ASYMPTOTICALLY EXACT A POSTERIORI ERROR ESTIMATORS, PART I: GRIDS WITH SUPERCONVERGENCE

RANDOLPH E. BANK* AND JINCHAO XU†

Abstract. In Part I of this work, we develop superconvergence estimates for piecewise linear finite element approximations on quasiuniform triangular meshes where most pairs of triangles sharing a common edge form approximate parallelograms. In particular, we first show a superconvergence of the gradient of the finite element solution u_h and to the gradient of the interpolant u_I. We then analyze a postprocessing gradient recovery scheme, showing that $Q_h \nabla u_h$ is a superconvergent approximation to ∇u. Here Q_h is the global L^2 projection. In Part II, we analyze a superconvergent gradient recovery scheme for general unstructured, shape regular triangulations. This is the foundation for an a posteriori error estimate and local error indicators.

Key words. Superconvergence, gradient recovery.

AMS subject classifications. 65N50, 65N30

1. Introduction. The study of superconvergence and a posteriori error estimates has been an area of active research; see the monographs by Verfürth [17], Chen and Huang [8], Wahlbin [18], Lin and Yan [16], Babuška and Strouboulis [3], and a recent article Lakhany, Marek and Whiteman [13] for overviews of the field. In this two-part work we study some new superconvergence results. In Part I, we develop some superconvergence results for finite element approximations of a general class of elliptic partial differential equations, based mainly on the geometry of the underlying triangular mesh. In Part II, we develop a gradient recovery techniques that can force superconvergence on general shape regular meshes. Patch recovery techniques have been studied by Zienkiewicz and Zhu and has itself evolved into an active subfield of research [25, 14, 23, 24, 9, 22]. Although our algorithm in some respects resembles this and other similar schemes [12, 19, 4, 6, 2, 10], it draws much of its motivation from multilevel iterative methods.

Let $\Omega \subset I\!R^2$ be a bounded domain with Lipschitz boundary $\partial\Omega$. For simplicity of exposition, we assume that Ω is a polygon. We assume that Ω is partitioned by a shape regular triangulation $\mathcal{T}_h$ of mesh size $h \in (0,1)$. Let $\mathcal{V}_h \subset H^1(\Omega)$ be the corresponding continuous piecewise linear finite element space associated with this triangulation $\mathcal{T}_h$, and $u_h \in \mathcal{V}_h$ be a finite elements approximation to a second order elliptic boundary value problem.

Our development has three main steps. In the first step, we prove a superconvergence result for $|u_h - u_I|_{1,\Omega}$, where u_I is the piecewise linear interpolant for u. In particular, we show in Theorem 3.1 that

$$|u_h - u_I|_{1,\Omega} \lesssim h^{1+\min(1,\sigma)} |\log h|^{1/2} \|u\|_{3,\infty,\Omega}. \tag{1.1}$$

Estimate (1.1) holds on quasi-uniform meshes, where an $O(h^2)$ approximate parallelogram property is satisfied for pairs of adjacent triangles in most parts of Ω except for a region of size $O(h^{2\sigma})$); see Section 2 for details.

*Department of Mathematics University of California, San Diego La Jolla, California 92093-0112. email:rbank@ucsd.edu. The work of this author was supported by the National Science Foundation under contracts DMS-9973276 and DMS-0208449.

†Center for Computational Mathematics and Applications, Department of Mathematics, The Pennsylvania State University, University Park, PA 16802. email:xu@math.psu.edu. The work of this author was supported by the National Science Foundation under contract DMS-0074299 and Center for Computational Mathematics and Applications, Penn State University.

The estimate (1.1) is well-known in the literature for the special case $\sigma = \infty$, namely the $O(h^2)$ approximate parallelogram property is satisfied for all pairs of adjacent triangles and it is also known for cases when the $O(h^2)$ approximate parallelogram property is satisfied except for triangles along a few lines, see Xu [20] and Lin and Xu [15] or except for triangles along the domain boundary, see Lin and Yan [16], Hlavacek and Krizek [11]. Lakhany, Marek and Whiteman [13] consider a less restrictive $O(h^{1+\alpha})$ approximate parallelogram property. Our new estimate (1.1) is a significant generalization of these known results. First, our analysis is based on local identities for each element that simplifies existing techniques. For example, our result can be extended in a straightforward fashion to the mesh in which an $O(h^{1+\alpha})$ (instead of $O(h^2)$) approximate parallelogram property holds for most pairs of triangles (see [13]). Second, the assumptions that we make are weaker than existing ones and should hold for many practical grids for some $\sigma > 0$, although in some cases σ could be very small.

One important case that our theory does not cover in this paper is locally refined grids. Lakhany, Marek and Whiteman [13] has some results on this topic for piecewise uniform grids (see also Lin and Xu [15]). Because of the local nature of our analysis, our technique can be extended to this type of grid. We will report this type of extension in future work.

Superconvergence results typically depend on delicate estimates involving cancellation of the lowest order terms in some asymptotic expansion of the local error. When one derives elementwise expressions using continuous finite element spaces, often one encounters boundary integrals involving the normal component of the gradient of the test function. Thus, although one can determine that some cancellation takes place between certain error local components, it is difficult to combine elementwise statements because the normal components of the gradient of $v_h \in \mathcal{V}_h$ are discontinuous. On the other hand, *tangential* components of ∇v_h along element edges are continuous. Thus our approach is to derive some expressions for the element error that involve only the tangential derivative of the test function on the element boundary. The key identity of this type is Lemma 2.3.

We also note that Lemma 2.3 is an identity rather than an estimate. Thus global versions of this identity give exact characterizations of the the error for arbitrary triangulations. In effect, one can see exactly cancellations that might occur even on completely unstructured meshes. The $O(h^2)$ approximate parallelogram property can be viewed in this context as one set of sufficient conditions for obtaining superconvergent bounds for those terms.

The techniques used in our analysis are related to but much more refined than many existing superconvergence techniques in the literature such as those summarized in [8, 16]. For example, the identity in Lemma 2.3 may be compared with the integral identities for rectangular elements [16]. In fact it was not known how the integral identities for rectangular elements in [16] could be generalized to triangular elements. The Lemma 2.3 offers clues for such generalizations and more work can obviously be done in this direction.

The second major component is a superconvergent approximation to ∇u. This approximation is generated by a gradient recovery procedure. In particular, in Theorem 4.2, we show

$$\|\nabla u - Q_h \nabla u_h\|_{0,\Omega} \lesssim h^{1+\min(1,\sigma)} |\log h|^{1/2} \|u\|_{3,\infty,\Omega}. \tag{1.2}$$

where Q_h is the L^2 projection. When the mesh does not satisfy the $O(h^2)$ parallelogram property or σ becomes very close to zero, then the superconvergence demon-

strated in (1.2) will be diminished. Intuitively, it appears that this is due mainly to high frequency errors introduced by the small nonuniformities of the mesh. Preferentially attenuating high frequency errors in mesh functions is of course a widely studied problem in multilevel iterative methods. Our proposal here is to apply these ideas in the present context. Thus, to enhance the superconvergence effect on general shape regular meshes, we compute $S^m Q_h \nabla u_h$, where S is an appropriate multigrid-like smoothing operator. In the second part of this manuscript [7], we analyze this procedure and prove superconvergence estimates somewhat like (1.2) for $\|u - S^m Q_h \nabla u_h\|_{0,\Omega}$.

In the third major component of our analysis, we use the recovered gradient to develop an a posteriori error estimate. An obvious choice is to use $(I - S^m Q_h)\nabla u_h$ to approximate the true error $\nabla(u - u_h)$. In [7], we show this is a good choice, and that in many circumstances we can expect the error estimate to be *asymptotically exact*; that is

$$\lim \frac{\|(I - S^m Q_h)\nabla u_h\|_{0,\Omega}}{\|\nabla(u - u_h)\|_{0,\Omega}} = 1.$$

as $h \to 0$ and $m \to \infty$ in an appropriate fashion.

We also use the recovered gradient to construct local approximations of interpolation errors to be used as local error indicators for adaptive meshing algorithms; see [7] for details.

We remark that both our gradient recovery scheme and our a posteriori error estimate are largely independent of the details of the partial differential equation. Indeed, all of the preliminary lemmas in Section 2 are also independent of the PDE. The PDE directly enters only in the proof of Theorem 3.1, and there the properties we assume are standard. This suggests that superconvergence can be expected for a wide variety of problems, as long as the adaptive meshing yields smoothly varying, shape regular meshes.

The rest of this paper is organized as follows: Section 2 contains technical identities and estimates that form the basis for the estimate (1.1). In Section 3, prove (1.1) for a general linear elliptic boundary value problems under standard assumptions. We also explore an application to nonlinear elliptic problems. In Section 4 we develop and analyze the superconvergent gradient recover scheme in the case of $O(h^2)$ parallelogram meshes. In Section 5 we present a few numerical examples illustrating the effectiveness of our procedures.

2. Preliminary Lemmas. We begin with some geometric identities for a canonical element τ. Let τ have vertices $\boldsymbol{p}_k^t = (x_k, y_k)$, $1 \le k \le 3$, oriented counterclockwise, and corresponding nodal basis functions (barycentric coordinates) $\{\phi_k\}_{k=1}^3$. Let $\{e_k\}_{k=1}^3$ denote the edges of element τ, $\{\theta_k\}_{k=1}^3$ the angles, $\{\boldsymbol{n}_k\}_{k=1}^3$ the unit outward normal vectors, $\{\boldsymbol{t}_k\}_{k=1}^3$ the unit tangent vectors with counterclockwise orientation, $\{\ell_k\}_{k=1}^3$ the edge lengths, and $\{d_k\}_{k=1}^3$ the perpendicular heights (see Fig.1). Let $\tilde{\boldsymbol{p}}$ be the point of intersection for the perpendicular bisectors of the three sides of τ. Let $|s_k|$ denote the distance between $\tilde{\boldsymbol{p}}$ and side k. If τ has no obtuse angles, then the s_k will be nonnegative; otherwise, the distance to the side opposite the obtuse angle will be negative.

There are many relationships among these quantities; in particular we note the

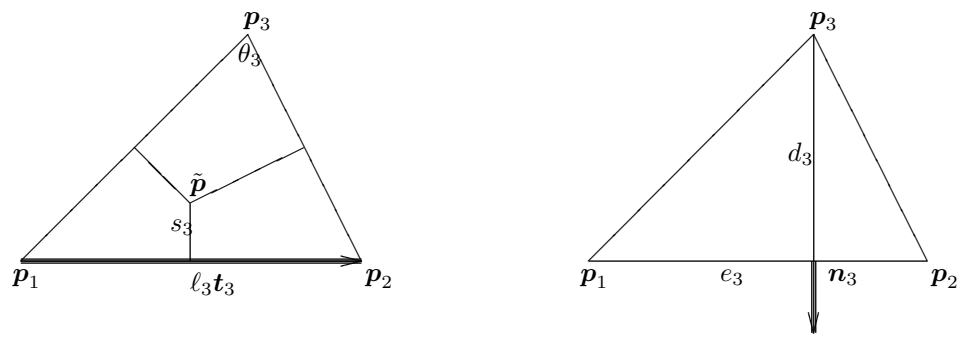

FIG. 2.1. *Parameters associated with the triangle τ*

following, which hold for $1 \le k \le 3$ and $k \pm 1$ permuted cyclically:

$$\begin{aligned}
\ell_k d_k &= \ell_{k+1}\ell_{k-1}\sin\theta_k = 2|\tau|,\\
2\ell_{k+1}\ell_{k-1}\cos\theta_k &= \ell_{k+1}^2 + \ell_{k-1}^2 - \ell_k^2,\\
\sin\theta_k &= \boldsymbol{n}_{k-1}\cdot\boldsymbol{t}_{k+1} = -\boldsymbol{n}_{k+1}\cdot\boldsymbol{t}_{k-1},\\
\cos\theta_k &= -\boldsymbol{t}_{k-1}\cdot\boldsymbol{t}_{k+1} = -\boldsymbol{n}_{k-1}\cdot\boldsymbol{n}_{k+1},\\
\nabla\phi_k &= -\boldsymbol{n}_k/d_k,\\
s_k &= -|\tau|\ \ell_k\nabla\phi_{k-1}\cdot\nabla\phi_{k+1} = \frac{\ell_k\cos\theta_k}{2\sin\theta_k}.
\end{aligned}$$

Let $\mathcal{D}_\tau$ be a symmetric 2×2 matrix with constant matrix entries. We define

$$\xi_k = -\boldsymbol{n}_{k+1}\cdot\mathcal{D}_\tau\boldsymbol{n}_{k-1}.$$

The important special case $\mathcal{D}_\tau = I$ corresponds to $-\Delta$, and in this case $\xi_k = \cos\theta_k$. Let $q_k = \phi_{k+1}\phi_{k-1}$ denote the quadratic bump function associated with edge e_k and let $\psi_k = \phi_k(1-\phi_k)$. In Lemma 2.1, we collect several simple identities that are used in the proof of Lemma 2.3.

LEMMA 2.1.

$$\sin\theta_k\nabla u\cdot\mathcal{D}_\tau\boldsymbol{n}_k = \xi_{k-1}\frac{\partial u}{\partial\boldsymbol{t}_{k-1}} - \xi_{k+1}\frac{\partial u}{\partial\boldsymbol{t}_{k+1}} \tag{2.1}$$

$$\frac{\partial u}{\partial\boldsymbol{t}_{k+1}} = -\cos\theta_{k-1}\frac{\partial u}{\partial\boldsymbol{t}_k} - \sin\theta_{k-1}\frac{\partial u}{\partial\boldsymbol{n}_k} \tag{2.2}$$

$$\frac{\partial u}{\partial\boldsymbol{t}_{k-1}} = -\cos\theta_{k+1}\frac{\partial u}{\partial\boldsymbol{t}_k} + \sin\theta_{k+1}\frac{\partial u}{\partial\boldsymbol{n}_k} \tag{2.3}$$

$$\int_\tau\frac{\partial u}{\partial\boldsymbol{t}_k} = -\sin\theta_{k+1}\int_{e_{k-1}}u + \sin\theta_{k-1}\int_{e_{k+1}}u \tag{2.4}$$

$$\sin\theta_k\int_{e_{k-1}}q_{k-1}u = \int_\tau\psi_{k+1}\frac{\partial u}{\partial\boldsymbol{t}_{k+1}} + \sin\theta_{k-1}\int_{e_k}q_k u \tag{2.5}$$

$$\sin\theta_k\int_{e_{k+1}}q_{k+1}u = -\int_\tau\psi_{k-1}\frac{\partial u}{\partial\boldsymbol{t}_{k-1}} + \sin\theta_{k+1}\int_{e_k}q_k u. \tag{2.6}$$

Proof. We note that (2.1) is an immediate consequence of

$$\mathcal{D}_\tau\boldsymbol{n}_k = \frac{\boldsymbol{n}_{k+1}\cdot\mathcal{D}_\tau\boldsymbol{n}_k}{\boldsymbol{n}_{k+1}\cdot\boldsymbol{t}_{k-1}}\boldsymbol{t}_{k-1} + \frac{\boldsymbol{n}_{k-1}\cdot\mathcal{D}_\tau\boldsymbol{n}_k}{\boldsymbol{n}_{k-1}\cdot\boldsymbol{t}_{k+1}}\boldsymbol{t}_{k+1} = \frac{\xi_{k-1}}{\sin\theta_k}\boldsymbol{t}_{k-1} - \frac{\xi_{k+1}}{\sin\theta_k}\boldsymbol{t}_{k+1}.$$

Proofs for (2.2)-(2.3) follow the same pattern. For (2.4), we note that from Green's Identity

$$\int_\tau \nabla u \cdot \boldsymbol{t}_k = \sum_{j=1}^{3} \boldsymbol{n}_j \cdot \boldsymbol{t}_k . \int_{e_k} u.$$

For (2.5)-(2.6), we note that ψ_k is constant along lines parallel to e_k, and $\partial\psi_k/\partial\boldsymbol{t}_k \equiv 0$. Thus

$$\frac{\partial(\psi_k u)}{\partial \boldsymbol{t}_k} = \psi_k \frac{\partial u}{\partial \boldsymbol{t}_k}.$$

Also, on edge e_k we have $q_k = \psi_{k+1} = \psi_{k-1}$. Equations (2.5)-(2.6) follow from these observations and (2.4). □

LEMMA 2.2. *Let $u \in W^{3,\infty}(\Omega)$. Let u_I and u_q be the continuous piecewise linear and piecewise quadratic interpolants, respectively, for u. Then*

$$\int_{e_k} (u - u_I) = \frac{\ell_k^2}{2} \int_{e_k} q_k \frac{\partial^2 u}{\partial \boldsymbol{t}_k^2}, \tag{2.7}$$

$$\int_\tau (u - u_I) = -\frac{1}{24} \int_\tau \sum_{k=1}^{3} \ell_k^2 \frac{\partial^2 u_q}{\partial \boldsymbol{t}_k^2} + \int_\tau (u - u_q). \tag{2.8}$$

Proof. Identity (2.7) is equivalent to the following:

$$\int_a^b u(s)ds - \frac{(b-a)}{2}(u(a) + u(b)) = \frac{1}{2}\int_a^b (s-a)(s-b)u''(s)ds$$

which follows by an integration by parts. To show, (2.8), we note that $u_q - u_I$ is a piecewise quadratic polynomial that is zero at all of the vertices in the mesh, and therefore can be expressed in terms of the quadratic bump functions. A simple calculations shows in a given element τ

$$u_q - u_I = \sum_{k=1}^{3} \ell_k^2 \boldsymbol{t}_k^t M_\tau \boldsymbol{t}_k \, q_k(x, y), \tag{2.9}$$

where

$$M_\tau = -\frac{1}{2} \begin{pmatrix} \partial_{11} u_q & \partial_{12} u_q \\ \partial_{21} u_q & \partial_{22} u_q \end{pmatrix}. \tag{2.10}$$

The matrix M_τ is constant since u_q is quadratic. Let $m_k = (p_{k+1} + p_{k-1})/2$ denote the midpoint of the k-th edge. Then

$$\frac{\boldsymbol{t}_k^t M_\tau \boldsymbol{t}_k}{2} = \frac{2u(m_k) - u(p_{k+1}) - u(p_{k-1})}{\ell_k^2}.$$

Identity (2.8) follows from

$$\begin{aligned}\int_\tau (u-u_I) &= \int_\tau (u_q-u_I) + \int_\tau (u-u_q)\\ &= -\frac{1}{2}\sum_{k=1}^{3}\ell_k^2\frac{\partial^2 u_q}{\partial \boldsymbol{t}_k^2}\int_\tau q_k + \int_\tau (u-u_q)\\ &= -\frac{|\tau|}{24}\sum_{k=1}^{3}\ell_k^2\frac{\partial^2 u_q}{\partial \boldsymbol{t}_k^2} + \int_\tau (u-u_q)\\ &= -\frac{1}{24}\int_\tau \sum_{k=1}^{3}\ell_k^2\frac{\partial^2 u_q}{\partial \boldsymbol{t}_k^2} + \int_\tau (u-u_q).\end{aligned}$$

□

The following is a fundamental identity in our analysis.

LEMMA 2.3. *Let $\mathcal{D}_\tau$ be a 2×2 symmetric matrix with constant entries. Then*

$$\begin{aligned}\int_\tau \nabla(u-u_I)\cdot\mathcal{D}_\tau\nabla v_h &= \sum_{k=1}^{3}\int_{e_k}\frac{\xi_k q_k}{2\sin\theta_k}\left\{(\ell_{k+1}^2-\ell_{k-1}^2)\frac{\partial^2 u}{\partial \boldsymbol{t}_k^2} + 4|\tau|\frac{\partial^2 u}{\partial \boldsymbol{t}_k\partial \boldsymbol{n}_k}\right\}\frac{\partial v_h}{\partial \boldsymbol{t}_k}\\ &- \int_\tau\sum_{k=1}^{3}\frac{\ell_k\xi_k}{2\sin^2\theta_k}\left\{\ell_{k+1}\psi_{k-1}\frac{\partial^3 u}{\partial^2\boldsymbol{t}_{k+1}\partial\boldsymbol{t}_{k-1}} + \ell_{k-1}\psi_{k+1}\frac{\partial^3 u}{\partial^2\boldsymbol{t}_{k-1}\partial\boldsymbol{t}_{k+1}}\right\}\frac{\partial v_h}{\partial \boldsymbol{t}_k}.\end{aligned}$$

Proof. Using Lemmas 2.1-2.2, we have

$$\begin{aligned}\int_\tau \nabla(u-u_I)\cdot\mathcal{D}_\tau\nabla v_h &= \sum_{k=1}^{3}\int_{e_k}(u-u_I)\nabla v_h v_h\cdot\mathcal{D}_\tau\boldsymbol{n}_k\\ &= \sum_{k=1}^{3}\int_{e_k}(u-u_I)\left\{\frac{\xi_{k-1}}{\sin\theta_k}\frac{\partial v_h}{\partial\boldsymbol{t}_{k-1}} - \frac{\xi_{k+1}}{\sin\theta_k}\frac{\partial v_h}{\partial\boldsymbol{t}_{k+1}}\right\}\\ &= \sum_{k=1}^{3}\left\{\frac{\xi_k}{\sin\theta_{k+1}}\int_{e_{k+1}}(u-u_I)\frac{\partial v_h}{\partial\boldsymbol{t}_k}\right\} - \left\{\frac{\xi_k}{\sin\theta_{k-1}}\int_{e_{k-1}}(u-u_I)\frac{\partial v_h}{\partial\boldsymbol{t}_k}\right\}\\ &= \sum_{k=1}^{3}\left\{\frac{\ell_{k+1}^2\xi_k}{2\sin\theta_{k+1}}\int_{e_{k+1}}q_{k+1}\frac{\partial^2 u}{\partial\boldsymbol{t}_{k+1}^2}\frac{\partial v_h}{\partial\boldsymbol{t}_k}\right\} - \left\{\frac{\ell_{k-1}^2\xi_k}{2\sin\theta_{k-1}}\int_{e_{k-1}}q_{k-1}\frac{\partial^2 u}{\partial\boldsymbol{t}_{k-1}^2}\frac{\partial v_h}{\partial\boldsymbol{t}_k}\right\}\\ &= \sum_{k=1}^{3}\frac{\ell_k\xi_k}{2\sin\theta_k}\left\{\ell_{k+1}\int_{e_{k+1}}q_{k+1}\frac{\partial^2 u}{\partial\boldsymbol{t}_{k+1}^2}\frac{\partial v_h}{\partial\boldsymbol{t}_k} - \ell_{k-1}\int_{e_{k-1}}q_{k-1}\frac{\partial^2 u}{\partial\boldsymbol{t}_{k-1}^2}\frac{\partial v_h}{\partial\boldsymbol{t}_k}\right\}\\ &= \sum_{k=1}^{3}\frac{\xi_k}{2\sin\theta_k}\int_{e_k}q_k\left\{\ell_{k+1}^2\frac{\partial^2 u}{\partial\boldsymbol{t}_{k+1}^2} - \ell_{k-1}^2\frac{\partial^2 u}{\partial\boldsymbol{t}_{k-1}^2}\right\}\frac{\partial v_h}{\partial\boldsymbol{t}_k}\\ &\quad - \int_\tau\sum_{k=1}^{3}\frac{\ell_k\xi_k}{2\sin^2\theta_k}\left\{\ell_{k+1}\psi_{k-1}\frac{\partial^3 u}{\partial\boldsymbol{t}_{k-1}\partial\boldsymbol{t}_{k+1}^2} + \ell_{k-1}\psi_{k+1}\frac{\partial^3 u}{\partial\boldsymbol{t}_{k+1}\partial\boldsymbol{t}_{k-1}^2} +\right\}\frac{\partial v_h}{\partial\boldsymbol{t}_k}.\end{aligned}$$

To complete the proof, we focus attention on the term

$$\ell_{k+1}^2\frac{\partial^2 u}{\partial\boldsymbol{t}_{k+1}^2} - \ell_{k-1}^2\frac{\partial^2 u}{\partial\boldsymbol{t}_{k-1}^2}.$$

Using Lemma 2.1 once again, we have

$$\frac{\partial^2 u}{\partial \boldsymbol{t}_{k+1}^2} = \cos^2\theta_{k-1}\frac{\partial^2 u}{\partial \boldsymbol{t}_k^2} + 2\cos\theta_{k-1}\sin\theta_{k-1}\frac{\partial^2 u}{\partial \boldsymbol{t}_k \partial \boldsymbol{n}_k} + \sin^2\theta_{k-1}\frac{\partial^2 u}{\partial \boldsymbol{n}_k^2},$$
$$\frac{\partial^2 u}{\partial \boldsymbol{t}_{k-1}^2} = \cos^2\theta_{k+1}\frac{\partial^2 u}{\partial \boldsymbol{t}_k^2} - 2\cos\theta_{k+1}\sin\theta_{k+1}\frac{\partial^2 u}{\partial \boldsymbol{t}_k \partial \boldsymbol{n}_k} + \sin^2\theta_{k+1}\frac{\partial^2 u}{\partial \boldsymbol{n}_k^2}.$$

We also need the following identities:

$$\ell_{k+1}^2\sin^2\theta_{k-1} - \ell_{k-1}^2\sin^2\theta_{k+1} = 0,$$
$$\ell_{k+1}^2\cos^2\theta_{k-1} - \ell_{k-1}^2\cos^2\theta_{k+1} = \ell_{k+1}^2 - \ell_{k-1}^2,$$
$$\ell_{k+1}^2 2\cos\theta_{k-1}\sin\theta_{k-1} + \ell_{k-1}^2 2\cos\theta_{k+1}\sin\theta_{k+1} = 4|\tau|.$$

Combining these equations leads to

$$\ell_{k+1}^2\frac{\partial^2 u}{\partial \boldsymbol{t}_{k+1}^2} - \ell_{k-1}^2\frac{\partial^2 u}{\partial \boldsymbol{t}_{k-1}^2} = (\ell_{k+1}^2 - \ell_{k-1}^2)\frac{\partial^2 u}{\partial \boldsymbol{t}_k^2} + 4|\tau|\frac{\partial^2 u}{\partial \boldsymbol{t}_k \partial \boldsymbol{n}_k},$$

completing the proof. □

Let e be an interior edge in the triangulation $\mathcal{T}_h$. Let τ and τ' be the two elements sharing e. We say that τ and τ' form an $O(h^2)$ approximate parallelogram if the lengths of any two opposite edges differ only by $O(h^2)$. Let x be a vertex lying on $\partial\Omega$, and let e and e' be the two boundary edges sharing x as an endpoint. Let τ and τ' be the two elements having e and e', respectively, as edges, and let $\boldsymbol{t}$ and and $\boldsymbol{t}'$ be the unit tangents. Take e and e' as one pair of corresponding edges, and make a clockwise traversal of $\partial\tau$ and $\partial\tau'$ to define two additional corresponding edge pairs. In this case, we say that τ and τ' form an $O(h^2)$ approximate parallelogram if $|\boldsymbol{t} - \boldsymbol{t}'| = O(h)$, and the lengths of any two corresponding edges differ only by $O(h^2)$.

Definition 2.4. *The triangulation $\mathcal{T}_h$ is $O(h^{2\sigma})$ irregular if:*

1. *Let $\mathcal{E} = \mathcal{E}_1 \oplus \mathcal{E}_2$ denote the set of interior edges in $\mathcal{T}_h$. For each $e \in \mathcal{E}_1$, τ and τ' form an $O(h^2)$ approximate parallelogram, while $\sum_{e\in\mathcal{E}_2}|\tau| + |\tau'| = O(h^{2\sigma})$.*
2. *Let $\mathcal{P} = \mathcal{P}_1 \oplus \mathcal{P}_2$ denote the set of boundary vertices. The elements associated with each $x \in \mathcal{P}_1$ form an $O(h^2)$ approximate parallelogram, and $|\mathcal{P}_2| = \kappa$, where κ is fixed independent of h.*

The boundary points $\mathcal{P}$ and the decomposition $\mathcal{P} = \mathcal{P}_1 \oplus \mathcal{P}_2$ are used only in the case of Neumann boundary conditions. Generally speaking, we expect $\mathcal{P}_2$ to consist of the geometric corners of Ω and perhaps a few other isolated points.

We can now state our main Lemma.

Lemma 2.5. *Let the triangulation $\mathcal{T}_h$ be $O(h^{2\sigma})$ irregular. Let $\mathcal{D}_\tau$ be a piecewise constant matrix function defined on $\mathcal{T}_h$, whose elements $\mathcal{D}_{\tau ij}$ satisfy*

$$|\mathcal{D}_{\tau ij}| \lesssim 1,$$
$$|\mathcal{D}_{\tau ij} - \mathcal{D}_{\tau' ij}| \lesssim h,$$

for $i = 1, 2$, $j = 1, 2$. Here τ and τ' are a pair of triangles sharing a common edge. Then

$$\left|\sum_{\tau\in\mathcal{T}_h}\int_\tau \nabla(u - u_I)\cdot\mathcal{D}_\tau\nabla v_h\right| \lesssim h^{1+\min(1,\sigma)}|\log h|^{1/2}\|u\|_{3,\infty,\Omega}|v_h|_{1,\Omega}. \tag{2.11}$$

Proof. Applying Lemma 2.3

$$\sum_{\tau\in\mathcal{T}_h}\int_\tau \nabla(u-u_I)\cdot\mathcal{D}_\tau\nabla v_h = I_1 + I_2 \tag{2.12}$$

where

$$I_1 = \sum_{\tau\in\mathcal{T}_h}\sum_{k=1}^{3}\int_{e_k}\frac{\xi_k q_k}{2\sin\theta_k}\left\{(\ell_{k+1}^2-\ell_{k-1}^2)\frac{\partial^2 u}{\partial \boldsymbol{t}_k^2}+4|\tau|\frac{\partial^2 u}{\partial \boldsymbol{t}_k\partial \boldsymbol{n}_k}\right\}\frac{\partial v_h}{\partial \boldsymbol{t}_k}$$

$$I_2 = -\sum_{\tau\in\mathcal{T}_h}\int_\tau\sum_{k=1}^{3}\frac{\ell_k\xi_k}{2\sin^2\theta_k}\left\{\ell_{k+1}\psi_{k-1}\frac{\partial^3 u}{\partial^2\boldsymbol{t}_{k+1}\partial\boldsymbol{t}_{k-1}}+\ell_{k-1}\psi_{k+1}\frac{\partial^3 u}{\partial^2\boldsymbol{t}_{k-1}\partial\boldsymbol{t}_{k+1}}\right\}\frac{\partial v_h}{\partial \boldsymbol{t}_k}$$

I_2 is easily estimated by

$$|I_2| \lesssim h^2\|u\|_{3,\Omega}|v_h|_{1,\Omega}. \tag{2.13}$$

To estimate I_1, let $\mathcal{E}=\mathcal{E}_1\oplus\mathcal{E}_2$ denote the set of interior edges. For each $e\in\mathcal{E}$, let τ and τ' share e as a common edge. Denote, with respect to τ,

$$\alpha_e = \frac{\xi_k}{2\sin\theta_k}(\ell_{k+1}^2-\ell_{k-1}^2),\quad \beta_e = \frac{\xi_k}{2\sin\theta_k}4|\tau|,$$

and with respect to τ',

$$\alpha_e' = \frac{\xi_{k'}}{2\sin\theta_{k'}}(\ell_{k'+1}^2-\ell_{k'-1}^2),\quad \beta_e' = \frac{\xi_{k'}}{2\sin\theta_{k'}}4|\tau'|.$$

Take $\boldsymbol{n}$ and $\boldsymbol{t}$ to correspond to τ. Then we can write

$$I_1 = I_{11}+I_{12}+I_{13},$$

where

$$I_{1j} = \sum_{e\in\mathcal{E}_j}\int_e q_e\left\{(\alpha_e-\alpha_e')\frac{\partial^2 u}{\partial \boldsymbol{t}^2}+(\beta_e-\beta_e')\frac{\partial^2 u}{\partial\boldsymbol{t}\partial\boldsymbol{n}}\right\}\frac{\partial v_h}{\partial \boldsymbol{t}}$$

for $j=1,2$, and

$$I_{13} = \sum_{e\subset\partial\Omega}\int_e q_e\left\{\alpha_e\frac{\partial^2 u}{\partial \boldsymbol{t}^2}+\beta_e\frac{\partial^2 u}{\partial\boldsymbol{t}\partial\boldsymbol{n}}\right\}\frac{\partial v_h}{\partial \boldsymbol{t}}.$$

Using the elementary identity

$$\left|\int_e f\right| \lesssim h^{-1}\int_\tau |f| + \int_\tau |\nabla f|,$$

we obtain (for $\boldsymbol{z}=\boldsymbol{t}$ and $\boldsymbol{z}=\boldsymbol{n}$)

$$\left|\int_e q_e\frac{\partial^2 u}{\partial\boldsymbol{t}\partial\boldsymbol{z}}\frac{\partial v_h}{\partial\boldsymbol{t}}\right| \lesssim h^{-1}\int_\tau|\nabla^2 u||\nabla v_h| + \int_\tau|\nabla^3 u||\nabla v_h|. \tag{2.14}$$

We can estimate this term in a slightly different way:

$$\left|\int_e q_e \frac{\partial^2 u}{\partial \boldsymbol{t}\partial \boldsymbol{z}} \frac{\partial v_h}{\partial \boldsymbol{t}}\right| \lesssim h^{-1}|u|_{2,\infty,\Omega} \int_\tau |\nabla v_h|. \tag{2.15}$$

For $e \in \mathcal{E}_1$,

$$|\alpha_e - \alpha_e'| \lesssim h^3,$$
$$|\beta_e - \beta_e'| \lesssim h^3.$$

Combining this with (2.14), we have

$$|I_{11}| \lesssim h^2 \int_\Omega (|\nabla^2 u| + h\nabla^3 u|)|\nabla v_h| \lesssim h^2 \|u\|_{3,\Omega}|v_h|_{1,\Omega}, \tag{2.16}$$

or, by (2.15), we have

$$|I_{11}| \lesssim h^2 |u|_{2,\infty,\Omega}|v_h|_{1,\Omega}.$$

Now we turn to the estimate for I_{12}. For $e \in \mathcal{E}_2$, we simply estimate

$$|\alpha_e - \alpha_e'| \le |\alpha_e| + |\alpha_e'| \lesssim h^2,$$
$$|\beta_e - \beta_e'| \le |\beta_e| + |\beta_e'| \lesssim h^2.$$

Using (2.15), this leads to

$$|I_{12}| \lesssim h^{1+\sigma}|u|_{2,\infty,\Omega}|v_h|_{1,\Omega}.$$

We now consider I_{13}. It is easy to see that, if $v_h = 0$ on $\partial\Omega$, then $I_{13} = 0$. In the general case, we set

$$B_e(u) = \alpha_e \frac{\partial^2 u}{\partial \boldsymbol{t}^2} + \beta_e \frac{\partial^2 u}{\partial \boldsymbol{t}\partial \boldsymbol{n}}$$

and

$$\overline{B}_e(u) = |e|^{-1} \int_e B_e(u).$$

Then

$$\begin{aligned} I_{13} &= \sum_{e\subset\partial\Omega} \int_e q_e B_e(u) \frac{\partial v_h}{\partial \boldsymbol{t}} \\ &= \sum_{e\subset\partial\Omega} \int_e q_e \overline{B}_e(u) \frac{\partial v_h}{\partial \boldsymbol{t}} + \sum_{e\subset\partial\Omega} \int_e q_e (B_e(u) - \overline{B}_e(u)) \frac{\partial v_h}{\partial \boldsymbol{t}}. \end{aligned}$$

For the second term, we have

$$\begin{aligned} \left|\sum_{e\subset\partial\Omega} \int_e q_e (B_e(u) - \overline{B}_e(u)) \frac{\partial v_h}{\partial \boldsymbol{t}}\right| &\lesssim h^3 |u|_{3,\infty,\Omega} \sum_{e\subset\partial\Omega} \int_e \left|\frac{\partial v_h}{\partial \boldsymbol{t}}\right| \\ &\lesssim h^{5/2}|u|_{3,\infty,\Omega}|v_h|_{1,\Omega}. \end{aligned}$$

We now estimate the first term. Let $\mathcal{P} = \mathcal{P}_1 \oplus \mathcal{P}_2$ denote the set of vertices on $\partial\Omega$. Then we have

$$\begin{aligned}\sum_{e\subset\partial\Omega}\int_e q_e\overline{B}_e(u)\frac{\partial v_h}{\partial \boldsymbol{t}} &= \sum_{e\subset\partial\Omega}\overline{B}_e(u)\frac{\partial v_h}{\partial \boldsymbol{t}}\int_e q_e\\ &= \sum_{e\subset\partial\Omega}\overline{B}_e(u)\frac{\partial v_h}{\partial \boldsymbol{t}}\frac{|e|}{6}\\ &= \frac{1}{6}\sum_{x\in\mathcal{P}}\left(\overline{B}_e(u)-\overline{B}_{e'}(u)\right)v_h(x)\end{aligned}$$

For $x \in \mathcal{P}_1$, we have

$$\begin{aligned}|\alpha_e - \alpha_{e'}| &\lesssim h^3,\\ |\beta_e - \beta_{e'}| &\lesssim h^3.\end{aligned}$$

Thus

$$\left|\overline{B}_e(u)-\overline{B}_{e'}(u)\right| \lesssim h^3|u|_{3,\infty,\Omega}.$$

For $x \in \mathcal{P}_2$, we have

$$\left|\overline{B}_e(u)-\overline{B}_{e'}(u)\right| \le \left|\overline{B}_e(u)\right| + \left|\overline{B}_{e'}(u)\right| \lesssim h^2|u|_{2,\infty,\Omega}.$$

Combining these estimates, we have

$$\begin{aligned}\left|\sum_{x\in\mathcal{P}}\left(\overline{B}_e(u)-\overline{B}_{e'}(u)\right)v_h(x)\right| &\lesssim h^2\left(|u|_{3,\infty,\Omega}+\kappa|u|_{2,\infty,\Omega}\right)\|v_h\|_{\infty,\partial\Omega}\\ &\lesssim h^2|\log h|^{1/2}\|u\|_{3,\infty,\Omega}\|v_h\|_{1,\Omega}.\end{aligned}$$

In the last step, we used the well known Sobolev inequality

$$\|v_h\|_{\infty,\Omega} \lesssim |\log h|^{1/2}\|v_h\|_{1,\Omega}.$$

$\|v_h\|_{1,\Omega}$ can be replaced by $|v_h|_{1,\Omega}$ by a standard argument. Thus our final estimate is

$$|I_{13}| \lesssim h^2|\log h|^{1/2}\|u\|_{3,\infty,\Omega}|v_h|_{1,\Omega}.$$

Consequently

$$|I_1| \lesssim h^{1+\min(1,\sigma)}|\log h|^{1/2}\|u\|_{3,\infty,\Omega}|v_h|_{1,\Omega}. \tag{2.17}$$

Combining (2.12) with (2.13) and (2.17), we obtain (2.11).
□

For pure Dirichlet boundary conditions, we have the following better estimate.

COROLLARY 2.6. *Assume the conditions of Lemma 2.5, except the second part of Definition 2.4 concerning regularity on the elements near the boundary. Then*

$$\begin{aligned}&\left|\sum_{\tau\in\mathcal{T}_h}\int_\tau \nabla(u-u_I)\cdot\mathcal{D}_\tau\nabla v_h\right|\\ &\qquad\lesssim h^{1+\min(1,\sigma)}(\|u\|_{3,\Omega}+\|u\|_{2,\infty,\Omega})|v_h|_{1,\Omega},\ v_h\in\mathcal{V}_h\cap H_0^1(\Omega).\end{aligned}$$

Proof. Use $I_{13} = 0$ in Lemma 2.5.□

In the general case, without the second part of Definition 2.4, we have the slightly weaker result.

COROLLARY 2.7. *Assume the conditions of Lemma 2.5, except the second part of Definition 2.4 concerning regularity on the elements near the boundary. Then*

$$\left| \sum_{\tau \in \mathcal{T}_h} \int_\tau \nabla(u - u_I) \cdot \mathcal{D}_\tau \nabla v_h \right| \lesssim h^{1+\min(1/2,\sigma)} (\|u\|_{3,\Omega} + \|u\|_{2,\infty,\Omega}) |v_h|_{1,\Omega}, \ v_h \in \mathcal{V}_h.$$

Proof. We always have the following estimate for I_{13}.

$$|I_{13}| \lesssim h^{3/2} |u|_{2,\infty,\partial\Omega} |v_h|_{1,\Omega}.$$

□

We conclude with a final technical result needed in Section 4.

LEMMA 2.8. *Let the triangulation $\mathcal{T}_h$ be $O(h^{2\sigma})$ irregular. Then*

$$\left| \sum_\tau \int_{\partial\tau} \sum_{k=1}^{3} \ell_k^2 \frac{\partial^2 u}{\partial \boldsymbol{t}_k^2} \boldsymbol{v}_h \cdot \boldsymbol{n} \right| \lesssim h^{1+\min(1,\sigma)} |\log h|^{1/2} \|u\|_{3,\infty,\Omega} \|\boldsymbol{v}_h\|_{0,\Omega}. \tag{2.18}$$

Proof. Let $e \equiv e_k$ be an arbitrary edge of element τ. We begin with the identity

$$\ell_k^2 \frac{\partial^2 u}{\partial \boldsymbol{t}_k^2} + \ell_{k+1}^2 \frac{\partial^2 u}{\partial \boldsymbol{t}_{k+1}^2} + \ell_{k-1}^2 \frac{\partial^2 u}{\partial \boldsymbol{t}_{k-1}^2} = (\alpha_e - \delta_e) \frac{\partial^2 u}{\partial \boldsymbol{t}_k^2} + \beta_e \frac{\partial^2 u}{\partial \boldsymbol{t}_k \partial \boldsymbol{n}_k} + \delta_e \frac{\partial^2 u}{\partial \boldsymbol{n}_k^2},$$

where

$$\begin{aligned} \alpha_e &= \ell_k^2 + \ell_{k+1}^2 + \ell_{k-1}^2, \\ \beta_e &= (\ell_{k+1}^2 - \ell_{k-1}^2) 4|\tau|/\ell_k^2, \\ \delta_e &= 8|\tau|^2/\ell_k^2. \end{aligned}$$

For $e \in \mathcal{E}$, let τ and τ' share e as a common edge. Take $\boldsymbol{n}$ and $\boldsymbol{t}$ to correspond to τ. Then we can write

$$\sum_\tau \int_{\partial\tau} \sum_{k=1}^{3} \ell_k^2 \frac{\partial^2 u}{\partial \boldsymbol{t}_k^2} \boldsymbol{v}_h \cdot \boldsymbol{n} = I_1 + I_2 + I_3,$$

where

$$I_j = \sum_{e \in \mathcal{E}_j} \int_e \left\{ (\alpha_e - \alpha_e') \frac{\partial^2 u}{\partial \boldsymbol{t}^2} + (\beta_e - \beta_e') \frac{\partial^2 u}{\partial \boldsymbol{t} \partial \boldsymbol{n}} \right\} \boldsymbol{v}_h \cdot \boldsymbol{n}$$

for $j = 1, 2$ and

$$I_3 = \sum_{e \in \partial\Omega} \int_e \left\{ (\alpha_e - \delta_e) \frac{\partial^2 u}{\partial \boldsymbol{t}^2} + \beta_e \frac{\partial^2 u}{\partial \boldsymbol{t} \partial \boldsymbol{n}} + \delta_e \frac{\partial^2 u}{\partial \boldsymbol{n}^2} \right\} \boldsymbol{v}_h \cdot \boldsymbol{n}.$$

Following the pattern of proof in Lemma 2.5, we estimate

$$\begin{aligned} |I_1| &\lesssim h^2 \|u\|_{3,\Omega} \|\boldsymbol{v}_h\|_{0,\Omega}, \\ |I_2| &\lesssim h^{1+\sigma} |u|_{2,\infty,\Omega} \|\boldsymbol{v}_h\|_{0,\Omega}, \\ |I_3| &\lesssim h^{1+\min(1,\sigma)} |\log h|^{1/2} \|u\|_{3,\infty,\Omega} \|\boldsymbol{v}_h\|_{0,\Omega}. \end{aligned}$$

(2.18) now follows directly from these estimates. □

3. Elliptic Boundary Value Problems. We consider the nonself-adjoint and possibly indefinite problem: find $u \in H^1(\Omega)$ such that

$$B(u,v) = \int_\Omega (\mathcal{D}\nabla u + \boldsymbol{b}u)\cdot\nabla v + cuv\,dx = f(v) \tag{3.1}$$

for all $v \in H^1(\Omega)$. Here $\mathcal{D}$ is a 2×2 symmetric, positive definite matrix, $\boldsymbol{b}$ a vector, and c a scalar, and $f(\cdot)$ is a linear functional. We assume all the coefficient functions are smooth.

In order to insure that (3.1) has a unique solution, we assume the bilinear form $B(\cdot,\cdot)$ satisfies the continuity condition

$$|B(\phi,\eta)| \le \nu\,\|\phi\|_{1,\Omega}\|\eta\|_{1,\Omega} \tag{3.2}$$

for all $\phi,\eta \in H^1(\Omega)$. We also assume the inf-sup conditions

$$\inf_{\phi\in H^1}\sup_{\eta\in H^1}\frac{B(\phi,\eta)}{\|\phi\|_{1,\Omega}\|\eta\|_{1,\Omega}} = \sup_{\phi\in H^1}\inf_{\eta\in H^1}\frac{B(\phi,\eta)}{\|\phi\|_{1,\Omega}\|\eta\|_{1,\Omega}} \ge \mu > 0, \tag{3.3}$$

Let $\mathcal{V}_h \subset H^1(\Omega)$ be the space of continuous piecewise linear polynomials associated with the triangulation $\mathcal{T}_h$, and consider the approximate problem: find $u_h \in \mathcal{V}_h$ such that

$$B(u_h,v_h) = f(v_h) \tag{3.4}$$

for all $v_h \in \mathcal{V}_h$. To insure a unique solution for (3.4) we assume the the inf-sup conditions

$$\inf_{\phi\in\mathcal{V}_h}\sup_{\eta\in\mathcal{V}_h}\frac{B(\phi,\eta)}{\|\phi\|_{1,\Omega}\|\eta\|_{1,\Omega}} = \sup_{\phi\in\mathcal{V}_h}\inf_{\eta\in\mathcal{V}_h}\frac{B(\phi,\eta)}{\|\phi\|_{1,\Omega}\|\eta\|_{1,\Omega}} \ge \mu > 0, \tag{3.5}$$

Xu and Zikatanov [21] have shown that under these assumptions,

$$\|u-u_h\|_{1,\Omega} \le \frac{\nu}{\mu}\inf_{v_h\in\mathcal{V}_h}\|u-v_h\|_{1,\Omega}.$$

See also Babuška and Aziz [1].

We define the piecewise constant matrix function $\mathcal{D}_\tau$ in terms of the diffusion matrix $\mathcal{D}$ as follows:

$$\mathcal{D}_{\tau ij} = \frac{1}{|\tau|}\int_\tau \mathcal{D}_{ij}\,dx.$$

Note that $\mathcal{D}_\tau$ is symmetric and positive definite.

THEOREM 3.1. *Assume that the solution of (3.1) satisfies $u \in W^{3,\infty}(\Omega)$. Further, assume the hypotheses of Lemma 2.5, with $\mathcal{D}_\tau$ defined as above. Then*

$$\|u_h - u_I\|_{1,\Omega} \lesssim h^{1+\min(1,\sigma)}|\log h|^{1/2}\|u\|_{3,\infty,\Omega}.$$

Proof. We begin with the identity

$$B(u-u_I,v_h) = \sum_{\tau\in\mathcal{T}_h}\int_\tau \nabla(u-u_I)\cdot\mathcal{D}_\tau\nabla v_h\,dx + \sum_{\tau\in\mathcal{T}_h}\int_\tau \nabla(u-u_I)\cdot(\mathcal{D}-\mathcal{D}_\tau)\nabla v_h\,dx$$
$$+\int_\Omega (u-u_I)(\boldsymbol{b}\cdot\nabla v_h + cv_h)\,dx = I_1 + I_2 + I_3.$$

The first term I_1 is estimated using Lemma 2.5. I_2 and I_3 can be easily estimated by

$$|I_2| + |I_3| \lesssim h^2 \|u\|_{2,\Omega} \|v_h\|_{1,\Omega}.$$

Thus

$$|B(u - u_I, v_h)| \lesssim h^{1+\min(1,\sigma)} |\log h|^{1/2} \|u\|_{3,\infty,\Omega} \|v_h\|_{1,\Omega}.$$

We complete the proof using the inf-sup condition in

$$\begin{aligned}
\mu \|u_h - u_I\|_{1,\Omega} &\leq \sup_{v_h \in \mathcal{V}_h} \frac{B(u_h - u_I, v_h)}{\|v_h\|_{1,\Omega}} \\
&= \sup_{v_h \in \mathcal{V}_h} \frac{B(u - u_I, v_h)}{\|v_h\|_{1,\Omega}} \\
&\lesssim h^{1+\min(1,\sigma)} |\log h|^{1/2} \|u\|_{3,\infty,\Omega}.
\end{aligned}$$

□

We now consider a more general nonlinear problem: find $u \in H^1(\Omega)$ such that

$$\mathcal{B}(u, v) = f(v) \tag{3.6}$$

for all $v \in H^1(\Omega)$. Here The form $\mathcal{B}(\cdot, \cdot)$ is assumed to be linear in its second argument, but nonlinear in its first. Once again, $f(v)$ is a linear functional. Let u_h be the finite element approximation: find $u_h \in \mathcal{V}_h$ such that

$$\mathcal{B}(u_h, v_h) = f(v_h) \tag{3.7}$$

for all $v_h \in \mathcal{V}_h$. We assume that $\mathcal{B}(\cdot, \cdot)$ is such that its linearization about u is a bilinear form $B(\cdot, \cdot)$ as in (3.1), although the coefficient functions will now generally depend on u. We assume that $B(\cdot, \cdot)$ satisfies the continuity and inf-sup conditions (3.2), (3.3) and (3.5), so that both (3.6) and (3.7) have unique solutions. The linearization process also satisfies

$$\mathcal{B}(u, v_h) - \mathcal{B}(u_h, v_h) = B(u - u_h, v_h) + \mathcal{Q}(u - u_h, v_h) = 0$$

for all $v_h \in \mathcal{V}_h$. The form $\mathcal{Q}(\cdot, \cdot)$ contains higher order truncation terms in the linearization process; as with $\mathcal{B}(\cdot, \cdot)$, it is linear in its second argument. We assume

$$|\mathcal{Q}(u - u_h, v_h)| \lesssim \|u - u_h\|_{1,\Omega}^2 \|v_h\|_{1,\Omega}. \tag{3.8}$$

THEOREM 3.2. *Assume the hypotheses of Theorem 3.1 and (3.8). Then*

$$\|u_h - u_I\|_{1,\Omega} \lesssim h^{1+\min(1,\sigma)} |\log h|^{1/2} \|u\|_{3,\infty,\Omega} + \|u - u_h\|_{1,\Omega}^2.$$

Proof. As in the proof of Theorem 3.1

$$\begin{aligned}
\mu \|u_h - u_I\|_{1,\Omega} &\leq \sup_{v_h \in \mathcal{V}_h} \frac{B(u_h - u_I, v_h)}{\|v_h\|_{1,\Omega}} \\
&\leq \sup_{v_h \in \mathcal{V}_h} \frac{B(u - u_I, v_h)}{\|v_h\|_{1,\Omega}} + \frac{\mathcal{Q}(u - u_h, v_h)}{\|v_h\|_{1,\Omega}} \\
&\lesssim h^{1+\min(1,\sigma)} |\log h|^{1/2} \|u\|_{3,\infty,\Omega} + \|u - u_h\|_{1,\Omega}^2.
\end{aligned}$$

□

If $\|u - u_h\|_{1,\Omega}$ is sufficiently small (e.g., $\|u - u_h\|_{1,\Omega} \leq C(u)h$), then we will observe superconvergence.

4. A Gradient Recovery Algorithm for $O(h^2)$ Approximate Parallelogram Meshes. In this section, we show that $Q_h\nabla u_I$ can superconverge to ∇u for meshes that are $O(h^{2\sigma})$ irregular.

THEOREM 4.1. *Let $u \in W^{3,\infty}(\Omega)$, and assume the hypotheses of Lemma 2.8. Then*

$$\|\nabla u - Q_h\nabla u_I\|_{0,\Omega} \lesssim h^{1+\min(1,\sigma)}|\log h|^{1/2}\|u\|_{3,\infty,\Omega}.$$

Proof. Given $\boldsymbol{v}_h \in \mathcal{V}_h \times \mathcal{V}_h$, we have

$$(Q_h\nabla(u-u_I),\boldsymbol{v}_h) = (\nabla(u-u_I),\boldsymbol{v}_h) = -((u-u_I),\nabla\cdot\boldsymbol{v}_h) + \int_{\partial\Omega}(u-u_I)\boldsymbol{v}_h\cdot\boldsymbol{n} \quad (4.1)$$

We estimate the two terms on the right hand side of (4.1). First,

$$\left|\int_{\partial\Omega}(u-u_I)\boldsymbol{v}_h\cdot\boldsymbol{n}\right| \lesssim h^{3/2}|u|_{2,\infty,\Omega}\|\boldsymbol{v}_h\|_{0,\Omega}.$$

For the other, we use Lemma 2.2 to get

$$\begin{aligned}
\int_\tau (u-u_I)\nabla\cdot\boldsymbol{v}_h &= -\frac{1}{24}\int_\tau\sum_{k=1}^3 \ell_k^2\frac{\partial^2 u_q}{\partial \boldsymbol{t}_k^2}\nabla\cdot\boldsymbol{v}_h + \int_\tau (u-u_q)\nabla\cdot\boldsymbol{v}_h\\
&= -\frac{1}{24}\int_\tau\sum_{k=1}^3 \ell_k^2\frac{\partial^2 u}{\partial \boldsymbol{t}_k^2}\nabla\cdot\boldsymbol{v}_h\\
&\qquad -\frac{1}{24}\int_\tau\sum_{k=1}^3 \ell_k^2\frac{\partial^2 (u_q-u)}{\partial \boldsymbol{t}_k^2}\nabla\cdot\boldsymbol{v}_h + \int_\tau (u-u_q)\nabla\cdot\boldsymbol{v}_h\\
&= -\frac{1}{24}\int_{\partial\tau}\sum_{k=1}^3 \ell_k^2\frac{\partial^2 u}{\partial \boldsymbol{t}_k^2}\boldsymbol{v}_h\cdot\boldsymbol{n} + \frac{1}{24}\int_\tau\sum_{k=1}^3 \ell_k^2\nabla\frac{\partial^2 u}{\partial \boldsymbol{t}_k^2}\boldsymbol{v}_h\\
&\qquad -\frac{1}{24}\int_\tau\sum_{k=1}^3 \ell_k^2\frac{\partial^2 (u_q-u)}{\partial \boldsymbol{t}_k^2}\nabla\cdot\boldsymbol{v}_h + \int_\tau (u-u_q)\nabla\cdot\boldsymbol{v}_h\\
&= I_1 + I_2 + I_3 + I_4
\end{aligned}$$

Easy estimates show

$$|I_3| + |I_4| \lesssim h^3\|u\|_{3,\tau}|\boldsymbol{v}_h|_{1,\tau} \lesssim h^2\|u\|_{3,\tau}\|\boldsymbol{v}_h\|_{0,\tau},$$
$$|I_2| \lesssim h^2\|u\|_{3,\tau}\|\boldsymbol{v}_h\|_{0,\tau}.$$

$|I_1|$ is estimated using Lemma 2.8. Consequently

$$|(Q_h\nabla(u-u_I),\boldsymbol{v}_h)| \lesssim h^{1+\min(1,\sigma)}|\log h|^{1/2}\|u\|_{3,\infty,\Omega}\|\boldsymbol{v}_h\|_{0,\Omega}.$$

Taking $\boldsymbol{v}_h = Q_h\nabla(u-u_I)$, it follows that

$$\|Q_h\nabla(u-u_I)\|_{0,\Omega} \lesssim h^{1+\min(1,\sigma)}|\log h|^{1/2}\|u\|_{3,\infty,\Omega}.$$

Theorem 4.1 now follows from the triangle inequality

$$\|\nabla u - Q_h\nabla u_I\|_{0,\Omega} \le \|\nabla u - Q_h\nabla u\|_{0,\Omega} + \|Q_h\nabla(u-u_I)\|_{0,\Omega}.$$

□

An immediate consequence of Theorems 3.1 and 4.1 is

THEOREM 4.2. *Let $u \in W^{3,\infty}(\Omega)$, and assume the hypotheses of Theorems 3.1 and 4.1. Then*

$$\|\nabla u - Q_h \nabla u_h\|_{0,\Omega} \lesssim h^{1+\min(1,\sigma)} |\log h|^{1/2} \|u\|_{3,\infty,\Omega}.$$

Proof. Using the triangle inequality

$$\begin{aligned} \|\nabla u - Q_h \nabla u_h\|_{0,\Omega} &\leq \|\nabla u - Q_h \nabla u_I\|_{0,\Omega} + \|Q_h \nabla (u_I - u_h)\|_{0,\Omega} \\ &\leq \|\nabla u - Q_h \nabla u_I\|_{0,\Omega} + \|\nabla (u_I - u_h)\|_{0,\Omega}. \end{aligned} \tag{4.2}$$

We estimate the two terms on the right hand side of (4.2) using Theorems 4.1 and 3.1. □

Finally, we would like to point out that many results presented above (such as Theorems 3.1, 3.2, 4.1 and 4.2) can be refined in many ways. Before the end of this section, let us give one such a refinement for piecewise $O(h^{2\sigma})$ irregular grid.

DEFINITION 4.3. *The triangulation $\mathcal{T}_h$ is piecewise $O(h^{2\sigma})$ irregular if Ω can be written as a union of a bounded number of polygonal subdomains and $\mathcal{T}_h$ is $O(h^{2\sigma})$ irregular on each of these subdomains.*

By applying Lemma 2.5 on each subdomain, we can easily get the following result.

THEOREM 4.4. *Lemma 2.5, Lemma 2.8, Theorem 3.1, Theorem 3.2, Theorem 4.1 and Theorem 4.2 are all valid for piecewise $O(h^{2\sigma})$ grids.*

The above theorem is related to superconvergence results on piecewise regular (or strongly regular) grid that were discussed in earlier literature, c.f. Xu [20] and Lin and Xu [15]. The significance of such an extension will be discussed in the following section.

5. Applications and Numerical Experiments. In this section, we develop a few simple applications of our results, and present some numerical examples. The numerical experiments were performed using the PLTMG package [5]. The experiments were done on an Linux PC using double precision arithmetic and the g77 compiler.

We begin out discussion with a very simple example of piecewise uniform grids. As shown in Figure 5.1, we began with a uniform 3×3 mesh with $nt = 8$ elements, and computed a sequence of uniformly refined meshes through regular refinement of each element of a given mesh into four similar triangles in the refined mesh by pairwise connecting the midpoints.

This grid is $O(h)$ irregular ($\sigma = 1/2$) by Definition 2.4 but piecewise $O(h^{2\sigma})$ irregular with $\sigma = \infty$ (namely piecewise regular) by Definition 4.3. Consequently, for this example, the result claimed by Theorem 4.4 is $O(h^{1/2})$ better than the corresponding result from previous sections. In our first experiment, we consider the problem

$$-\Delta u + u = f \tag{5.1}$$

$\Omega = (0,1) \times (0,1)$ with either Dirichlet or Neumann boundary conditions. The right hand side f and the boundary conditions were chosen such that $u = e^{x+y}$ was the exact solution. In this experiment, we begin with the uniform 3×3 mesh with eight triangles described above, and make seven levels of uniform regiment. The results are

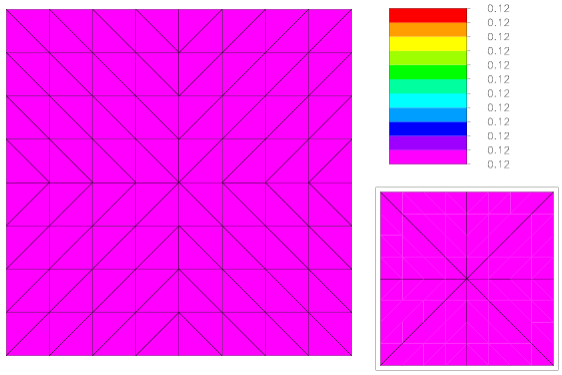

FIG. 5.1. *A (globally) $O(h^{2\sigma})$ irregular grid with $\sigma = 1/2$, but piecewise $O(h^{2\sigma})$ irregular grid with $\sigma = \infty$*

reported in Table 5.1. In Table 5.1 and subsequent tables,

$$H_1 = \|\nabla(u - u_h)\|_{0,\Omega},$$
$$\widetilde{H}_1 = \|\nabla(u_I - u_h)\|_{0,\Omega},$$
$$\overline{H}_1 = \|\nabla u - Q_h \nabla u_h)\|_{0,\Omega},$$

where u_I is the linear interpolant of u. In the last line, the order of convergence was estimated from the reported data using a least squares technique.

TABLE 5.1
Results for square domain, uniform refinement.

	Dirichlet Problem			Neumann Problem		
nt	H_1	$\widetilde{H}_1$	$\overline{H}_1$	H_1	$\widetilde{H}_1$	$\overline{H}_1$
8	1.2e 0	2.7e-1	6.0e-1	9.5e-1	7.2e-1	6.7e-1
32	6.0e-1	8.1e-2	2.4e-1	5.5e-1	2.4e-1	3.0e-1
128	3.0e-1	2.3e-2	8.8e-2	2.9e-1	7.5e-2	1.1e-1
512	1.5e-1	6.1e-3	3.2e-2	1.5e-1	2.2e-2	3.7e-2
2048	7.5e-2	1.6e-3	1.1e-2	7.5e-2	6.1e-3	1.3e-2
8192	3.8e-2	4.4e-4	4.0e-3	3.8e-2	1.7e-3	4.3e-3
32768	1.9e-2	1.2e-4	1.4e-3	1.9e-2	4.5e-4	1.5e-3
131072	9.4e-3	3.0e-5	5.1e-4	9.4e-3	1.2e-4	5.2e-4
order	1.01	1.95	1.51	1.01	1.91	1.55

In Table 5.1, we see quite clearly the first order convergence of $\|\nabla(u - u_h)\|_{0,\Omega}$, and the superconvergence of $\|\nabla(u_I - u_h)\|_{0,\Omega}$. In the latter case, the rate is nearly second order, which is consistent with Theorem 4.4. We also note superconvergence of $\|\nabla u - Q_h \nabla u_h)\|_{0,\Omega}$, with order close to 3/2. This perhaps is the result of most practical significance.

We then repeated the experiment, replacing uniform refinement with the adaptive refinement procedure in PLTMG. This adaptive refinement procedure is based on longest-edge bisection, and also includes a mesh smoothing phase that allows the vertices in the mesh to move. The result was a sequence of unstructured, nonuniform, nonnested, shape regular meshes. The target values for the the adaptive procedure

were selected to produce a sequence of meshes with approximately the same numbers of elements as the uniform refinement case. The results are shown in Table 5.2.

For the adaptive meshes, the story is a quite different; $\|\nabla(u_I - u_h)\|_{0,\Omega}$ and $\|\nabla u - Q_h \nabla u_h)\|_{0,\Omega}$ show less superconvergence In this case $\sigma > 0$, but it is clearly much smaller than in the uniform refinement case. In Part II of this work [7], we show how to obtain strong superconvergence for such meshes using $S^m Q_h \nabla u_h$ as the recovered gradient. Here S is a multigrid-like smoothing operator, and m is a small integer ($m = 1$ or $m = 2$ is usually satisfactory). Analysis and a complete description are deferred to Part II of this work.

TABLE 5.2
Results for square domain, adaptive refinement.

	Dirichlet Problem				Neumann Problem		
nt	H_1	$\widetilde{H}_1$	$\overline{H}_1$	nt	H_1	$\widetilde{H}_1$	$\overline{H}_1$
8	1.2e 0	2.7e-1	6.0e-1	8	9.5e-1	7.2e-1	6.7e-1
34	5.8e-1	1.0e-1	2.3e-1	36	4.6e-1	2.7e-1	2.3e-1
136	2.2e-1	7.2e-2	7.7e-2	134	2.5e-1	9.3e-2	1.0e-1
528	1.2e-1	3.4e-2	3.8e-2	526	1.2e-1	4.0e-2	3.8e-2
2079	6.0e-2	1.7e-2	1.7e-2	2080	6.0e-2	1.8e-2	1.6e-2
8254	2.9e-2	6.9e-3	7.1e-3	8257	2.9e-2	7.9e-3	7.2e-3
32888	1.4e-2	3.2e-3	3.0e-3	32890	1.4e-2	3.7e-3	3.2e-3
131301	6.9e-3	1.5e-3	1.4e-3	131311	7.0e-3	1.8e-3	1.5e-3
order	1.05	1.10	1.17		1.04	1.11	1.13

In our second experiment, we solved (5.1) on a domain Ω in the shape of Lake Superior. The true solution u in this case was chosen to be $u = \sin x \sin y$. In this case, the initial mesh with $nt = 2765$ elements was unstructured and nonuniform, but shape regular. This mesh is shown in Figure 5.2. As in the first example, we first computed a sequence of uniformly refined meshes through regular refinement of each element of a given mesh into four similar triangles. The results are shown in Table 5.3.

TABLE 5.3
Lake Superior domain, uniform refinement.

	Dirichlet Problem			Neumann Problem		
nt	H_1	$\widetilde{H}_1$	$\overline{H}_1$	H_1	$\widetilde{H}_1$	$\overline{H}_1$
2765	9.2e-1	1.5e-1	2.5e-1	9.1e-1	1.6e-1	2.6e-1
11060	4.6e-1	4.5e-2	1.0e-1	4.6e-1	4.8e-2	1.0e-1
44240	2.3e-1	1.3e-2	3.5e-2	2.3e-1	1.4e-2	3.5e-2
176960	1.2e-1	3.6e-3	1.2e-2	1.2e-1	3.8e-3	1.2e-2
order	1.02	1.88	1.54	1.02	1.89	1.54

In Table 5.3, we see quite clearly the first order convergence of $\|\nabla(u - u_h)\|_{0,\Omega}$, and the superconvergence of $\|\nabla(u_I - u_h)\|_{0,\Omega}$. In the latter case, the rate is nearly second order, which is again consistent with Theorem 4.4. Evidently, the many small uniform patches were sufficient to produce a very strong superconvergence effect. We also see superconvergence of $Q_h \nabla u_h$ to ∇u, similar to the first example.

We then repeated the experiment, replacing uniform with adaptive refinement. The target values for the the adaptive procedure once again were selected to produce a

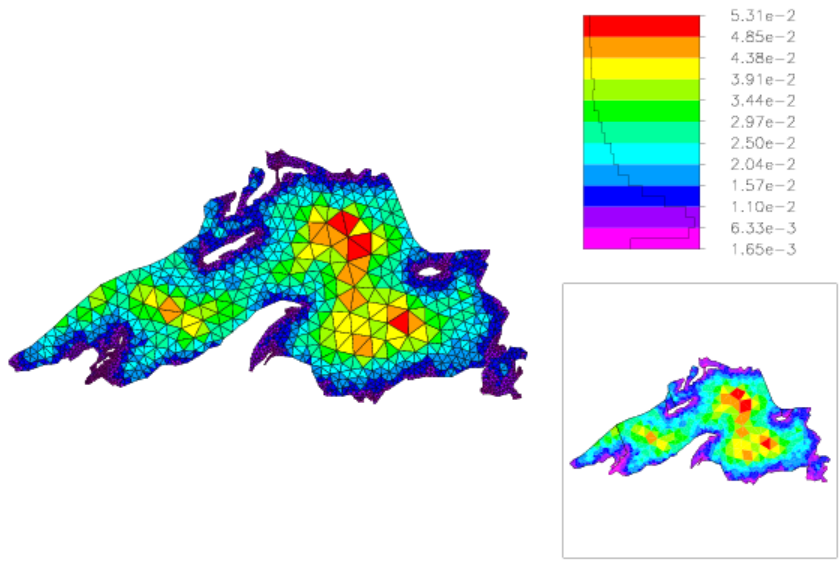

FIG. 5.2. *Lake Superior mesh with* $nt = 2765$. *Elements are colored according to size.*

TABLE 5.4
Lake Superior domain, adaptive refinement.

	Dirichlet Problem				Neumann Problem		
nt	H_1	$\widetilde{H}_1$	$\overline{H}_1$	nt	H_1	$\widetilde{H}_1$	$\overline{H}_1$
2765	9.2e-1	1.5e-1	2.5e-1	2765	9.1e-1	1.6e-1	2.6e-1
11565	2.5e-1	4.0e-2	5.3e-2	11560	2.5e-1	4.3e-2	5.3e-2
45524	1.2e-1	1.8e-2	2.2e-2	45521	1.2e-1	1.9e-2	2.2e-2
179655	6.1e-2	8.3e-3	9.9e-3	179666	6.1e-2	8.7e-3	9.9e-3
order	1.12	1.22	1.32		1.12	1.26	1.32

sequence of meshes with approximately the same numbers of elements as the uniform refinement case. The results are shown in Table 5.4.

The results here are qualitatively similar to the first example. We note in Table 5.4 slightly elevated estimates for the estimated order of convergence of $\|\nabla(u - u_h)\|_{0,\Omega}$. This is an artifact of the least squares procedure. Notice that in the first adaptive step, there was an unusually large decrease in $\|\nabla(u - u_h)\|_{0,\Omega}$. This was because the initial nonuniform mesh was adapted mainly to the complex geometry of Ω and not to the character of the solution. In subsequent adaptive refinement steps, the error is rapidly approaching first order behavior. The orders for $H1$ and $\overline{H1}$ are also slightly elevated by unusually large decreases in the first adaptive step.

Acknowledgment. The ideas for this manuscript germinated while the authors were attending the workshop on *A Posteriori Error Estimation and Adaptive Approaches in the Finite Element Method,* held at the Mathematical Sciences Research Institute, University of California, Berkeley, April 3-14, 2000. We are grateful for the opportunity to attend this workshop and for the stimulating scientific environment at MSRI. We also thank the reviewers for their careful reviews and their suggestions for improvements.

REFERENCES

[1] A. K. Aziz and I. Babuška, *Part I, survey lectures on the mathematical foundations of the finite element method*, in The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations, Academic Press, New York, 1972, pp. 1–362.

[2] I. Babuška and W. C. Rheinboldt, *A posteriori error error estimates for the finite element method*, Internat. J. Numer. Methods Engrg., 12 (1978), pp. 1597–1615.

[3] I. Babuška and T. Strouboulis, *The finite element method and its reliability*, Numerical Mathematics and Scientific Computation, Oxford Science Publications, 2001.

[4] I. Babuška, T. Strouboulis, and C. S. Upadhyay, *η%-superconvergence of finite element approximations in the interior of general meshes of triangles*, Comput. Methods Appl. Mech. Engrg., 122 (1995), pp. 273–305.

[5] R. E. Bank, *PLTMG: A Software Package for Solving Elliptic Partial Differential Equations, Users' Guide 8.0*, Software, Environments and Tools, Vol. 5, SIAM, Philadelphia, 1998.

[6] R. E. Bank and A. Weiser, *Some a posteriori error estimators for elliptic partial differential equations*, Mathematics of Computation, 44 (1985), pp. 283–301.

[7] R. E. Bank and J. Xu, *Asymptotically exact a posteriori error estimators, part II: General unstructured grids*, SIAM J. Numerical Analysis, (submitted).

[8] C. Chen and Y. Huang, *High accuracy theory of finite element methods*, Hunan Science Press, Hunan, China, 1995. in Chinese.

[9] L. Du and N. Yan, *Gradient recovery type a posteriori error estimate for finite element approximation on non-uniform meshes*, Adv. Comput. Math., 14 (2001), pp. 175–193.

[10] R. Durán, M. A. Muschietti, and R. Rodríguez, *On the asymptotic exactness of error estimators for linear triangular finite elements*, Numer. Math., 59 (1991), pp. 107–127.

[11] I. Hlaváček and M. Křížek, *On a superconvergent finite element scheme for elliptic systems. I. Dirichlet boundary condition*, Apl. Mat., 32 (1987), pp. 131–154.

[12] W. Hoffmann, A. H. Schatz, L. B. Wahlbin, and G. Wittum, *Asymptotically exact a posteriori estimators for the pointwise gradient error on each element in irregular meshes. I. A smooth problem and globally quasi-uniform meshes*, Math. Comp., 70 (2001), pp. 897–909 (electronic).

[13] A. M. Lakhany, I. Marek, and J. R. Whiteman, *Superconvergence results on mildly structured triangulations*, Comput. Methods Appl. Mech. Engrg., 189 (2000), pp. 1–75.

[14] B. Li and Z. Zhang, *Analysis of a class of superconvergence patch recovery techniques for linear and bilinear finite elements*, Numer. Methods Partial Differential Equations, 15 (1999), pp. 151–167.

[15] Q. Lin and J. Xu, *Linear finite elements with high accuracy*, J. Comp. Math., 3 (1985), pp. 115–133.

[16] Q. Lin and N. Yan, *The construction and analysis of high efficiency finite elements*, Hebei University Press, Hunan, China, 1996. in Chinese.

[17] R. Verfürth, *A Posteriori Error Estimation and Adaptive Mesh Refinement Techniques*, Teubner Skripten zur Numerik, B. G. Teubner, Stuttgart, 1995.

[18] L. B. Wahlbin, *Superconvergence in Galerkin finite element methods*, Springer-Verlag, Berlin, 1995.

[19] ———, *General principles of superconvergence in Galerkin finite element methods*, in Finite element methods (Jyväskylä, 1997), Dekker, New York, 1998, pp. 269–285.

[20] J. Xu, *The error analysis and the improved algorithms for the infinite element method*, in Proceedings of the 1984 Beijing Symposium on Differential Geometry and Differential Equations, Beijing, China, 1985, Science Press, pp. 326–331.

[21] J. Xu and L. Zikatanov, *Some observations on Babuška and Brezzi theories*, Numerische Mathematik, (to appear).

[22] N. Yan and A. Zhou, *Gradient recovery type a posteriori error estimates for finite element approximations on irregular meshes*, Comput. Methods Appl. Mech. Engrg., 190 (2001), pp. 4289–4299.

[23] Z. Zhang and H. D. Victory, Jr., *Mathematical analysis of Zienkiewicz-Zhu's derivative patch recovery technique*, Numer. Methods Partial Differential Equations, 12 (1996), pp. 507–524.

[24] Z. Zhang and J. Z. Zhu, *Superconvergence of the derivative patch recovery technique and a posteriori error estimation*, in Modeling, mesh generation, and adaptive numerical methods for partial differential equations (Minneapolis, MN, 1993), Springer, New York, 1995, pp. 431–450.

[25] J. Z. Zhu and O. C. Zienkiewicz, *Superconvergence recovery technique and a posteriori error*

estimators, Internat. J. Numer. Methods Engrg., 30 (1990), pp. 1321–1339.