ASYMPTOTICALLY EXACT A POSTERIORI ERROR ESTIMATORS, PART II: GENERAL UNSTRUCTURED GRIDS

RANDOLPH E. BANK* AND JINCHAO XU†

Abstract. In Part I of this work, we analyzed superconvergence for piecewise linear finite element approximations on triangular meshes where most pairs of triangles sharing a common edge form approximate parallelograms. In this work, we consider superconvergence for general unstructured but shape regular meshes. We develop a post-processing gradient recovery scheme for the finite element solution u_h, inspired in part by the smoothing iteration of the multigrid method. This recovered gradient superconverges to the gradient of the true solution, and becomes the basis of a global a posteriori error estimate that is often asymptotically exact. Next, we use the superconvergent gradient to approximate the Hessian matrix of the true solution, and form local error indicators for adaptive meshing algorithms. We provide several numerical examples illustrating the effectiveness of our procedures.

Key words. Superconvergence, gradient recovery, a posteriori error estimates.

AMS subject classifications. 65N50, 65N30

1. Introduction. In Part I of this work [10], we developed some superconvergence estimates and a gradient recovery algorithm appropriate for piecewise linear finite element approximations of elliptic boundary problems. In that work, we restricted attention to triangular meshes that are $O(h^{2\sigma})$ irregular [10]. In this work, we extend the gradient recovery scheme to more general meshes and develop an a posteriori error estimate and local error indicator for use in adaptive meshing algorithms. See [15, 23, 24, 25, 14, 22, 12, 16, 20, 5, 13, 1] for related work, and in particular the monographs by Verfürth [18], Babuška and Strouboulis [4], Chen and Huang [11], Lin and Yan [17], Wahlbin [19] for recent surveys of the field as a whole.

Our overall development has three major steps. Let $\mathcal{V}_h \subset H^1(\Omega)$ be the finite element subspace consisting of continuous piecewise linear polynomials associated with a shape regular triangulation $\mathcal{T}_h$. Let $u_h \in \mathcal{V}_h$ be the finite element solution of an appropriate linear or nonlinear elliptic boundary value problem. In the first component of our development, we prove a superconvergence result for $|u_h - u_I|_{1,\Omega}$, where u_I is the piecewise linear interpolant for u. In particular, in Part I of this manuscript [10], we prove 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 nonuniform meshes, where most pairs of adjacent triangles satisfy an $O(h^2)$ approximate parallelogram property. $\sigma > 0$ in some sense measures the extent to which this condition is violated; see [10] for details.

The second major component is a superconvergent approximation to ∇u. This approximation is generated by a gradient recovery procedure. In particular, in Section 2 of this manuscript we compute $S^m Q_h \nabla u_h$, where S is an appropriate smoothing operator, and Q_h is the L^2 projection operator. In words, the discontinuous, piecewise

*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.

constant gradient ∇u_h is projected into the space of continuous piecewise linear polynomials, and then smoothed, using a multigrid-like smoothing operator. Although the L^2 projection operator is global, the overall work estimate is still $O(N)$ for a mesh with N vertices. In the case of a small number of smoothing steps (the most interesting case), Theorem 2.7 shows that

$$\|\nabla u - S^m Q_h \nabla u_h\|_{0,\Omega} \lesssim h \left\{ \min\left(h^{\min(1,\sigma)} |\log h|, \left[\frac{\kappa-1}{\kappa}\right]^m \right) + m\, h^{1/2} \right\} \|u\|_{3,\infty,\Omega}. \tag{1.2}$$

Here $\kappa > 1$ is a constant independent of h and u. The term $(1-\kappa^{-1})^m$ illustrates the well-known effectiveness of a few smoothing steps, and is reminiscent of terms arising in connection with multigrid convergence analysis [7]. If σ is sufficiently large, then the L^2 projection itself ($m = 0$) is sufficient to produce superconvergence. The purpose of smoothing is to improve the performance when $\sigma \approx 0$ and the mesh is shape regular.

In the third major component of our analysis, presented in section 3 of this manuscript, 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 Theorem 3.1 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. This is motivated by noting that under certain circumstances, $|u_q - u_I|_{1,\Omega}$ is an asymptotically exact estimate of $|u - u_h|_{1,\Omega}$. Here u_q is the piecewise quadratic interpolant for u. Thus $u_q - u_I$ is a locally defined quadratic polynomial with value zero at all vertices of the mesh. On a given element τ, $u_q - u_I$ can be expressed as a linear combination of quadratic "bump functions" q_k associated with the edge midpoints of τ,

$$u_q - u_I = \sum_{k=1}^{3} \ell_k^2 \boldsymbol{t}_k^t M_\tau \boldsymbol{t}_k \, q_k(x,y) \tag{1.3}$$

where ℓ_k is the length of edge k, $\boldsymbol{t}_k$ is the unit tangent, and

$$M_\tau = -\frac{1}{2} \begin{pmatrix} \partial_{11} u_q & \partial_{12} u_q \\ \partial_{21} u_q & \partial_{22} u_q \end{pmatrix} \tag{1.4}$$

is the Hessian matrix (for details, see Section 3). For convenience in notation, we let $\partial_i u$ denote the partial derivative $\partial u / \partial x_i$. All terms on the right hand side of (1.3) are known except for the second derivatives appearing in the Hessian matrix M_τ. In our local error indicator, we simply replace $\partial_{ij} u_q$ by $\partial_i S^m Q_h \partial_j u_h$. Let ϵ_τ denote this locally defined a posteriori error estimate. In Theorem 3.2, we prove the superconvergence estimate

$$\|\partial_i(\partial_k u - S^m Q_h \partial_k u_h)\|_{0,\Omega} \lesssim \left\{ \min\left(h^{\min(1,\sigma)} |\log h|, \left[\frac{\kappa-1}{\kappa}\right]^m \right) + m\, h^{1/2} \right\} \|u\|_{3,\infty,\Omega}. \tag{1.5}$$

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. 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.

It also is interesting to note that the superconvergent global approximation to ∇u emphasizes once again a classic dilemma in error estimation. On the one hand, generally it seems quite advantageous to take the superconvergent approximation $S^m Q_h \nabla u_h$ as the "accepted" approximation to ∇u. Not only is it of higher order than ∇u_h, it is globally continuous and differentiable, often desirable properties. On the other hand, the a posteriori error estimates and resulting adaptive meshing algorithms use $S^m Q_h \nabla u_h$ to estimate the error in ∇u_h. In some respects, the situation is analogous to adaptive time step selection schemes for initial value problems where order p and $p+1$ approximations are computed to estimate the local error in the order p approximation, which is then used to control the time step.

Finally, as a point of practical interest, since the gradient recovery and a posteriori error estimates are independent of the PDE, a single implementation can be used across a broad spectrum of problems. There is no need to have special implementations for each problem class, as is typical of schemes that involve the solution of local problems in each element or patch of elements [3, 9].

The rest of this paper is organized as follows: In Section 2 we first provide some notation, describe our gradient recovery scheme, and summarize the main superconvergence estimates of [10] for the case of $O(h^{2\sigma})$ irregular meshes. We then extend the gradient recovery scheme to more general meshes through the use of a multigrid smoother. In Section 3, we develop and analyze our a posteriori error estimate, and prove (1.5). Finally, in Section 4, we present several numerical examples, involving both uniform and non-uniform (adaptive) meshes, with some solutions that satisfy our smoothness assumptions and some that do not. In the latter cases, we observe superconvergence away from singularities for adaptive meshes, although this effect is not covered by our current analysis.

2. A Gradient Recovery Algorithm for Shape Regular Triangulations. 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 polyhedron. We assume that Ω is partitioned by a shape regular triangulation $\mathcal{T}_h$ of mesh size h. Let $\mathcal{V}_h \subset H^1(\Omega)$ be the corresponding continuous piecewise linear finite element space associated with this triangulation $\mathcal{T}_h$.

We consider the non self-adjoint and possibly indefinite problem: find $u \in H^1(\Omega)$ such that

$$B(u,v) = \int_\Omega (\mathcal{D}\nabla u + \boldsymbol{b}u) \cdot \nabla v + cuv \, dx = f(v) \tag{2.1}$$

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. Choosing $H^1(\Omega)$ as trial space implies Neumann boundary conditions, a choice made for convenience. In [10], we also analyzed more general nonlinear PDEs and boundary conditions. However, since the details of the PDE do not strongly influence our gradient recovery scheme, here we consider only the most simple case.

In order to insure that (2.1) has a unique solution, we assume the bilinear form $B(\cdot,\cdot)$ satisfies the continuity condition

$$|B(\phi,\eta)| \leq \nu \, \|\phi\|_{1,\Omega} \|\eta\|_{1,\Omega} \tag{2.2}$$

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

$$\inf_{\phi\in H^1}\sup_{\eta\in H^1}\frac{B(\phi,\eta)}{\|\phi\|_{1,\Omega}\|\eta\|_{1,\Omega}} = \sup_{\phi\in H^1}\inf_{\eta\in H^1}\frac{B(\phi,\eta)}{\|\phi\|_{1,\Omega}\|\eta\|_{1,\Omega}} \geq \mu > 0, \tag{2.3}$$

For simplicity, we assume that μ and ν are such that the standard Galerkin finite element approximation is an appropriate discretization. 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

$$B(u_h, v_h) = f(v_h) \tag{2.4}$$

for all $v_h \in \mathcal{V}_h$. To insure a unique solution for (2.4) we assume the inf-sup conditions

$$\inf_{\phi\in\mathcal{V}_h}\sup_{\eta\in\mathcal{V}_h}\frac{B(\phi,\eta)}{\|\phi\|_{1,\Omega}\|\eta\|_{1,\Omega}} = \sup_{\phi\in\mathcal{V}_h}\inf_{\eta\in\mathcal{V}_h}\frac{B(\phi,\eta)}{\|\phi\|_{1,\Omega}\|\eta\|_{1,\Omega}} \geq \mu > 0, \tag{2.5}$$

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

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

See also Babuška and Aziz [2]. In this situation, we have standard a priori estimates of the form

$$\|u - u_h\|_{\alpha,\Omega} \lesssim h^{2-\alpha}\|u\|_{2,\Omega}$$

for $0 \leq \alpha \leq 1$.

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. The following results are proved in [10].

THEOREM 2.1. *Let the triangulation $\mathcal{T}_h$ be $O(h^{2\sigma})$ irregular [10]. Assume $\mathcal{D}_\tau$ defined above satisfies*

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

for $i = 1, 2$, $j = 1, 2$. Here τ and τ' are a pair of triangles sharing a common edge. Assume that the solution of (2.1) satisfies $u \in W^{3,\infty}(\Omega)$ and the $u_h \in \mathcal{V}_h$ if the solution of (2.4). Then

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

When the mesh is not $O(h^{2\sigma})$ irregular or σ becomes very close to zero, then the superconvergence demonstrated in Theorem 2.1 will be diminished. Intuitively,

it appears that superconvergence of $Q_h \nabla u_h$ is diminished mainly because of 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. In particular, we construct a multigrid smoother S and take $S^m Q_h \nabla u_h$ as our recovered gradient. As with multigrid methods, we expect a very small number of smoothing steps will suffice; in our code, we take $m = 2$ as default.

Our post-processing gradient recovery scheme is based on the following bilinear form:

$$a(u, v) = (\nabla u, \nabla v) + (u, v). \tag{2.6}$$

We introduce the discrete operator $A : v_h \mapsto \mathcal{V}_h$ defined by

$$(Au_h, v_h) = a(u_h, v_h), \quad \forall u_h, v_h \in \mathcal{V}_h.$$

We note that A is symmetric positive definite on $\mathcal{V}_h$ and

$$\lambda \equiv \rho(A) \eqsim h^{-2}. \tag{2.7}$$

Using A, we introduce the smoothing operator S defined by

$$S = I - \lambda^{-1} A.$$

The usual multigrid convergence function

$$f(\alpha, \beta) = \frac{\alpha^\alpha \beta^\beta}{(\alpha + \beta)^{(\alpha+\beta)}}$$

$\alpha, \beta > 0$, plays an important role. Here we summarize some standard properties of $f(\alpha, \beta)$. Let $p, \alpha, \beta > 0$. Then

$$\begin{aligned} \sup_{x \in [0,1]} x^\alpha (1 - x)^\beta &= f(\alpha, \beta), \\ f(\alpha, \beta)^p &= f(p\alpha, p\beta), \\ f(\alpha, \beta) &= f(\beta, \alpha). \end{aligned}$$

For convenience in notation, we let $\partial_i u$ denote the partial derivative $\partial u / \partial x_i$. We now state and prove some preliminary lemmas leading up the main Theorem 2.7 in this section.

LEMMA 2.2. *For any $z \in \mathcal{V}_h$,*

$$\|(I - S^m) z\|_{0,\Omega} \lesssim mh \left(\|z - \partial_i u\|_{1,\Omega} + h \|u\|_{3,\Omega} + h^{1/2} |u|_{2,\infty,\partial\Omega} \right)$$

Proof. We note from the definition of S,

$$\begin{aligned} \|(I - S^m) z\| &= \lambda^{-1} \|(I - S^m)(I - S)^{-1} A z\| \\ &\leq \lambda^{-1} \max_{s \in [0,1]} [(1 - s^m)(1 - s)^{-1}] \|Az\| \\ &\leq \lambda^{-1} m \|Az\| \\ &\lesssim mh^2 \|Az\| \end{aligned}$$

Let $w = Az$. By definition

$$(w, \phi) = (\nabla z, \nabla \phi) + (z, \phi) \tag{2.8}$$

for all $\phi \in \mathcal{V}_h$. We take $\phi = w$ in(2.8) and estimate the terms on the right hand side. The critical term is $(\nabla z, \nabla w)$, where we have

$$\begin{aligned}
(\nabla z, \nabla w) &= (\nabla(z - \partial_i u), \nabla w) + (\nabla \partial_i u, \nabla w) \\
&\lesssim \|\nabla(z - \partial_i u)\| \|\nabla w\| - (\Delta \partial_i u, w) + \int_{\partial\Omega} \nabla \partial_i u \cdot \boldsymbol{n}\, w \, ds \\
&\lesssim (h^{-1}\|z - \partial_i u\|_{1,\Omega} + \|u\|_{3,\Omega})\|w\|_{0,\Omega} + |u|_{2,\infty,\partial\Omega} \int_{\partial\Omega} |w| \, ds \\
&\lesssim \left(h^{-1}\|z - \partial_i u\|_{1,\Omega} + \|u\|_{3,\Omega} + h^{-1/2}|u|_{2,\infty,\partial\Omega}\right) \|w\|_{0,\Omega}.
\end{aligned}$$

Also

$$(z, w) = (z - \partial_i u, w) + (\partial_i u, w) \lesssim (h^{-1}\|z - \partial_i u\|_{1,\Omega} + \|u\|_{3,\Omega})\|w\|_{0,\Omega}.$$

Thus for $z \in \mathcal{V}_h$,

$$\|Az\| \lesssim h^{-1}\|z - \partial_i u\|_{1,\Omega} + \|u\|_{3,\Omega} + h^{-1/2}|u|_{2,\infty,\partial\Omega},$$

completing the proof. □

LEMMA 2.3. *Suppose that for $v \in \mathcal{V}_h$, and some $0 < \alpha \leq 1$, we have*

$$\begin{aligned}
\|v\| &\leq \omega(h, v), \\
\|v\|_{-\alpha} \equiv \|A^{-\alpha/2} v\| &\leq (\mathcal{C}h)^{\alpha} \omega(h, v).
\end{aligned}$$

Then

$$\|S^m v\| \leq \varepsilon_m \, \omega(h, v),$$

where

$$\varepsilon_m = \begin{cases} \kappa^{\alpha/2} f(m, \alpha/2) \lesssim m^{-\alpha/2} & \text{for } m > (\kappa - 1)\alpha/2 \\ [(\kappa - 1)/\kappa]^m & \text{for } m \leq (\kappa - 1)\alpha/2 \end{cases}$$

and $\kappa = (\mathcal{C}h)^2 \lambda$.

Proof. Let $0 \leq \beta \leq \alpha$. Then from the Hölder inequality

$$\|v\|_{-\beta} \leq \|v\|_{-\alpha}^{\beta/\alpha} \|v\|^{1 - \beta/\alpha}$$

and the hypotheses of the lemma, it follows that

$$\|v\|_{-\beta} \leq (\mathcal{C}h)^{\beta} \omega(h, v)$$

for $0 \leq \beta \leq \alpha$.

Now,

$$\begin{aligned}
\|S^m v\| &= \lambda^{\beta/2} \|S^m (I - S)^{\beta/2} A^{-\beta/2} v\| \\
&\leq \lambda^{\beta/2} \max_{s \in [0,1]} [s^m (1 - s)^{\beta/2}] \|A^{-\beta/2} v\| \\
&\leq \lambda^{\beta/2} f(m, \beta/2) (\mathcal{C}h)^{\beta} \omega(h, v) \\
&\leq \kappa^{\beta/2} f(m, \beta/2) \omega(h, v).
\end{aligned}$$

where $\kappa = (\mathcal{C}h)^2\lambda$. We now minimize this bound with respect to β on the interval $0 \le \beta \le \alpha$.

$$\begin{aligned}\frac{\partial \kappa^{\beta/2} f(m,\beta/2)}{\partial \beta} &= \frac{1}{2}\log\{(\kappa\beta)/(2m+\beta)\}\cdot\kappa^{\beta/2}f(m,\beta/2) = 0\\ &\Leftrightarrow \kappa\beta/(2m+\beta) = 1\\ &\Rightarrow \beta = 2m/(\kappa-1)\end{aligned}$$

There are two cases: the first is when $2m/(\kappa-1) > \alpha$. Here the minimum occurs at $\beta = \alpha$. Hence, for $m > (\kappa-1)\alpha/2$,

$$\varepsilon_m = \kappa^{\alpha/2} f(m,\alpha/2).$$

The second case is when $2m/(\kappa-1) \le \alpha$. Here $\beta = 2m/(\kappa-1)$ and

$$\varepsilon_m = \left(\frac{\kappa-1}{\kappa}\right)^m.$$

□

LEMMA 2.4. *Let $w \in H^1(\Omega)$. Then, for $1/2 < \alpha \le 1$,*

$$\|S^m Q_h \partial_i w\|_{0,\Omega} \lesssim \varepsilon_m \left(h^{-1}\|w\|_{0,\Omega} + \|w\|_{1,\Omega} + h^{-\alpha}\|w\|_{0,\infty,\partial\Omega}\right),$$

with ε_m defined as in Lemma 2.3.

Proof. Our plan is to apply Lemma 2.3 to $v = Q_h\partial_i w$. Note

$$\|v\|_{-\alpha} = \|Q_h\partial_i w\|_{-\alpha} = \sup_{\phi\in\mathcal{V}_h}\frac{(Q_h\partial_i w,\phi)}{\|\phi\|_\alpha} = \sup_{\phi\in\mathcal{V}_h}\frac{(\partial_i w,\phi)}{\|\phi\|_\alpha}.$$

Using integration by parts,

$$\begin{aligned}(\partial_i w,\phi) &= -(w,\partial_i\phi) + \int_{\partial\Omega} w\phi n_i\,ds\\ &\lesssim \|w\|_{0,\Omega}\|\phi\|_{1,\Omega} + \|w\|_{0,\infty,\partial\Omega}\int_{\partial\Omega}|\phi|\,ds\\ &\lesssim h^{\alpha-1}\|w\|_{0,\Omega}\|\phi\|_{\alpha,\Omega} + \|w\|_{0,\infty,\partial\Omega}\|\phi\|_{\alpha,\Omega}\\ &\lesssim (h^{\alpha-1}\|w\|_{0,\Omega} + \|w\|_{0,\infty,\partial\Omega})\|\phi\|_{\alpha,\Omega}.\end{aligned}$$

Thus

$$\|v\|_{-\alpha,\Omega} \lesssim h^\alpha\omega(h,v)$$

with

$$\omega(h,v) = h^{-1}\|w\|_{0,\Omega} + \|w\|_{1,\Omega} + h^{-\alpha}\|w\|_{0,\infty,\partial\Omega}.$$

Since,

$$\|v\|_{0,\Omega} = \|Q_h\partial_i w\|_{0,\Omega} \le \omega(h,v),$$

the desired estimate now follows from Lemma 2.3. □

LEMMA 2.5. *Let $u \in H^3(\Omega) \cap W^{2,\infty}(\Omega)$. Then for any $v_h \in \mathcal{V}_h$, and $1/2 < \alpha \leq 1$, we have*

$$\|\nabla u - S^m Q_h \nabla v_h\|_{0,\Omega} \lesssim m h^{3/2} \left(h^{1/2} \|u\|_{3,\Omega} + |u|_{2,\infty,\partial\Omega} \right) + \varepsilon_m \left(h^{-1} \|u - v_h\|_{0,\Omega} + \|u - v_h\|_{1,\Omega} + h^{-\alpha} \|u - v_h\|_{0,\infty,\partial\Omega} \right),$$

with ε_m defined as in Lemma 2.3.

Proof. By the triangle inequality:

$$\|\partial_i u - S^m Q_h \partial_i v_h\|_{0,\Omega} \leq \|(I - Q_h)\partial_i u\|_{0,\Omega} + \|(I - S^m) Q_h \partial_i u\|_{0,\Omega} + \|S^m Q_h \partial_i (u - v_h)\|_{0,\Omega}.$$

We now estimate these three terms. The first term is easy; by standard arguments

$$\|(I - Q_h)\partial_i u\|_{0,\Omega} \lesssim h^2 \|u\|_{3,\Omega}.$$

The second is estimated by Lemma 2.2 with $z = Q_h \partial_i u$. For the third, we apply Lemma 2.4 with $w = u - v_h$. ∎

In the case that $v_h = u_h \in \mathcal{V}_h \cap H_0^1(\Omega)$ is the finite element approximation to $u \in H_0^1(\Omega)$, the boundary terms vanish and

$$\|\nabla u - S^m Q_h \nabla v_h\|_{0,\Omega} \lesssim h(mh + \varepsilon_m) \|u\|_{3,\Omega}.$$

In the more general case, if $v_h = u_h \in \mathcal{V}_h$ and $1/2 < \alpha < 1$, we use the well-known L^∞ norm estimate for the linear finite element approximation to obtain:

$$h^{-\alpha} \|u - u_h\|_{0,\infty,\partial\Omega} \lesssim h^{1-\alpha} |\log h| h |u|_{2,\infty,\Omega} \lesssim h |u|_{2,\infty,\Omega}$$

and, hence

$$\|\nabla u - S^m Q_h \nabla v_h\|_{0,\Omega} \lesssim h(mh^{1/2} + \varepsilon_m)(\|u\|_{3,\Omega} + |u|_{2,\infty,\Omega}).$$

Similar estimates hold for the case $v = u_I$. We now turn to the main theorems in this section. This theorem is based only on the results developed in this section, and summarizes the above discussion.

THEOREM 2.6. *Let $u \in H^3(\Omega) \cap W^{2,\infty}(\Omega)$ and $u_h \in \mathcal{V}_h$ be an approximation of u satisfying*

$$\|u - u_h\|_{k,\Omega} \lesssim h^{2-k} |u|_{2,\Omega}, \quad k = 0, 1,$$
$$\|u - u_h\|_{0,\infty\Omega} \lesssim h^2 |\log h| |u|_{2,\infty\Omega}.$$

Then

$$\|\nabla u - S^m Q_h \nabla u_h\|_{0,\Omega} \lesssim h(mh^{1/2} + \varepsilon_m) \left(\|u\|_{3,\Omega} + |u|_{2,\infty,\Omega} \right),$$

where ε_m is defined as in Lemma 2.3 and $1/2 < \alpha < 1$.

This theorem combines results from this section with our earlier superconvergence results.

THEOREM 2.7. *Let $u \in W^{3,\infty}(\Omega)$ and assume the hypotheses of Theorem 2.1. Then*

$$\|\nabla u - S^m Q_h \nabla u_I\|_{0,\Omega} \lesssim h \left(\min(h^{\min(1,\sigma)} |\log h|, \varepsilon_m) + m\, h^{1/2} \right) \|u\|_{3,\infty,\Omega}, \tag{2.9}$$

$$\|\nabla u - S^m Q_h \nabla u_h\|_{0,\Omega} \lesssim h \left(\min(h^{\min(1,\sigma)} |\log h|, \varepsilon_m) + m\, h^{1/2} \right) \|u\|_{3,\infty,\Omega}. \tag{2.10}$$

where ε_m is defined as in Lemma 2.3 and $1/2 < \alpha < 1$.

Proof. Our proof combines Lemma 2.5 and Theorem 2.1. We first use the triangle inequality

$$\|\partial_i u - S^m Q_h \partial_i u_I\|_{0,\Omega} \leq \|(I-Q_h)\partial_i u\|_{0,\Omega} + \|(I-S^m)Q_h\partial_i u\|_{0,\Omega} + \|S^m Q_h \partial_i (u-u_I)\|_{0,\Omega}.$$

The first two terms are estimated as in Lemma 2.5. For the third term, we can first use Theorem 2.1 as

$$\begin{aligned}\|S^m Q_h \partial_i (u-u_I)\|_{0,\Omega} &\lesssim \|Q_h \partial_i (u-u_I)\|_{0,\Omega} \\ &\lesssim \|\partial_i u - Q_h \partial_i u_I\|_{0,\Omega} + \|(I-Q_h)\partial_i u\|_{0,\Omega} \\ &\lesssim h^{1+\min(1,\sigma)} |\log h|^{1/2} \|u\|_{3,\infty,\Omega} + h^2 \|u\|_{3,\Omega}.\end{aligned}$$

The third term can also be estimated as in Lemma 2.5. Taken together, these estimates establish (2.9). The proof of (2.10) is identical. □

We conclude with a few implementation details. First, with respect to the selection of the critical parameter m: Balancing the terms

$$\left(\frac{\kappa-1}{\kappa}\right)^m \approx m h^{-1/2}$$

suggests m should grow in a logarithmic-like fashion as the mesh is refined. On the other hand, in our empirical investigations, we have found that taking $m \leq 2$ has been adequate for scalar PDE equations involving $O(10^5)$ unknowns, which suggests that a simple fixed strategy is good enough for most purposes.

Second, with respect to the L^2 projection: This linear system is solved approximately by an iterative method, in our case, symmetric Gauss-Seidel with conjugate gradient acceleration (SGSCG). The mass matrix is assembled in the standard nodal basis and is sparse and diagonally dominant, so convergence is very rapid; typically 4–6 iterations are sufficient. In the context of an adaptive refinement feedback loop, the initial guess is taken as zero for the first (coarsest) mesh, and interpolated from the previous mesh at all subsequent refinement steps. The overall complexity of this step is thus $O(N)$ for a mesh with N vertices. If necessary, this step could be made more efficient (in terms of the size of the constant, not the order of complexity) by using some standard mass lumping scheme to construct a diagonal approximation to the mass matrix. This would also make the calculation *local* rather than global.

Third, with respect to the smoothing steps: We do not compute the constant λ exactly. In fact, we use a Jacobi-conjugate gradient (JCG) iteration. The stiffness matrix A corresponding to the operator $-\Delta$ is assembled in the nodal basis; this matrix is symmetric, positive semi-definite with a one dimensional kernel corresponding the the constant function. (We used the complete H^1 inner product in our analysis to avoid the technical complications introduced by a nontrivial kernel). Then m JCG steps are applied to the linear system $Ax = 0$, with initial guess corresponding the the finite element function $Q_h \partial_i u_h$. Our default choice is $m = 2$ iterations; thus this step also has complexity $O(N)$.

3. An A Posteriori Error Estimator. In this section we use the recovered gradient to develop an a posteriori error estimator. The obvious choice for a global a posteriori error estimator is to approximate $\|\nabla(u-u_h)\|_{0,\Omega}$ by $\|(I-S^m Q_h)\nabla u_h\|_{0,\Omega}$. In Theorem 3.1, we show that this is indeed a good approximation.

THEOREM 3.1. *Assume the hypotheses of Theorem 2.7.*

$$\|\nabla(u-u_h)\|_{0,\Omega} \le \|(I-S^mQ_h)\nabla u_h\|_{0,\Omega} + Ch\left(\min(h^{\min(1,\sigma)}|\log h|,\varepsilon_m) + m\,h^{1/2}\right)\|u\|_{3,\infty,\Omega}, \quad (3.1)$$

$$\|(I-S^mQ_h)\nabla u_h\|_{0,\Omega} \le \|\nabla(u-u_h)\|_{0,\Omega} + Ch\left(\min(h^{\min(1,\sigma)}|\log h|,\varepsilon_m) + m\,h^{1/2}\right)\|u\|_{3,\infty,\Omega}, \quad (3.2)$$

where ε_m is defined as in Lemma 2.3 for $1/2 < \alpha < 1$. Furthermore, if there exists a positive constant $c_0(u)$ independent of h such that

$$\|\nabla(u-u_h)\|_{0,\Omega} \ge c_0(u)h, \quad (3.3)$$

then

$$\left|\frac{\|(I-S^mQ_h)\nabla u_h\|_{0,\Omega}}{\|\nabla(u-u_h)\|_{0,\Omega}} - 1\right| \lesssim \min(h^{\min(1,\sigma)}|\log h|,\varepsilon_m) + m\,h^{1/2}. \quad (3.4)$$

Proof. The proof of (3.1)-(3.2) is just a simple application of the triangle inequalities

$$\|\nabla(u-u_h)\|_{0,\Omega} \le \|(I-S^mQ_h)\nabla u_h\|_{0,\Omega} + \|\nabla u - S^mQ_h\nabla u_h\|_{0,\Omega},$$
$$\|(I-S^mQ_h)\nabla u_h\|_{0,\Omega} \le \|\nabla(u-u_h)\|_{0,\Omega} + \|\nabla u - S^mQ_h\nabla u_h\|_{0,\Omega},$$

and Theorem 2.7. Estimate (3.4) follows from (3.1)-(3.2) and the assumption (3.3). ∎

Taken together, (3.1)-(3.2) show that if the true error is first order, $\|\nabla(u-u_h)\|_{0,\Omega} = O(h)$, then the a posteriori error estimate $\|(I-S^mQ_h)\nabla u_h\|_{0,\Omega}$ will also be $O(h)$. In particular, given a superconvergent approximation to ∇u, we can expect the effectivity ratio $\|(I-S^mQ_h)\nabla u_h\|_{0,\Omega}/\|\nabla(u-u_h)\|_{0,\Omega}$ to be close to unity. Furthermore, in this case Theorem 3.1 shows the a posteriori error estimate will be asymptotically exact.

In terms of local error indicators, an obvious choice would be to estimate the local error in a given element τ by $\|(I-S^mQ_h)\nabla u_h\|_{0,\tau}$. For practical reasons discussed below, we prefer an alternative approach where the recovered gradient is used to approximate the Hessian matrix of second derivatives of u. By way of motivation, we note that for an $O(h^{2\sigma})$ irregular mesh.

$$\begin{aligned}\|\nabla(u-u_h)\|_{0,\Omega} &\le \|\nabla(u-u_q)\|_{0,\Omega} + \|\nabla(u_q-u_I)\|_{0,\Omega} + \|\nabla(u_I-u_h)\|_{0,\Omega}\\ &\le C(u)h^{1+\min(1,\sigma)}|\log h| + \|\nabla(u_q-u_I)\|_{0,\Omega},\\ \|\nabla(u_q-u_I)\|_{0,\Omega} &\le \|\nabla(u_q-u)\|_{0,\Omega} + \|\nabla(u-u_h)\|_{0,\Omega} + \|\nabla(u_h-u_I)\|_{0,\Omega}\\ &\le C(u)h^{1+\min(1,\sigma)}|\log h| + \|\nabla(u-u_h)\|_{0,\Omega}.\end{aligned}$$

From this pair of estimates, it follows that $\|\nabla(u_q-u_I)\|_{0,\Omega} = O(h)$ if and only if $\|\nabla(u-u_h)\|_{0,\Omega} = O(h)$, and that in this case $\|\nabla(u_q-u_I)\|_{0,\Omega}$ is asymptotically exact.

The function $u_q - u_I$ is a locally defined, piecewise quadratic polynomial with value zero at all vertices of the mesh. Let a canonical element $\tau \in \mathcal{T}_h$ have vertices $\boldsymbol{p}_k^t = (x_k, y_k)$, $1 \le k \le 3$, oriented counterclockwise, and corresponding nodal basis

functions (barycentric coordinates) $\{\psi_k\}_{k=1}^3$. Let $\{e_k\}_{k=1}^3$ denote the edges of element τ, $\{\boldsymbol{n}_k\}_{k=1}^3$ the unit outward normal vectors, $\{\boldsymbol{t}_k\}_{k=1}^3$ the unit tangent vectors with counterclockwise orientation, and $\{\ell_k\}_{k=1}^3$ the edge lengths (see Figure 3.1). Let $q_k = \psi_{k+1}\psi_{k-1}$ denote the quadratic bump function associated with edge k of τ, where $(k-1, k, k+1)$ is a cyclic permutation of $(1, 2, 3)$. Thus in element τ, $u_q - u_I$ is a linear combination of the quadratic bump functions associated with the edge midpoints of the element,

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

where

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

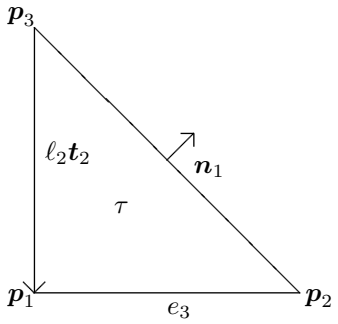

FIG. 3.1. *Parameters associated with the triangle τ.*

In our local error indicator, we simply approximate the second derivatives in the Hessian matrix M_τ using gradients of $S^m Q_h \partial_i u_h$. In particular, let

$$\tilde{M}_\tau = -\frac{1}{2} \begin{pmatrix} \partial_1 S^m Q_h \partial_1 u_h & \partial_1 S^m Q_h \partial_2 u_h \\ \partial_2 S^m Q_h \partial_1 u_h & \partial_2 S^m Q_h \partial_2 u_h \end{pmatrix},$$
$$\bar{M}_\tau = \frac{\alpha_\tau}{2} (\tilde{M}_\tau + \tilde{M}_\tau^t),$$

where $\alpha_\tau > 0$ is a constant described below. For the case of meshes that are $O(h^{2\sigma})$ irregular, we can have $m = 0$, but for general shape regular meshes, we have $m > 0$. In either case, the local error estimate ϵ_τ is given by

$$\epsilon_\tau = \sum_{k=1}^{3} \ell_k^2 \boldsymbol{t}_k^t \bar{M}_\tau \boldsymbol{t}_k \, q_k(x, y). \tag{3.5}$$

The normalization constant α_τ is chosen such that the local error indicator η_τ satisfies

$$\eta_\tau \equiv \|\nabla \epsilon_\tau\|_{0,\tau} = \|(I - S^m Q_h) \nabla u_h\|_{0,\tau}.$$

Normally we expect that $\alpha_\tau \approx 1$, which is likely to be the case in regions where the Hessian matrix for the true solution is well defined. Near singularities, u is not smooth and we anticipate difficulties in estimating the Hessian. For elements near such singularities, α_τ provides a heuristic for partly compensating for poor approximation.

The form of our a posteriori error estimate (3.5) is quite useful in practice. It explicitly shows the dependence on the shape, size, and orientation of the elements, as well as the dependence on the second derivatives of u. This leads to many interesting algorithms for adaptive mesh smoothing and topology modification (e.g., "edge flipping") [8]. For example, if $\bar{M}_\tau$ is assumed constant, then ϵ_τ and η_τ^2 are rational functions of vertex locations, and derivatives with respect to the vertex locations are easily computed.

Using ϵ_τ also provides a simple, robust, and elegant solution to an important practical problem for adaptive mesh refinement schemes: how to provide error estimates for the refined elements without immediately resolving the global problem. In the past, most schemes were based on crude (or not-so-crude) extrapolation ideas, using the error indicator of the parent element as a basis. In most such schemes it is difficult to take into account the details of the geometry, and they tend to become very inaccurate after only a few levels of refinement. On the other hand, with our error estimator, the children elements inherit only the Hessian matrix from the parent, and all the geometrical information is derived from the refined elements themselves. Thus it is possible to have many levels of refinement before the approximation breaks down and a new global solution is required. This has proved to be very effective in the PLTMG 8.0 package, which employs this scheme, but uses a different a posteriori error estimate to compute the approximate Hessian matrix [8]. In Theorem 3.2, we show that our recovered gradients can provide reasonable approximation to the Hessian.

THEOREM 3.2. *Assume the hypotheses of Theorem 2.7. Then*

$$\|\partial_i(\partial_k u - S^m Q_h \partial_k u_h)\|_{0,\Omega} \lesssim \left(\min(h^{\min(1,\sigma)}|\log h|, \varepsilon_m) + m\, h^{1/2}\right) \|u\|_{3,\infty,\Omega}, \quad (3.6)$$

where ε_m is defined as in Lemma 2.3 for $1/2 \le \alpha < 1$.

Proof. Let $z = I_h \partial_k u \in \mathcal{V}_h$. Then

$$\begin{aligned}
\|\partial_i(\partial_k u - S^m Q_h \partial_k u_h)\|_{0,\Omega} &\le \|\partial_i(\partial_k u - z)\|_{0,\Omega} + \|\partial_i(z - S^m Q_h \partial_k u_h)\|_{0,\Omega} \\
&\lesssim h\|u\|_{3,\Omega} + h^{-1}\|z - S^m Q_h \partial_k u_h\|_{0,\Omega} \\
&\lesssim h\|u\|_{3,\Omega} + h^{-1}\left(\|z - \partial_k u\|_{0,\Omega} + \|\partial_k u - S^m Q_h \partial_k u_h\|_{0,\Omega}\right) \\
&\lesssim \left(\min(h^{\min(1,\sigma)}|\log h|, \varepsilon_m) + m\, h^{1/2}\right) \|u\|_{3,\infty,\Omega}.
\end{aligned}$$

□

4. Numerical Experiments. In this section, we present some numerical illustrations of our recovery scheme in the cases of uniform and adaptively refined (nonuniform) meshes. Our gradient recovery scheme and a posteriori error estimate were implemented in the PLTMG package [6], which was then used for our numerical experiments. The experiments were done on an SGI Octane using double precision arithmetic.

In our first example, we consider the solution of the problem

$$\begin{aligned}
-\Delta u &= f \qquad \text{in } \Omega = (0,1) \times (0,1), \\
u &= g \qquad \text{on } \partial\Omega,
\end{aligned}$$

where f and g are chosen such that $u = e^{x+y}$ is the exact solution. This is a very smooth solution that satisfies all the assumptions of our theory. Here we will compare

the recovery scheme with $m = 2$ smoothing steps, for the case of uniform and adaptive meshes. We begin with a uniform 3×3 mesh consisting of eight right triangles as shown in Figure 4.1. Elements in Figure 4.1 are colored according to size; this allows one to obtain some impression of the structure of highly refined meshes with many elements, even if individual elements can no longer be resolved.

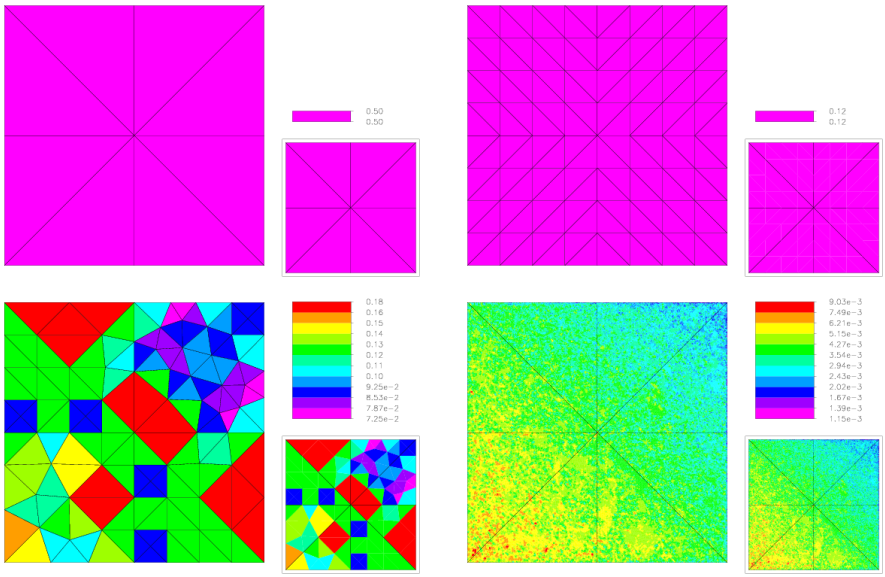

FIG. 4.1. *Top left:* 3×3 *initial mesh. Top right: uniform refinement with* $nt = 128$. *Bottom left: adaptive refinement with* $nt = 134$. *Bottom right: adaptive refinement with* $nt = 130961$. *Elements are colored according to size.*

In Table 4.1, we record the results of the computation. We give the error as a function of the number of elements, choosing targets for the adaptive refinement procedure to produce adaptive meshes with similar numbers of elements to the uniform refinement case. The values are defined as follows:

$$
\begin{aligned}
L2 &= \|u - u_h\|_{0,\Omega}, \\
H1 &= \|\nabla u - S^m Q_h \nabla u_h\|_{0,\Omega}, \\
\widetilde{H1} &= \|\nabla (u - u_h)\|_{0,\Omega}, \\
Ef &= \frac{\|(I - S^m Q_h)\nabla u_h\|_{0,\Omega}}{\|\nabla (u - u_h)\|_{0,\Omega}}.
\end{aligned}
$$

For the cases of $L2$, $H1$ and $\widetilde{H1}$, we made a least squares fit of the data to a function of the form $F(N) = CN^{-p/2}$ to estimate the order of convergence p. All integrals were approximated using a 12-point order 7 quadrature formula applied to each triangle.

What is most striking is the similarity in the data. $L2$ is approximately the same for both uniform and adaptive refinement, while $H1$ is slightly better in the adaptive case. This is consistent with our strategy which adaptively refines with respect to $\|\nabla \epsilon_\tau\|_{0,\tau}$. Nonetheless, both cases exhibit some superconvergence for the recovered gradients. This is further supported by noting that the effectivity ratios Ef suggest asymptotic exactness of the a posteriori error estimates.

TABLE 4.1
Error estimates for the case $m = 2$.

adaptive meshes				uniform meshes			
nt	$L2$	$H1$	Ef	nt	$L2$	$H1$	Ef
8	1.5e-1	1.7e 0	1.68	8	1.5e-1	1.7e 0	1.68
34	4.9e-2	5.9e-1	1.36	32	3.8e-2	8.7e-1	1.74
134	1.3e-2	2.9e-1	1.54	128	9.6e-3	3.6e-1	1.50
510	2.4e-3	7.1e-2	1.17	512	2.4e-3	1.6e-1	1.41
2037	5.2e-4	2.4e-2	1.09	2048	6.0e-4	6.7e-2	1.30
8148	1.1e-4	7.1e-3	1.04	8192	1.5e-4	2.6e-2	1.20
32683	2.7e-5	2.0e-3	1.01	32768	3.8e-5	1.0e-2	1.12
130961	7.0e-6	6.2e-4	1.00	131072	9.4e-6	3.7e-3	1.07
	$L2$	$H1$	$\widetilde{H1}$		$L2$	$H1$	$\widetilde{H1}$
order	2.06	1.76	1.04		2.03	1.42	1.01

In Table 4.2, we show the effect of varying the number of smoothing steps. To reduce the amount of data, we report only the case of adaptive meshes. Since the a posteriori error estimates are used to create the meshes, the meshes differ for each value of m, but at level K have $nt \approx 2^{2K+1}$ elements. For $m = 0$, we note only slight superconvergence; thus although the meshes are shape regular and quasi-uniform, apparently $\sigma \approx 0$. In contrast, uniform meshes for $m = 0$ have a computed order of convergence for $H1$ of 1.52, essentially that predicted by our theory. However, the data show that the situation for adaptive meshes improves dramatically for $m = 1, 2$. For $m = 10$, one can see the effects of "too many" smoothings; Ef become more erratic, and $H1$ *increases* for some of the coarser refinement steps. But even in this case, for more refined meshes (e.g. $K = 8$), Ef again appears to be converging towards one. This is likely due to well known (and in this case extremely useful) effect of the smoothing iteration "slowing down" quickly as h becomes smaller.

TABLE 4.2
Order of convergence as a function of m for adaptive meshes.

	m=0		m=1		m=2		m=3		m=10	
K	$H1$	Ef	$H1$	Ef	$H1$	Ef	$H1$	Ef	$H1$	Ef
1	6.1e 0	0.83	7.8e-1	1.16	1.7e 0	1.68	1.7e 0	1.70	1.7e 0	1.70
2	2.3e-1	0.92	3.2e-1	1.00	5.9e-1	1.36	1.3e 0	2.34	1.8e 0	3.67
3	7.8e-2	0.94	1.1e-1	1.06	2.9e-1	1.54	3.8e-1	1.64	1.8e 0	6.75
4	3.8e-2	0.95	3.4e-2	1.02	7.1e-2	1.17	1.5e-1	1.49	1.1e 0	7.09
5	1.7e-2	0.96	1.1e-2	1.01	2.4e-2	1.09	4.8e-2	1.26	3.2e-1	4.05
6	7.2e-3	0.97	3.7e-3	1.00	7.1e-3	1.04	1.3e-2	1.09	1.2e-1	3.63
7	3.0e-3	0.98	1.4e-3	1.00	2.0e-3	1.01	3.4e-3	1.03	3.2e-2	2.27
8	1.5e-3	0.98	5.7e-4	1.00	6.2e-4	1.00	8.9e-4	1.01	7.0e-3	1.34
	$H1$	$\widetilde{H1}$	$H1$	$\widetilde{H1}$	$H1$	$\widetilde{H1}$	$H1$	$\widetilde{H1}$	$H1$	$\widetilde{H1}$
order	1.16	1.04	1.39	1.03	1.76	1.04	1.92	1.07	2.01	1.08

In Table 4.3, we explore the effect of "lumping" the mass matrix in the L^2 projection step. In particular, the mass matrix was replaced by a diagonal matrix with diagonal entries given by the sum of all nonzero entries of the corresponding row of

the mass matrix. In Table 4.3, we see results that are quite comparable to those of Table 4.1, although the gradient errors are generally slightly larger. Nonetheless, these results suggest that our gradient recovery algorithm could be modified to use only local calculations without much loss in effectiveness.

TABLE 4.3
The effect of a lumped *mass matrix.*

adaptive meshes				uniform meshes			
nt	$L2$	$H1$	Ef	nt	$L2$	$H1$	Ef
8	1.5e-1	1.7e 0	1.69	8	1.5e-1	1.7e 0	1.69
34	4.9e-2	8.2e-1	1.70	32	3.8e-2	1.1e 0	2.02
134	1.3e-2	3.0e-1	1.57	128	9.6e-3	5.2e-1	1.96
514	2.8e-3	8.9e-2	1.23	512	2.4e-3	2.2e-1	1.74
2036	5.5e-4	2.7e-2	1.11	2048	6.0e-4	9.3e-2	1.55
8148	1.2e-4	7.7e-3	1.04	8192	1.5e-4	3.7e-2	1.37
32676	2.8e-5	2.3e-3	1.01	32768	3.8e-5	1.4e-2	1.23
130904	6.8e-6	8.2e-4	1.01	131072	9.4e-6	5.1e-3	1.13
	$L2$	$H1$	$\widetilde{H1}$		$L2$	$H1$	$\widetilde{H1}$
order	2.10	1.64	1.05		2.03	1.43	1.01

In our second example, we consider the nonlinear problem

$$\begin{aligned} -\nabla \cdot (a\nabla u) + e^u &= f \qquad \text{in } \Omega = (0,1) \times (0,1), \\ u &= 0 \qquad \text{on } \partial\Omega, \end{aligned}$$

where a is the 2×2 diagonal matrix

$$a = \begin{pmatrix} .01 & \\ & 1 \end{pmatrix}.$$

The function f is chosen such that $u = x(1-x)^3 y^5 (1-y)$ is the exact solution. We repeat the same computations as in the first example, with uniform and adaptive meshes. The uniform meshes are identical to those of the first example. Some of the adaptive meshes are shown in Figure 4.2. The numerical results are summarized in Table 4.4.

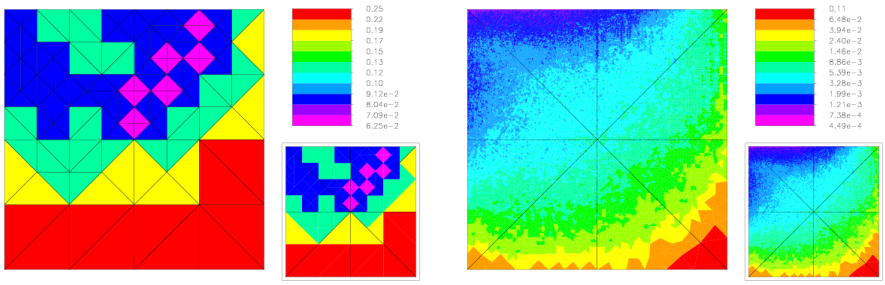

FIG. 4.2. *Left: adaptive refinement with* $nt = 138$. *Right: adaptive refinement with* $nt = 131112$. *Elements are colored according to size.*

This problem is more difficult than the first in several respects. The diffusion is anisotropic and the operator is nonlinear. The solution is smooth but generally has

TABLE 4.4
Error estimates for the case $m = 2$.

adaptive meshes				uniform meshes			
nt	$L2$	$H1$	Ef	nt	$L2$	$H1$	Ef
8	1.9e-3	1.7e-2	0.30	8	1.9e-3	1.7e-2	0.30
32	1.0e-3	1.6e-2	0.84	32	1.0e-3	1.6e-2	0.84
138	2.7e-4	1.1e-2	1.46	128	3.8e-4	1.2e-2	1.42
531	2.4e-3	3.7e-3	1.67	512	1.1e-4	7.8e-3	1.89
2060	4.4e-5	1.0e-3	1.32	2048	3.0e-5	4.1e-3	2.09
8203	1.2e-5	2.6e-4	1.09	8192	7.7e-6	1.9e-3	2.01
32736	8.6e-7	7.6e-5	1.00	32768	1.9e-6	7.7e-4	1.79
131112	2.5e-7	3.0e-5	0.98	131072	4.9e-7	3.0e-4	1.53
	$L2$	$H1$	$\widetilde{H1}$		$L2$	$H1$	$\widetilde{H1}$
order	1.83	1.58	1.05		2.01	1.30	1.02

larger derivatives than the first example. Nonetheless, we see a similar behavior of the gradient recovery scheme and a posteriori error estimate. In this example, the adaptive meshes are more strongly graded than in the first example, suggesting that localization and equilibration of the error are more important effects for superconvergence of our gradient recovery procedure than geometric uniformity in the mesh.

In our third example, we consider the problem

$$\begin{aligned} -\Delta u &= 0 \qquad && \text{in } \Omega, \\ u &= g \qquad && \text{on } \partial\Omega_1, \\ u_n &= 0 \qquad && \text{on } \partial\Omega_2, \end{aligned}$$

where Ω is a circle of radius one centered at the origin, and with a crack along the positive x-axis $0 \le x \le 1$. $\partial\Omega_2$ is the bottom edge of the crack, and $\partial\Omega_1 = \partial\Omega - \partial\Omega_2$. The function g is chosen such that the exact solution is $u = r^{1/4}\sin(\theta/4)$, the leading term of the singularity associated with the interior angle of 2π and change in boundary conditions at the origin. In Figure 4.3 we illustrate the initial mesh, and several of the uniformly and adaptively refined meshes.

Convergence results for uniform and adaptive refinement are reported in Table 4.5. The solution u is not smooth in this case ($u \in H^{5/4-\epsilon}(\Omega)$), and this is reflected in the results. For the case of uniform refinement, the 0.25 order of convergence of the gradient coincides with the smoothness of the solution. For the adaptive meshes, the order of convergence improves and seems to be approaching order one for the gradient. This sort of behavior is typical of a reasonable adaptive refinement procedure. However, even in this case, there is no apparent superconvergence.

Let Ω' now denote the union of all triangles in Ω with a least one vertex outside a circle of $r = 0.1$; note that this definition of Ω' is mesh dependent, but eventually Ω' excludes small triangles close to the singularity. In Table 4.6, we report results for the same computations but with the error calculation restricted to Ω'. For $L2'$, the results do not change much. However, for the gradients ($H1'$), the results are quite striking. For the case of adaptive refinement, the improvement is quite dramatic, in that away from the singularity, the gradient recovery scheme exhibits the same sort of behavior as for a smooth problem. For the case of uniform refinement, there is also some improvement in order of convergence, but the recovered gradient does

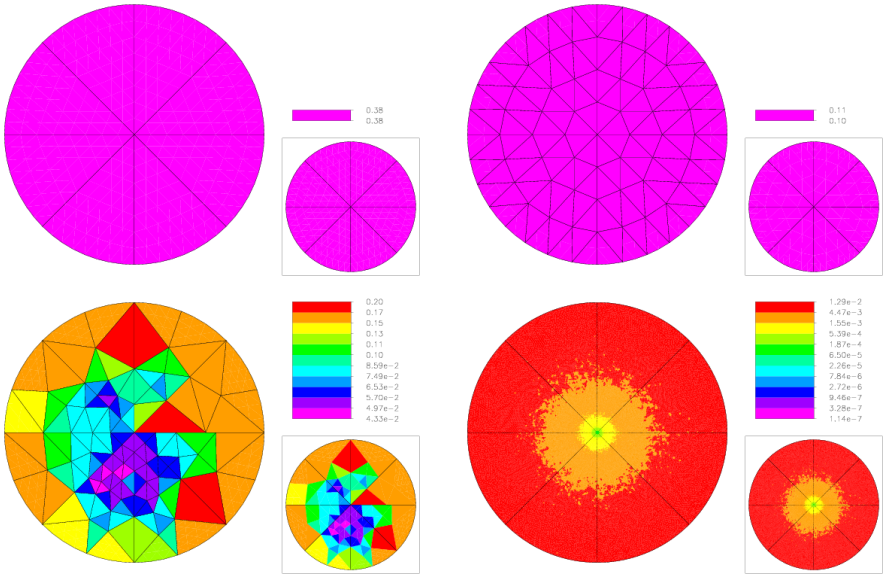

FIG. 4.3. *Top left: the initial mesh. Top right: uniform refinement with* $nt = 128$. *Bottom left: adaptive refinement with* $nt = 138$. *Bottom right: adaptive refinement with* $nt = 131105$. *Elements are colored according to size.*

TABLE 4.5
Error estimates for the case $m = 2$.

adaptive meshes				uniform meshes			
nt	$L2$	$H1$	Ef	nt	$L2$	$H1$	Ef
8	3.0e-1	7.1e-1	1.32	8	3.0e-1	7.1e-1	1.32
31	1.9e-1	6.0e-1	1.14	32	1.8e-1	5.9e-1	1.18
138	8.3e-2	5.3e-1	1.18	128	1.1e-1	5.4e-1	1.16
535	3.3e-2	3.9e-1	1.26	512	7.3e-2	4.6e-1	1.14
2091	1.1e-2	2.3e-1	1.22	2048	4.9e-2	3.9e-1	1.12
8242	2.9e-3	1.2e-1	1.28	8192	3.4e-2	3.3e-1	1.11
32832	6.8e-4	4.9e-2	1.08	32768	2.3e-2	2.8e-1	1.10
131105	1.3e-4	2.2e-2	1.11	131072	1.6e-2	2.3e-1	1.10
	$L2$	$H1$	$\widetilde{H1}$		$L2$	$H1$	$\widetilde{H1}$
order	2.23	1.15	1.10		0.54	0.25	0.27

not appear significantly more accurate than ∇u_h. We also note the quite different behavior of Ef' compared with Ef for the case on uniform refinement. Taken as a whole, it seems likely that the error in the uniform refinement case is not localized and pollution effects are still dominant even on the most refined meshes.

In our fourth example, we consider a problem with discontinuous coefficients

$$\begin{aligned} -\nabla \cdot (a \nabla u) &= f \qquad \text{in } \Omega, \\ u &= 0 \qquad \text{on } \partial\Omega, \end{aligned}$$

where Ω is once again the unit square $(0,1) \times (0,1)$. The scalar coefficient function $a(p) = 1$ for $p \in \Omega_1 \equiv (0,1/2) \times (0,1/2) \cup (1/2,1) \times (1/2,1)$, and $a(p) = 10^{-2}$ for

TABLE 4.6
The effect of the singularity.

adaptive meshes				uniform meshes			
nt	$L2'$	$H1'$	Ef'	nt	$L2'$	$H1'$	Ef'
8	3.0e-1	7.1e-1	1.32	8	3.0e-1	7.1e-1	1.32
31	1.9e-1	6.0e-1	1.14	32	1.8e-1	5.9e-1	1.18
138	8.3e-2	5.3e-1	1.18	128	1.1e-1	5.4e-1	1.16
535	3.0e-2	2.4e-1	0.92	512	7.3e-2	4.6e-1	1.14
2091	1.0e-2	4.3e-2	0.78	2048	4.6e-2	1.9e-1	0.69
8242	2.7e-3	1.2e-2	0.94	8192	3.1e-2	1.1e-1	0.28
32832	6.2e-4	3.0e-3	0.99	32768	2.1e-2	7.8e-2	0.19
131105	1.2e-4	7.9e-4	1.00	131072	1.5e-2	5.3e-2	0.13
	$L2'$	$H1'$	$\widetilde{H1}'$		$L2'$	$H1'$	$\widetilde{H1}'$
order	2.24	1.95	1.12		0.56	0.60	0.61

$p = \Omega - \Omega_1$. The function f is given by $f = 8\pi^2 \sin(2\pi x)\sin(2\pi y)$, and the exact solution u is given by $u = a^{-1}\sin(2\pi x)\sin(2\pi y)$. The initial mesh is that same as in the first two examples. In Table 4.7, we give the results for both uniform and adaptive meshes. Here we note no superconvergence for $H1$ in either case; indeed, in the uniform mesh case, $S^m Q_h \nabla u_h$ is much less accurate than ∇u_h in terms of order of convergence. The gradient of the exact solution is discontinuous along the lines $x = 1/2$ and $y = 1/2$. Since the mesh is aligned with the discontinuity, ∇u_h is able to capture this discontinuity with no problem. Since $Q_h \nabla u_h$ is continuous, the lack of superconvergence is due to the global L^2 projection; making local L^2 projections in each of the four subregions where $a(p)$ is constant would allow the projected gradient to remain discontinuous along the interfaces. In Table 4.8, we show the results for such a calculation. We computed $S^m \hat{Q}_h \nabla u_u$, where $\hat{Q}_h$ corresponds to the four local L^2 projections. Since we used a different projection, the adaptive meshes change slightly. Nonetheless, we see quite clearly allowing for the discontinuity in ∇u corrects the problems. We observe superconvergence in $H1'$ for both the uniform and adaptive meshes, as well as significant improvement in the effectivity ratios Ef'. We note in passing that away from the discontinuities, the approximation $S^m Q_h \nabla u_h$ is superconvergent. In this respect, the behavior is similar to the third example.

These numerical examples show that the effectiveness of our adaptive scheme does not necessarily depend critically on either the quasi-uniformity of the mesh or on the global regularity of the solution. But theoretically, there is still much work to do fill in the gaps. We believe that our quasi-uniformity assumption on the triangular meshes can be removed with some extra effort. Our results are very local in the sense that the domain Ω in our Theorems could be any subdomain of the actual physical domain. It will be much more challenging to have obtain theoretical results with more realistic assumptions on continuous solution's regularity. But heuristically speaking, at places where the solution is singular, the solution will have large (or infinite) W^3_∞ norm and, as a result, our error indicator will be large in those regions and hence the grid will be refined there.

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

TABLE 4.7
Error estimates for the case $m = 2$.

adaptive meshes				uniform meshes			
nt	$L2$	$H1$	Ef	nt	$L2$	$H1$	Ef
8	3.5e 1	3.2e 2	0.00	8	3.5e 1	3.2e 2	0.00
32	1.7e 1	2.9e 2	1.97	32	1.7e 1	3.0e 2	1.07
136	2.6e 0	2.2e 2	2.72	128	5.6e 0	2.5e 2	1.96
528	7.4e-1	1.1e 2	2.81	512	1.5e 0	1.7e 2	2.76
2071	2.2e-1	5.8e 1	2.67	2048	3.8e-1	9.0e 1	3.04
8143	6.9e-2	2.8e 1	2.42	8192	9.5e-2	5.4e 1	3.64
33102	2.0e-2	1.4e 1	2.36	32768	2.4e-2	3.7e 1	4.85
130809	6.1e-3	7.3e 0	2.31	131072	6.0e-3	2.6e 1	6.72
	$L2$	$H1$	$\widetilde{H1}$		$L2$	$H1$	$\widetilde{H1}$
order	1.78	1.00	0.94		2.04	0.58	1.02

TABLE 4.8
The effect of the discontinuity.

adaptive meshes				uniform meshes			
nt	$L2'$	$H1'$	Ef'	nt	$L2'$	$H1'$	Ef'
8	3.5e 1	3.2e 2	0.00	8	3.5e 1	3.2e 2	0.00
32	7.1e 0	3.1e 2	2.11	32	1.7e 1	3.1e 2	1.11
136	3.5e 0	1.7e 2	1.83	128	5.6e 0	2.3e 2	1.81
532	9.1e-1	7.6e 1	1.85	512	1.5e 0	1.5e 2	2.49
2071	1.6e-1	2.4e 1	1.57	2048	3.8e-1	7.2e 1	2.47
8092	4.2e-2	6.3e 0	1.20	8192	9.5e-2	2.4e 1	1.84
32486	9.4e-3	1.3e 0	1.05	32768	2.4e-2	7.3e 0	1.37
130586	2.4e-3	3.0e-1	1.01	131072	6.0e-3	2.2e 0	1.15
	$L2$	$H1$	$\widetilde{H1}$		$L2$	$H1$	$\widetilde{H1}$
order	2.10	2.16	1.09		2.04	1.68	1.02

opportunity to attend this workshop and for the stimulating scientific environment at MSRI. We also thank Lars B. Wahlbin for helpful discussions during the aforementioned workshop and the reviewers for their careful reviews and their suggestions for improvements.

REFERENCES

[1] M. AINSWORTH AND A. CRAIG, *A posteriori error estimators in the finite element method*, Numer. Math., 60 (1992), pp. 429–463.

[2] 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.

[3] 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.

[4] I. BABUŠKA AND T. STROUBOULIS, *The finite element method and its reliability*, Numerical Mathematics and Scientific Computation, Oxford Science Publications, 2001.

[5] 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.

[6] 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.
[7] R. E. Bank and C. C. Douglas, *Sharp estimates for multigrid rates of convergence with general smoothing and acceleration*, SIAM J. Numerical Analysis, 22 (1985), pp. 617–633.
[8] R. E. Bank and R. K. Smith, *Mesh smoothing using a posteriori error estimates*, SIAM J. Numerical Analysis, 34 (1997), pp. 979–997.
[9] R. E. Bank and A. Weiser, *Some a posteriori error estimators for elliptic partial differential equations*, Mathematics of Computation, 44 (1985), pp. 283–301.
[10] R. E. Bank and J. Xu, *Asymptotically exact a posteriori error estimators, part I: Grids with superconvergence*, SIAM J. Numerical Analysis, (submitted).
[11] C. Chen and Y. Huang, *High accuracy theory of finite element methods*, Hunan Science Press, Hunan, China, 1995. in Chinese.
[12] 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.
[13] 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.
[14] 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).
[15] 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.
[16] J. Li, *Convergence and superconvergence analysis of finite element methods on highly nonuniform anisotropic meshes for singularly perturbed reaction-diffusion problems*, Appl. Numer. Math., 36 (2001), pp. 129–154.
[17] Q. Lin and N. Yan, *The construction and analysis of high efficiency finite elements*, Hebei University Press, Hunan, China, 1996. in Chinese.
[18] R. Verfürth, *A Posteriori Error Estimation and Adaptive Mesh Refinement Techniques*, Teubner Skripten zur Numerik, B. G. Teubner, Stuttgart, 1995.
[19] L. B. Wahlbin, *Superconvergence in Galerkin finite element methods*, Springer-Verlag, Berlin, 1995.
[20] ———, *General principles of superconvergence in Galerkin finite element methods*, in Finite element methods (Jyväskylä, 1997), Dekker, New York, 1998, pp. 269–285.
[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.