

$$J_i|_S = J_S := [\![\mathbf{A}(\mathbf{x}_S)\cdot\nabla u_{\mathcal{T}}]\!]_S\cdot\mathbf{n}_S = J(\mathbf{x}_S) \qquad \forall S\subset\Lambda_i. \tag{4.2}$$

For each node $\mathbf{x}_i$ consider the auxiliary (continuous) function $\varphi_i = \sum_{S\subset\Lambda_i} \beta_S\varphi_S \in H^1_0(\Omega_i)$ where $\beta_S \in \mathbf{R}$ satisfies

$$\beta_S \langle\!\langle J_S, \varphi_S\rangle\!\rangle_S = h_S\|J_S\|^2_{L^2(S)} \qquad \forall\ S\subset\Lambda_i. \tag{4.3}$$

Since φ_S is a nonnegative quadratic polynomial with $\varphi_S(\mathbf{x}_S) = 1$ and J_S is constant, we deduce that the integral in (4.3) is equivalent to $h_S^{1/2}\|J_S\|_{L^2(S)}$. This yields $|\beta_S| \leq C h_S^{1/2}\|J_S\|_{L^2(S)}$, which results in

$$\|\nabla\varphi_i\|_{L^2(\Omega_i)} \leq \sum_{S\subset\Lambda_i} |\beta_S|\ \|\nabla\varphi_S\|_{L^2(\Omega_i)} \leq C\left(\sum_{S\subset\Lambda_i} h_S\|J_S\|^2_{L^2(S)}\right)^{1/2}. \tag{4.4}$$

Let R_i denote a convenient constant approximation of R in Ω_i to be specified later. We claim that

$$(4.5)\quad E_i^2 := \sum_{S\subset\Lambda_i} h_S\|J_S\|^2_{L^2(S)} \le Ch_i^2\|R-R_i\|^2_{L^2(\Omega_i)} + C\sum_{T\subset\Omega_i}\|\nabla\omega_T\|^2_{L^2(T)} + \left|\langle R_i,\varphi_i\rangle_{\Omega_i}\right|.$$

In view of (4.3), we can write

$$\begin{aligned}
E_i^2 &= \sum_{S\subset\Lambda_i}\langle\!\langle J_S,\varphi_i\rangle\!\rangle_S\\
&= \sum_{T\subset\Omega_i}\Big(\langle R,\varphi_i\rangle_T + \tfrac{1}{2}\langle\!\langle J,\varphi_i\rangle\!\rangle_{\partial T}\Big) + \langle R_i - R,\varphi_i\rangle_{\Omega_i} - \langle R_i,\varphi_i\rangle_{\Omega_i} + \langle\!\langle J_i - J,\varphi_i\rangle\!\rangle_{\Lambda_i}\\
&= I + II + III + IV,
\end{aligned}$$

and proceed to examine each term separately. In view of (2.9), (3.2), and (4.4), we get

$$\begin{aligned}
I &= \sum_{T\subset\Omega_i}\langle\mathbf{A}\cdot\nabla\omega_T,\nabla\varphi_i\rangle_T\\
&\le C\|\nabla\varphi_i\|_{L^2(\Omega_i)}\left(\sum_{T\subset\Omega_i}\|\nabla\omega_T\|^2_{L^2(T)}\right)^{1/2} \le CE_i\left(\sum_{T\subset\Omega_i}\|\nabla\omega_T\|^2_{L^2(T)}\right)^{1/2}
\end{aligned}$$

and

$$II \le Ch_i\|\nabla\varphi_i\|_{L^2(\Omega_i)}\|R-R_i\|_{L^2(\Omega_i)} \le Ch_iE_i\|R-R_i\|_{L^2(\Omega_i)}.$$

Since $\mathbf{A}$ is both positive definite and Lipschitz continuous, we have the local bound

$$(4.6)\quad \sum_{S\subset\Lambda_i} h_S\|J-J_S\|^2_{L^2(S)} \le \sum_{S\subset\Lambda_i} h_S\|\mathbf{A}-\mathbf{A}(\mathbf{x}_S)\|^2_{L^\infty(S)}\|[\![\nabla u_{\mathcal{T}}]\!]_S\cdot\mathbf{n}_S\|^2_{L^2(S)} \le Ch_i^2E_i^2.$$

We make use again of (3.2) and (4.4), together with (4.6), to estimate IV as follows:

$$IV \le \Big(\sum_{S\subset\Lambda_i} h_S\|J-J_S\|^2_{L^2(S)}\Big)^{1/2}\Big(\sum_{S\subset\Lambda_i} h_S^{-1}\|\varphi_i\|^2_{L^2(S)}\Big)^{1/2} \le Ch_iE_i\|\nabla\varphi_i\|_{L^2(\Omega_i)} \le Ch_iE_i^2.$$

Collecting the above estimates and choosing $h_{\mathcal{T}} \le h^*$ sufficiently small, we easily obtain the asserted inequality (4.5).

2. We now deal with the remaining term III and relate it to the error equation (2.6). To do so we have to convert III into an expression involving ϕ_i, that is to replace the piecewise quadratic function φ_i by the piecewise linear ϕ_i. Since R_i is *constant* in Ω_i, (3.2) implies

$$III \le \left|\langle R_i,\varphi_i\rangle_{\Omega_i}\right| \le Ch_i^2|R_i|\|\nabla\varphi_i\|_{L^2(\Omega_i)} \le CE_i\left|\langle R_i,\phi_i\rangle_{\Omega_i}\right|.$$

Applying (2.6) leads to

$$\langle R_i, \phi_i\rangle_{\Omega_i} = \langle R_i - R, \phi_i\rangle_{\Omega_i} + \langle R, \phi_i\rangle_{\Omega_i} = \langle R_i - R, \phi_i\rangle_{\Omega_i} + \langle\!\langle J_i - J, \phi_i\rangle\!\rangle_{\Lambda_i} - \langle\!\langle J_i, \phi_i\rangle\!\rangle_{\Lambda_i}.$$

We then see, by simply arguing as above with terms II and IV, that

$$\left|\langle R_i, \varphi_i\rangle_{\Omega_i}\right| \le (\delta + Ch_i)E_i^2 + Ch_i^2\|R - R_i\|^2_{L^2(\Omega_i)} + C\left|\langle\!\langle J_i, \phi_i\rangle\!\rangle_{\Lambda_i}\right|^2,$$

where $\delta > 0$ is as small as we wish. Combining this estimate with (4.5), we arrive at

$$\sum_{S\subset\Lambda_i} h_S\|J_S\|^2_{L^2(S)} \le C\sum_{T\subset\Omega_i}\|\nabla\omega_T\|^2_{L^2(T)} + Ch_i^2\|R - R_i\|^2_{L^2(\Omega_i)} + C\left|\langle\!\langle J_i, \phi_i\rangle\!\rangle_{\Lambda_i}\right|^2, \tag{4.7}$$

which is the desired local version of (4.1). The rightmost term is amenable to the error equation (2.6), which in turn relates it with the local BW indicators ω_T. This connection is explored in steps 3 and 4. The middle term, instead, will lead to an asymptotically negligible contribution, as shown in step 5.

3. We intend to express $\left|\langle\!\langle J_i, \phi_i\rangle\!\rangle_{\Lambda_i}\right|^2$ in terms of $\sum_{T\subset\Omega_i}\|\nabla\omega_T\|^2_{L^2(T)}$. We claim that

$$\left|\langle\!\langle J_i, \phi_i\rangle\!\rangle_{\Lambda_i}\right|^2 \le Ch_i^2E_i^2 + Ch_i^2\|R - R_i\|^2_{L^2(\Omega_i)} + C\sum_{T\subset\Omega_i}\|\nabla\omega_T\|^2_{L^2(T)}. \tag{4.8}$$

In contrast to the second step, we would like now to replace ϕ_i by a piecewise quadratic function and thereby use (2.9). Set $\zeta_i := \sum_{S\subset\Lambda_i}\varphi_S \in H_0^1(\Omega_i)$ and write

$$\begin{aligned}\sum_{T\subset\Omega_i}\langle \mathbf{A}\cdot\nabla\omega_T, \nabla\zeta_i\rangle_T &= \sum_{T\subset\Omega_i}\Big(\langle R, \zeta_i\rangle_T + \tfrac{1}{2}\langle\!\langle J, \zeta_i\rangle\!\rangle_{\partial T}\Big)\\ &= \langle R - R_i, \zeta_i\rangle_{\Omega_i} + \langle R_i, \zeta_i\rangle_{\Omega_i} + \langle\!\langle J - J_i, \zeta_i\rangle\!\rangle_{\Lambda_i} + \langle\!\langle J_i, \zeta_i\rangle\!\rangle_{\Lambda_i}\\ &= V + VI + VII + VIII.\end{aligned} \tag{4.9}$$

Invoking the mid-point quadrature rule [5], which is exact for quadratics, we deduce that $\langle\varphi_S, 1\rangle_{\Omega_i} = |T_S|/3$ with T_S being the union of the two elements sharing S, and thus $\langle\zeta_i, 1\rangle_{\Omega_i} = 2|\Omega_i|/3$. Using also that $\langle\phi_i, 1\rangle_{\Omega_i} = |\Omega_i|/3$, we obtain

$$\langle\zeta_i, 1\rangle_{\Omega_i} = 2\langle\phi_i, 1\rangle_{\Omega_i}$$

which, with the aid of (2.6), results in

$$VI = R_i\langle 1, \zeta_i\rangle_{\Omega_i} = 2\langle R_i, \phi_i\rangle_{\Omega_i} = 2\langle R_i - R, \phi_i\rangle_{\Omega_i} + 2\langle\!\langle J_i - J, \phi_i\rangle\!\rangle_{\Lambda_i} - 2\langle\!\langle J_i, \phi_i\rangle\!\rangle_{\Lambda_i}.$$

We can associate the first two terms above with V and VII to give rise to modified expressions for V and VII. They in turn can be bounded as follows by simply reasoning as before:

$$V \le Ch_i\|R - R_i\|_{L^2(\Omega_i)}, \qquad VII \le Ch_iE_i.$$

Replacing back into (4.9), we realize that we still have to tackle the difference $\langle\langle J_i, \zeta_i \rangle\rangle_{\Lambda_i} - 2\langle\langle J_i, \phi_i \rangle\rangle_{\Lambda_i}$. We resort to the Simpson's rule, which is known to be exact for cubic polynomials. Since this is the case for both $J_i\zeta_i$ and $J_i\phi_i$ on each $S \subset \Lambda_i$, we have

$$\begin{aligned}\langle\langle J_i, \zeta_i \rangle\rangle_{\Lambda_i} &= \sum_{S\subset\Lambda_i} \langle\langle J_S, \varphi_S \rangle\rangle_S = \tfrac{2}{3} \sum_{S\subset\Lambda_i} h_S J(\mathbf{x}_S)\\ \langle\langle J_i, \phi_i \rangle\rangle_{\Lambda_i} &= \sum_{S\subset\Lambda_i} \langle\langle J_S, \phi_i \rangle\rangle_S = \tfrac{1}{2} \sum_{S\subset\Lambda_i} h_S J(\mathbf{x}_S),\end{aligned}$$

whence

$$\langle\langle J_i, \zeta_i - 2\phi_i \rangle\rangle_{\Lambda_i} = -\tfrac{2}{3} \langle\langle J_i, \phi_i \rangle\rangle_{\Lambda_i}.$$

Inserting this key equality into (4.9), together with the previous estimates, and recalling that $\|\nabla\zeta_i\|_{L^2(\Omega_i)} \le C$ yields (4.8).

4. Coupling (4.7) and (4.8), adding on $1 \le i \le I$, and using the geometric property that the cardinal of $\mathcal{A}_i := \{j : \Lambda_i \cap \Lambda_j \neq \emptyset\}$ is bounded independently of i and $\mathcal{T}$, we obtain

$$\sum_{S\subset\mathcal{S}} h_S \|J_S\|^2_{L^2(S)} \le C \sum_{T\subset\mathcal{T}} \|\nabla\omega_T\|^2_{L^2(T)} + C \sum_{i=1}^{I} h_i^2 \|R - R_i\|^2_{L^2(\Omega_i)}.$$

Utilizing now (4.6) we can replace J_S by J on the above estimate, which then reads

$$\sum_{S\subset\mathcal{S}} h_S \|J\|^2_{L^2(S)} \le C \sum_{T\subset\mathcal{T}} \|\nabla\omega_T\|^2_{L^2(T)} + C \sum_{i=1}^{I} h_i^2 \|R - R_i\|^2_{L^2(\Omega_i)}.$$

5. To conclude the proof of (4.1) it only remains to select R_i in such a way that

$$\sum_{i=1}^{I} h_i^2 \|R - R_i\|^2_{L^2(\Omega_i)} \le \varepsilon h_{\mathcal{T}}^2, \tag{4.10}$$

for a given $\varepsilon > 0$ and $h_{\mathcal{T}} \le \delta$, where $\delta > 0$ depends on $\varepsilon, f, \mathbf{A}, \Omega$, and the minimum angle of $\mathcal{T}$ but is otherwise independent of u and $\mathcal{T}$. This is to be achieved without any additional conditions on f and $\mathbf{A}$ besides (2.1) and (2.2). To this end, we split R as follows:

$$R = \big(f + \mathbf{div}\,\mathbf{A}\cdot\nabla u\big) - \mathbf{div}\,\mathbf{A}\cdot\nabla e_{\mathcal{T}} =: r - \rho.$$

The convergence of $u_{\mathcal{T}}$ to u in $H_0^1(\Omega)$ is known to be uniform, namely independent of u, for f's which are uniformly bounded in $L^2(\Omega)$ [10]. In view of this, (2.1) and (2.2), given $\varepsilon > 0$ there thus exists $\delta > 0$ as required above such that $\|\rho\|_{L^2(\Omega)} \le \varepsilon$ provided $h_{\mathcal{T}} \le \delta$.

For each $1 \le i \le I$ and $\xi \in L^2(\Omega)$ we define now $\xi_i = P_i(\xi)$ as the L^2-projection of ξ onto the space of constants over Ω_i; thus $\|\xi_i\|_{L^2(\Omega_i)} \le \|\xi\|_{L^2(\Omega_i)}$. It turns out that

$$\sum_{i=1}^{I} h_i^2 \|\rho - \rho_i\|_{L^2(\Omega_i)}^2 \le C h_{\mathcal{T}}^2 \sum_{i=1}^{I} \|\rho\|_{L^2(\Omega_i)}^2 \le C h_{\mathcal{T}}^2 \|\rho\|_{L^2(\Omega)}^2 \le \varepsilon h_{\mathcal{T}}^2,$$

for $h_{\mathcal{T}} \le \delta$ as a consequence of card $\mathcal{A}_i \le C$ being independent of i. This property again, combined now with a simple density argument valid for any $\xi \in L^2(\Omega)$, yields the existence of $\delta > 0$ depending on ε and ξ such that

$$\begin{aligned}\sum_{i=1}^{I} \|\xi - \xi_i\|_{L^2(\Omega_i)}^2 &\le C \sum_{i=1}^{I} \Big(\|\xi - \phi\|_{L^2(\Omega_i)}^2 + \|\phi - P_i(\phi)\|_{L^2(\Omega_i)}^2 + \|P_i(\xi - \phi)\|_{L^2(\Omega_i)}^2 \Big) \\ &\le C\|\xi - \phi\|_{L^2(\Omega)}^2 + C h_{\mathcal{T}}^2 \|\nabla \phi\|_{L^2(\Omega)}^2 \le \varepsilon^2\end{aligned}$$

for $h_{\mathcal{T}} \le \delta$, where ϕ is a suitable smooth approximation of ξ. To complete the argument we have to show, for $\xi = r$, that δ does not depend on u. We note, as in [10], that the set $\mathcal{W}$ of all ∇u's corresponding to uniformly bounded f's in $L^2(\Omega)$ is precompact in $[L^2(\Omega)]^2$. Therefore there exists a finite subset $\{\mathbf{w}_k\}_{k=1}^{K}$ of $\mathcal{W}$ such that K depends only on ε and data, and all elements $\mathbf{w} \in \mathcal{W}$ satisfy $\|\mathbf{w} - \mathbf{w}_k\|_{L^2(\Omega)} \le \varepsilon$ for some k. Finally, on writing

$$r - r_i = (f - f_i) + \mathbf{div}\mathbf{A}\cdot(\nabla u - \mathbf{w}_k) + (\mathbf{div}\mathbf{A}\cdot\mathbf{w}_k - P_i(\mathbf{div}\mathbf{A}\cdot\mathbf{w}_k)) + P_i(\mathbf{div}\mathbf{A}\cdot(\mathbf{w}_k - \nabla u)),$$

using the above estimate for f and $\mathbf{div}\mathbf{A}\cdot\mathbf{w}_k$ for $1 \le k \le K$ in place of ξ, and setting δ to be the minimum of the various δ's involved, we conclude that $\sum_{i=1}^{I} h_i^2 \|r - r_i\|_{L^2(\Omega_i)}^2 \le \varepsilon h_{\mathcal{T}}^2$ for $h_{\mathcal{T}} \le \delta$ independent of u and $\mathcal{T}$. This leads to (4.10), as desired.

The above reasoning to derive (2.11) is applicable even for piecewise constant coefficients $\mathbf{A}$ over $\mathcal{T}$, because (2.14) still holds. The case of discontinuous but variable coefficients is discussed below.

5. Discontinuous Coefficients. Let Γ be a polygonal in Ω that is matched by $\mathcal{T}$, namely Γ is the union of sides of $\mathcal{S}$. Instead of (2.2) we now assume the weaker condition

$$\mathbf{A} \in [W^{1,\infty}(\Omega \backslash \Gamma)]^{2\times 2} \quad \text{is positive definite.} \tag{5.1}$$

In extending (2.11) to this more general situation, we need (2.14) for both $\mathcal{E}_2(\mathcal{T})$ and $\mathcal{E}_3(\mathcal{T})$, which we now prove. It turns out that the following assertion plays a relevant role

$$\sum_{S \in \mathcal{S}} h_S \|J - J_S\|_{L^2(S)}^2 \le o(h_{\mathcal{T}}^2) + C h_{\mathcal{T}}^2 \big(\mathcal{E}_3(\mathcal{T})\big)^2. \tag{5.2}$$

The explicit form (4.2) of J_S changes in case $S \subset \Gamma$, thereby becoming

$$J_S = \big(\mathbf{A}^+(\mathbf{x}_S)\cdot\nabla u_{\mathcal{T}}^+ - \mathbf{A}^-(\mathbf{x}_S)\cdot\nabla u_{\mathcal{T}}^-\big)\cdot\mathbf{n}_S \qquad \forall S \subset \Gamma,$$

where the signs $\pm$ refer to values on both sides of S. Hence, since $u_{\mathcal{T}}$ is piecewise linear,

$$
\begin{aligned}
\|J - J_S\|_{L^2(S)} &\le \|(\mathbf{A}^+ - \mathbf{A}^+(\mathbf{x}_S))\cdot\nabla u_{\mathcal{T}}^+\|_{L^2(S)} \\
&\quad + \|(\mathbf{A}^- - \mathbf{A}^-(\mathbf{x}_S))\cdot\nabla u_{\mathcal{T}}^-\|_{L^2(S)} \le C h_S^{1/2} \|\nabla u_{\mathcal{T}}\|_{L^2(T_S)} \qquad \forall\ S \subset \Gamma.
\end{aligned} \tag{5.3}
$$

The estimates of Babuška and Miller, without any assumptions on $\mathbf{A}$ and f, read [1]

$$
C_1\mathcal{E}_2(\mathcal{T}) - C_3\Big(\sum_{S\in\mathcal{S}} h_S\|J-J_S\|^2_{L^2(S)} + \sum_{T\in\mathcal{T}} h_T^2\|R-R_T\|^2_{L^2(T)}\Big)^{1/2} \le \|\nabla e_{\mathcal{T}}\|_{L^2(\Omega)} \le C_2\mathcal{E}_2(\mathcal{T}),
$$

where R_T stands for a constant approximation of R in T. We then see how to exploit (5.2), in conjunction with an argument similar to that in step 5 above, to show that the factor multiplying C_3 is of order $o(h_{\mathcal{T}}^2)$, and so negligible in account of (2.5). The equivalence (2.14) is thus valid for $\mathcal{E}_2(\mathcal{T})$, along with a lower bound for $\mathcal{E}_3(\mathcal{T})$ because $\mathcal{E}_3(\mathcal{T}) \le \mathcal{E}_2(\mathcal{T})$. The remaining upper bound for $\mathcal{E}_3(\mathcal{T})$ follows from (2.6) which, used as in step 2, yields

$$
\sum_{T\in\mathcal{T}} h_T^2\|R\|^2_{L^2(T)} \le C\sum_{i=1}^{I} h_i^2\|R-R_i\|^2_{L^2(\Omega_i)} + C\sum_{S\in\mathcal{S}} h_S\|J\|^2_{L^2(S)} \le o(h_{\mathcal{T}}^2) + C\mathcal{E}_3(\mathcal{T}).
$$

Note that (4.10) is still valid even though $\mathbf{A}$ is now discontinuous across Γ because no regularity on u, beyond H^1, was necessary in step 5.

With (2.14) at hand we can repeat the proof of (4.1), but this time employing both (4.6) and (5.2) in dealing with terms IV and VII. Therefore to complete the argument we only have to demonstrate (5.2).

Let $\Gamma_{\mathcal{T}}$ denote the union of all triangles touching Γ. We infer that

$$
\text{meas}\ (\Gamma_{\mathcal{T}}) \le C h_{\mathcal{T}}, \tag{5.4}
$$

because Γ is a polygonal. We decompose $\mathcal{S}$ into those sides contained in $\Gamma_{\mathcal{T}}$ and the complement. For the latter (4.6) still holds, whereas for the former we use the absolute continuity of the integral, (5.3), (5.4), and the set $\{\mathbf{w}_k\}_{k=1}^K \subset \mathcal{W}$ introduced in step 5 to deduce

$$
\begin{aligned}
\sum_{S\subset\Gamma} h_S\|J-J_S\|^2_{L^2(S)} &\le C h_{\mathcal{T}}^2\|\nabla u_{\mathcal{T}}\|^2_{L^2(\Gamma_{\mathcal{T}})} \\
&\le C h_{\mathcal{T}}^2 \max_{1\le k\le K}\|\mathbf{w}_k\|^2_{L^2(\Gamma_{\mathcal{T}})} + C h_{\mathcal{T}}^2\|\nabla u - \mathbf{w}_k\|^2_{L^2(\Omega)} + C h_{\mathcal{T}}^2\|\nabla e_{\mathcal{T}}\|^2_{L^2(\Omega)} \le \varepsilon h_{\mathcal{T}}^2,
\end{aligned}
$$

for $h_{\mathcal{T}} \le \delta$ independent of u and $\mathcal{T}$. This concludes the proof of (5.2). We summarize our derivation in the following

THEOREM 2. *If (2.2) is replaced by (5.1), then the assertion of Theorem 1 is valid.*

REFERENCES

[1] I. Babuška and A. Miller, *A feedback finite element method with a posteriori error estimation: Part I. The finite element method and some basic properties of the a posteriori error estimator*, Comput. Methods Appl. Mech. Engrg., 61 (1987), pp. 1–40.

[2] R.E. Bank, *PLTMG: A Software Package for Solving Elliptic Partial Differential Equations, User's Guide 6.0*, Frontiers in Applied Mathematics SIAM, Philadelphia, 1990.

[3] R.E. Bank and B. Velfert, *A posteriori error estimates for the Stokes problem*, SIAM J. Numer. Anal., 28 (1991), pp. 591-623.

[4] R.E. Bank and A. Weiser, *Some a posteriori error estimators for elliptic partial differential equations*, Math. Comp., 44 (1985), pp. 285–301.

[5] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North Holland, Amsterdam, 1978.

[6] R. Durán and R. Rodriguez, *On the asymptotic exactness of Bank-Weiser estimator*, Numer. Math., 62 (1992), pp. 297–303.

[7] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Pitman, Boston, 1985.

[8] C. Johnson, *Adaptive finite element methods for diffusion and convection problems*, Comput. Methods Appl. Mech. Engrg., 82 (1990), pp. 301–322.

[9] R.H. Nochetto, *Pointwise a posteriori error estimators for elliptic problems on highly graded meshes*, (to appear).

[10] A.H. Schatz and J. Wang, *Some new error estimates for Ritz-Galerkin methods with minimal regularity assumptions*, (to appear).

[11] R. Verfürth, *A posteriori error estimators for the Stokes equations*, Numer. Math., 55 (1989), pp. 309–325.