SUPERCONVERGENT DERIVATIVE RECOVERY FOR LAGRANGE TRIANGULAR ELEMENTS OF DEGREE P ON UNSTRUCTURED GRIDS

RANDOLPH E. BANK[*], JINCHAO XU[†], AND BIN ZHENG[‡]

Abstract. In this paper, we develop a postprocessing derivative recovery scheme for the finite element solution u_h on general unstructured but shape regular triangulations. In the case of continuous piecewise polynomials of degree $p \geq 1$, by applying the global L^2 projection (Q_h) and a smoothing operator (S_h), the recovered p-th derivatives ($S_h^m Q_h \partial^p u_h$) superconverge to the exact derivatives ($\partial^p u$). Based on this technique we are able to derive a local error indicator depending only on the geometry of corresponding element and the $(p+1)$-st derivatives approximated by $\partial S_h^m Q_h \partial^p u_h$. We provide several numerical examples illustrating the effectiveness of our procedures. We also observe that higher order elements are likely to require more conservative refinement strategies to create meshes corresponding to optimal orders of convergence.

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

AMS subject classifications. 65N50, 65N30

1. Introduction. In this work we introduce a derivative recovery scheme for Lagrange triangular elements of degree p. It is an extension of the gradient recovery scheme for linear elements proposed by Bank and Xu [3]. The recovered p-th derivatives are shown to be superconvergent to the exact ones for general shape regular meshes. Due to the superconvergent property of this scheme, some a posteriori error estimates and local error indicators can be derived for mesh adaptation.

The recovery techniques for finite element analysis have been studied extensively in the literature [6, 7, 10, 12, 13, 16, 17]. The main goal of the recovery techniques is to construct better approximations of the solution function or derivative using certain postprocessing procedures. Typically these techniques involve some kind of local or global averaging, including local or global L^2 projection. Due to the superconvergence property, recovery techniques are often used to construct a posteriori error estimators (see e.g. [5, 11, 14, 15, 18]) which are asymptotically exact. For the literature regarding superconvergence analysis of recovery techniques we refer to [3] and the references therein.

Most recovery schemes are only concerned with the recovery of the gradient, the finite element solution itself and the second order derivatives. There was also some work on the recovery of higher order derivatives on uniform grids (see e.g. [4]). It is the purpose of the current work to recover $(p+1)$-st derivatives for Lagrange elements of degree $p \geq 1$ on unstructured grids.

Our development has two major components. First, we develop a postprocessing derivative recovery scheme for the finite element solution u_h on general shape regular triangulations. In particular, in Section 2 of this paper we compute $S_h^m Q_h \partial^p u_h$,

where S_h is an appropriate smoothing operator, $m \in \{1, 2, \ldots\}$ is the number of smoothing steps, and Q_h is the L^2 projection operator. The recovered p-th derivatives superconverge to the exact ones. In the case of a small number of smoothing steps (the most interesting case), Theorem 2.5 shows that

$$\|\partial^p u - S_h^m Q_h \partial^p u_h\|_{0,\Omega} \lesssim h \left(m h^{1/2} + \left[\frac{\kappa - 1}{\kappa} \right]^m \right) \left(\|u\|_{p+2,\Omega} + |u|_{p+1,\infty,\Omega} \right).$$

Here $\kappa > 1$ is a constant independent of h and u.

The second major component, presented in Section 3 of this paper, is the development of a posteriori error estimates based on the derivative recovery scheme. As an example, we discuss quadratic finite elements in detail. For the case of quadratic elements, we define our local error indicator as

$$\epsilon_\tau = \frac{1}{12} \prod_{k=1}^3 \left(\ell_{k+1} \partial_{k+1} - \ell_{k-1} \partial_{k-1} \right) \bar{u}_3 \phi_0 + \frac{1}{12} \sum_{k=1}^3 \ell_k^3 \partial_k^3 \bar{u}_3 \phi_k,$$

where $\bar{u}_3$ is any cubic polynomial with third derivatives equal to $\partial S_h^m Q_h \partial^2 u_h$, ℓ_k are the edge lengths of the triangular element, and ϕ_k's are hierarchical basis functions for the 4-dimensional space of cubic polynomials that are zero at the vertices and midpoints of the element. Note that the above local error indicator depends only on the geometry of corresponding element and the gradients of the recovered second order derivatives.

The rest of this paper is organized as follows: In Section 2, we describe our derivative scheme, and give superconvergence estimates of the p-th derivatives for shape regular meshes. In Section 3, we develop and analyze our a posteriori error estimate. Finally, in Section 4, we present several numerical examples, involving both uniform and adaptively refined (nonuniform) meshes, with some solutions that satisfy our smoothness assumptions and some that do not. In the latter case, we observe that high order elements require conservative refinement strategies to create meshes corresponding to optimal orders of convergence.

2. A derivative recovery scheme for shape regular triangulations. Let $\Omega \subset \mathbb{R}^2$ be a bounded domain with Lipschitz boundary $\partial \Omega$.[1] For simplicity of exposition, we assume that Ω is a polygon. Let $\mathcal{V}_h^{(p)}$ denote the finite element space consisting of C^0 piecewise polynomials of degree p associated with a shape regular triangulation $\mathcal{T}_h$, and $u_h \in \mathcal{V}_h^{(p)}$ be the finite element approximation to a (possibly nonlinear) second order elliptic boundary value problem.

We analyze a superconvergent approximation to the p-th order derivatives of u. This approximation is generated by applying the global L^2 projection operator Q_h and a multigrid smoothing operator S_h to the discrete p-th order derivatives of the finite element solution u_h, and it can be represented as $S_h^m Q_h \partial_h^p u_h$.

We first recall that the L^2 projection $Q_h u \in \mathcal{V}_h^{(1)}$ of a given function $u \in L^2(\Omega)$ is defined by solving the variational problem

$$(Q_h u, v_h) = (u, v_h), \ \forall \ v_h \in \mathcal{V}_h^{(1)}. \tag{2.1}$$

Here $(\cdot, \cdot)$ denotes the inner product on $L^2(\Omega)$.

---

[1] It is easy to see our theory in the paper is also valid for domains with cracks, such as the slit domain in the third example of Section 4.

Consider the following bilinear form:

$$a(u, v) = (\nabla u, \nabla v) + (u, v). \tag{2.2}$$

By the Riesz representation theorem, $a(\cdot, \cdot)$ induces a bounded linear operator $A_h :$ $\mathcal{V}_h^{(1)} \to \mathcal{V}_h^{(1)}$ uniquely determined by

$$(A_h u_h, v_h) = a(u_h, v_h), \ \forall \ u_h, v_h \in \mathcal{V}_h^{(1)},$$

it follows that the operator A_h is symmetric with respect to the L^2-inner product. We further notice that the discrete operator A_h is symmetric positive definite on the finite dimensional space $\mathcal{V}_h^{(1)}$ and

$$\lambda \equiv \rho(A_h) \simeq h^{-2}.$$

Using A_h, we introduce the smoothing operator S_h defined by

$$S_h = I - \lambda^{-1} A_h.$$

The usual multigrid convergence function

$$f(\alpha, \beta) = \frac{\alpha^\alpha \beta^\beta}{(\alpha + \beta)^{(\alpha+\beta)}} = \sup_{x \in [0,1]} x^\alpha (1 - x)^\beta,$$

$\alpha, \beta > 0$, plays an important role [3].

For convenience in notation, we let $\partial^p u$ denote some p-th order derivative of u and $\partial^p u_h$ denote some discrete p-th order derivative of u_h. We also use the notation $\| \cdot \|'_\Omega$ to indicate discrete norm $\sum_{K \in \mathcal{T}_h} \| \cdot \|_K$. We now state and prove some preliminary lemmas leading to the main Theorem 2.5 in this section.

LEMMA 2.1. *For any* $z \in \mathcal{V}_h^{(1)}$, $u \in H^{p+2}(\Omega)$,

$$\|(I - S_h^m)z\|_{0,\Omega} \lesssim mh(\|z - \partial^p u\|_{1,\Omega} + h\|u\|_{p+2,\Omega} + h^{1/2}|u|_{p+1,\partial\Omega}).$$

Proof. We note, from the definition of S_h,

$$\begin{aligned}
\|(I - S_h^m)z\|_{0,\Omega} &= \lambda^{-1} \|(I - S_h^m)(I - S_h)^{-1} A_h z\|_{0,\Omega} \\
&\leq \lambda^{-1} \max_{s \in [0,1]} [(1 - s^m)(1 - s)^{-1}] \|A_h z\|_{0,\Omega} \\
&\leq \lambda^{-1} m \|A_h z\|_{0,\Omega} \\
&\lesssim mh^2 \|A_h z\|_{0,\Omega}.
\end{aligned}$$

Let $w = A_h z$. By definition,

$$(w, \varphi) = (\nabla z, \nabla \varphi) + (z, \varphi) \tag{2.3}$$

for all $\varphi \in \mathcal{V}_h^{(1)}$. We take $\varphi = w$ in (2.3),

$$\|w\|_{0,\Omega}^2 = (w, w) = (\nabla z, \nabla w) + (z, w).$$

We estimate the terms on the right-hand side

$$(\nabla z, \nabla w) = (\nabla(z - \partial^p u), \nabla w) + (\nabla \partial^p u, \nabla w)$$
$$\lesssim \|\nabla(z - \partial^p u)\|_{0,\Omega}\|\nabla w\|_{0,\Omega} - (\triangle \partial^p u, w) + \int_{\partial\Omega} \nabla \partial^p u \cdot n w ds$$
$$\lesssim h^{-1}\|\nabla(z - \partial^p u)\|_{0,\Omega}\|w\|_{0,\Omega} + \|u\|_{p+2,\Omega}\|w\|_{0,\Omega} + |u|_{p+1,\partial\Omega}\|w\|_{0,\partial\Omega}$$
$$\lesssim (h^{-1}\|z - \partial^p u\|_{1,\Omega} + \|u\|_{p+2,\Omega} + h^{-1/2}|u|_{p+1,\partial\Omega})\|w\|_{0,\Omega}.$$

Also

$$(z, w) = (z - \partial^p u, w) + (\partial^p u, w)$$
$$\lesssim (\|z - \partial^p u\|_{0,\Omega} + \|u\|_{p,\Omega})\|w\|_{0,\Omega}.$$

Thus for $z \in \mathcal{V}_h^{(1)}$,

$$\|A_h z\|_{0,\Omega} = \|w\|_{0,\Omega} \lesssim h^{-1}\|z - \partial^p u\|_{1,\Omega} + \|u\|_{p+2,\Omega} + h^{-1/2}|u|_{p+1,\partial\Omega},$$

completing the proof. □

LEMMA 2.2. *[3] Suppose that for $v \in \mathcal{V}_h^{(1)}$ and some $0 < \alpha \le 1$ we have*

$$\|v\| \le \omega(h, v),$$
$$\|v\|_{-\alpha} \equiv \|A_h^{-\alpha/2}v\| \le (\mathcal{C}h)^\alpha \omega(h, v).$$

Then

$$\|S_h^m v\| \le \varepsilon_m \, \omega(h, v),$$

where

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

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

LEMMA 2.3. *Let $w|_K \in H^p(K) \cap W^{p-1,\infty}(K)$, for all $K \in \mathcal{T}_h$. Then, for $1/2 < \alpha \le 1$,*

$$\|S_h^m Q_h \partial^p w\|_{0,\Omega} \lesssim \varepsilon_m (h^{-1}\|w\|'_{p-1,\Omega} + \|w\|'_{p,\Omega} + h^{-\alpha}\|w\|'_{p-1,\infty,\Omega}),$$

with ε_m defined as in Lemma 2.2.

Proof. Our plan is to apply Lemma 2.2 to $v = Q_h \partial^p w$. Note that

$$\|v\|_{-\alpha} = \|Q_h \partial^p w\|_{-\alpha} = \sup_{\phi \in \mathcal{V}_h^{(1)}} \frac{(Q_h \partial^p w, \phi)}{\|\phi\|_\alpha} = \sup_{\phi \in \mathcal{V}_h^{(1)}} \frac{(\partial^p w, \phi)}{\|\phi\|_\alpha}.$$

Using integration by parts,

$$(\partial^p w, \phi) = -(\partial^{p-1} w, \partial\phi) + \sum_{K \in \mathcal{T}_h} \int_{\partial K} \partial^{p-1} w \phi n_i ds$$
$$\le \|w\|'_{p-1,\Omega}\|\phi\|_{1,\Omega} + \|w\|'_{p-1,\infty,\Omega}\|\phi\|_{\alpha,\Omega}$$
$$\lesssim (h^{\alpha-1}\|w\|'_{p-1,\Omega} + \|w\|'_{p-1,\infty,\Omega})\|\phi\|_{\alpha,\Omega}.$$

Thus

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

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

Since

$$\|v\|_{0,\Omega} \le \|\partial^p w\|'_{0,\Omega} \le \omega(h,v),$$

the desired estimate now follows from Lemma 2.2. $\square$

LEMMA 2.4. *Let* $u \in H^{p+2}(\Omega) \cap W^{p+1,\infty}(\Omega)$. *Then for any* $v_h \in \mathcal{V}_h^{(p)}$ *and* $1/2 < \alpha \le 1$ *we have*

$$\|\partial^p u - S_h^m Q_h \partial^p v_h\|_{0,\Omega} \lesssim mh^{3/2}(h^{1/2}\|u\|_{p+2,\Omega} + |u|_{p+1,\partial\Omega})$$
$$+ \varepsilon_m(h^{-1}\|u - v_h\|'_{p-1,\Omega} + h^{-\alpha}\|u - v_h\|'_{p-1,\infty,\Omega}),$$

with ε_m *defined as in Lemma* 2.2.

Proof. By the triangle inequality,

$$\|\partial^p u - S_h^m Q_h \partial^p v_h\|_{0,\Omega} \le \|(I - Q_h)\partial^p u\|_{0,\Omega} + \|(I - S_h^m)Q_h \partial^p u\|_{0,\Omega}$$
$$+ \|S_h^m Q_h \partial^p(u - v_h)\|_{0,\Omega}.$$

By standard arguments, the first term

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

The second term is estimated by Lemma 2.1. For the third term, we apply Lemma 2.3. $\square$

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

$$\|\partial^p u - S_h^m Q_h \partial^p u_h\|_{0,\Omega} \lesssim h(mh + \varepsilon_m)\|u\|_{p+2,\Omega}.$$

In the more general case, we have the following theorem based only on the results developed in this section.

THEOREM 2.5. *Let* $u \in H^{p+2}(\Omega) \cap W^{p+1,\infty}(\Omega)$ *and* $u_h \in \mathcal{V}_h^{(p)}$ *be an approximation of* u *satisfying*

$$\|u - u_h\|'_{p-1,\Omega} \lesssim h^2 |u|_{p+1,\Omega},$$

$$\|u - u_h\|'_{p-1,\infty,\Omega} \lesssim h^2 |\log h| |u|_{p+1,\infty,\Omega}.$$

Then

$$\|\partial^p u - S_h^m Q_h \partial^p u_h\|_{0,\Omega} \lesssim h(mh^{1/2} + \varepsilon_m)(\|u\|_{p+2,\Omega} + |u|_{p+1,\infty,\Omega}),$$

where ε_m *is defined as in Lemma* 2.2 *and* $1/2 < \alpha < 1$.

We can easily derive the following estimate for $(p+1)$-st order derivatives with the help of Theorem 2.5.

THEOREM 2.6. *Assume the hypotheses of Theorem* 2.5. *Then*

$$\|\partial(\partial^p u - S_h^m Q_h \partial^p u_h)\|_{0,\Omega} \lesssim (mh^{1/2} + \varepsilon_m)(\|u\|_{p+2,\Omega} + |u|_{p+1,\infty,\Omega}),$$

where ε_m is defined as in Lemma 2.2 and $1/2 < \alpha < 1$.

    Proof. Let $z = I_h \partial^p u \in \mathcal{V}_h^{(1)}$. Then

$$
\begin{aligned}
\|\partial(\partial^p u - S_h^m Q_h \partial^p u_h)\|_{0,\Omega} &\leq \|\partial(\partial^p u - z)\|_{0,\Omega} + \|\partial(z - S_h^m Q_h \partial^p u_h\|_{0,\Omega} \\
&\lesssim h\|u\|_{p+2,\Omega} + h^{-1}\|z - S_h^m Q_h \partial^p u_h\|_{0,\Omega} \\
&\lesssim h\|u\|_{p+2,\Omega} + h^{-1}(\|z - \partial^p u\|_{0,\Omega} + \|\partial^p u - S_h^m Q_h \partial^p u_h\|_{0,\Omega}) \\
&\lesssim (mh^{1/2} + \varepsilon_m)(\|u\|_{p+2,\Omega} + |u|_{p+1,\infty,\Omega}).
\end{aligned}
$$

$\square$

    3. A Posteriori Error Estimates. We begin with a description of our a posteriori error estimator. Our approach follows the development in [3] for the case of piecewise linear finite elements. For the general case of Lagrange elements of degree p, our goal is to find an expression for the error that involves only (approximate) derivatives of order $p + 1$ of u and known parameters describing the geometry of a given element τ. Let a canonical element $\tau \in \mathcal{T}_h$ have vertices $\boldsymbol{p}_k^t = (x_k, y_k)$, $1 \leq k \leq 3$, oriented counterclockwise, and corresponding linear 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).

    As an example, we now restrict attention to quadratic finite elements, since it is the quadratic space that is used in the numerical illustrations. We first seek an expression for $\hat{u}_3 - u_2$ on τ, where u_2 is the quadratic Lagrange interpolant and $\hat{u}_3$ is the cubic hierarchical extension. Thus $\hat{u}_3 - u_2$ is a cubic polynomial zero at vertices and edge midpoints of τ. A hierarchical basis for this 4-dimensional space is given by

$$
\begin{aligned}
\phi_0 &= \psi_1 \psi_2 \psi_3, \\
\phi_k &= \psi_{k-1} \psi_{k+1}(\psi_{k+1} - \psi_{k-1}),
\end{aligned}
$$

for $1 \leq k \leq 3$, and $(k-1, k, k+1)$ is a cyclic permutation of $(1,2,3)$. Let $\partial_k u$ denote the directional derivative in the direction $\boldsymbol{t}_k$. Then

$$
\hat{u}_3 - u_2 = \frac{1}{12} \prod_{k=1}^3 (\ell_{k+1}\partial_{k+1} - \ell_{k-1}\partial_{k-1})\,\hat{u}_3 \phi_0 + \frac{1}{12}\sum_{k=1}^3 \ell_k^3 \partial_k^3 \hat{u}_3 \phi_k. \qquad (3.1)
$$

Equation (3.1) can be verified using the identities

$$
\begin{aligned}
\psi_1 + \psi_2 + \psi_3 &= 1, \\
\nabla\psi_1 + \nabla\psi_2 + \nabla\psi_3 &= 0, \\
\ell_1 \boldsymbol{t}_1 + \ell_2 \boldsymbol{t}_2 + \ell_3 \boldsymbol{t}_3 &= 0, \\
\begin{pmatrix} \ell_1 \boldsymbol{t}_1^t \\ \ell_2 \boldsymbol{t}_2^t \\ \ell_3 \boldsymbol{t}_3^t \end{pmatrix} \begin{pmatrix} \nabla\psi_1 & \nabla\psi_2 & \nabla\psi_3 \end{pmatrix} &= \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}.
\end{aligned}
$$

In our local error indicator, we simply approximate the third derivatives needed
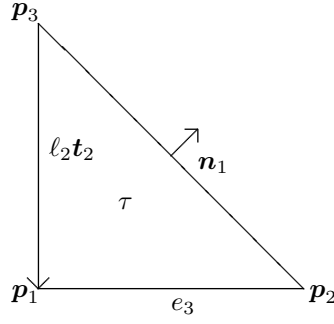
FIG. 3.1. *Parameters associated with the triangle τ.*

to compute the directional derivatives appearing in (3.1) by

$$
\begin{aligned}
\partial_{xxx}\hat{u}_3 &\approx \alpha_\tau \partial_x S_h^m Q_h \partial_{xx} u_h, \\
\partial_{xxy}\hat{u}_3 &\approx \frac{\alpha_\tau}{2}(\partial_y S_h^m Q_h \partial_{xx} u_h + \partial_x S_h^m Q_h \partial_{xy} u_h), \\
\partial_{xyy}\hat{u}_3 &\approx \frac{\alpha_\tau}{2}(\partial_y S_h^m Q_h \partial_{xy} u_h + \partial_x S_h^m Q_h \partial_{yy} u_h), \\
\partial_{yyy}\hat{u}_3 &\approx \alpha_\tau \partial_y S_h^m Q_h \partial_{yy} u_h,
\end{aligned}
\tag{3.2}
$$

where $\alpha_\tau > 0$ is a constant described below. Let $\bar{u}_3$ be any cubic polynomial with third derivatives given by the right-hand sides of (3.2). Then our local error indicator is given by

$$
\epsilon_\tau = \frac{1}{12}\prod_{k=1}^{3}(\ell_{k+1}\partial_{k+1} - \ell_{k-1}\partial_{k-1})\,\bar{u}_3\phi_0 + \frac{1}{12}\sum_{k=1}^{3}\ell_k^3\partial_k^3\bar{u}_3\phi_k.
\tag{3.3}
$$

The normalization constant α_τ is chosen such that

$$
\begin{aligned}
|\epsilon_\tau|_{2,\tau}^2 = \|(I - S_h^m Q_h)\partial_{xx}^2 u_h\|_{0,\tau}^2 &+ 2\|(I - S_h^m Q_h)\partial_{xy}^2 u_h\|_{0,\tau}^2 \\
&+ \|(I - S_h^m Q_h)\partial_{yy}^2 u_h\|_{0,\tau}^2 \equiv |u_h - R(u_h)|_{2,\tau}^2.
\end{aligned}
$$

Normally we expect that $\alpha_\tau \approx 1$, which is likely to be the case in regions where the third derivatives of the true solution are well defined. Near singularities, u is not smooth and we anticipate difficulties in estimating the third derivatives. For elements near such singularities, α_τ provides a heuristic for partly compensating for poor approximation. Note that ϵ_τ is a cubic polynomial on each element depending only on the geometry of τ and the approximate third derivatives derived from our superconvergent approximations.

In the general case, $\hat{u}_{p+1} - u_p$ on element τ is a polynomial of degree $p+1$ that is zero at the degrees of freedom defining u_p. One can express this polynomial in terms of hierarchical basis functions depending on the geometry of τ, and the derivatives of order $p+1$ of $\hat{u}_{p+1}$. The derivatives can be approximated by $\partial S_h^m Q_h \partial^p u_h$ as in the example above. This yields a polynomial ϵ_τ of degree $p+1$ for each element. The local error indicator η_τ is given by

$$
\eta_\tau = \|\nabla\epsilon_\tau\|_{0,\tau}.
\tag{3.4}
$$

Since ϵ_τ is a discontinuous piecewise polynomial on all of Ω, we can also formally approximate errors in global norms and other functionals using ϵ_τ. For example

$$\|u - u_h\|_{0,\Omega}^2 \approx \sum_\tau \|\epsilon_\tau\|_{0,\tau}^2,$$

$$|u - u_h|_{1,\Omega}^2 \approx \sum_\tau |\epsilon_\tau|_{1,\tau}^2 = \sum_\tau \eta_\tau^2.$$

In the case of $|\cdot|_{1,\Omega}$ there is a bit of theory. In particular

$$|u - u_h|_{1,\Omega} \le |u - \hat{u}_{p+1}|_{1,\Omega} + |u_p - u_h|_{1,\Omega} + |u_p - \hat{u}_{p+1}|_{1,\Omega}, \tag{3.5}$$

$$|u_p - \hat{u}_{p+1}|_{1,\Omega} \le |u - \hat{u}_{p+1}|_{1,\Omega} + |u_p - u_h|_{1,\Omega} + |u - u_h|_{1,\Omega}. \tag{3.6}$$

Suppose $|u - u_h|_{1,\Omega} \ge ch^p$, and $|u - \hat{u}_{p+1}|_{1,\Omega} \le Ch^{p+1}$. Then if $|u_p - u_h|_{1,\Omega}$ is also higher order, estimates (3.5)–(3.6) show $|u_p - \hat{u}_{p+1}|_{1,\Omega}$ to be an asymptotically exact estimate for $|u - u_h|_{1,\Omega}$. Such super-approximation results for $|u_p - u_h|_{1,\Omega}$ are known for $p = 1, 2$, see [2, 3, 8]. However, super-approximation estimates for $|u_p - u_h|_{1,\Omega}$ are known *not* to hold for $p \ge 3$, see [9]. For general p, estimate (3.5) can be replaced by

$$|u - u_h|_{1,\Omega} \le C\left(|u - \hat{u}_{p+1}|_{1,\Omega} + |\hat{u}_{p+1} - u_p|_{1,\Omega}\right) \tag{3.7}$$

due to the best approximation property for the energy norm and norm comparability of $\|\cdot\|_\Omega$ and $|\cdot|_{1,\Omega}$. Here we lose asymptotic exactness, but still have a useful upper bound for the error. Insofar as we know, the lower bound for $p > 2$ is still an open question, as are general norms and functionals. Nonetheless, this informal analysis suggests that the error indicators η_τ will provide a useful and reliable basis for adaptive meshing algorithms.

4. Numerical experiments. We now present some numerical illustrations of our recovery scheme in the cases of uniform and adaptively refined (nonuniform) meshes. The gradient recovery scheme and a posteriori error estimate described above for the case of continuous piecewise quadratic elements were implemented in the PLTMG package [1], which was then used for our numerical experiments. The experiments were done on a dual Opteron Linux workstation, using the *g77* compiler and double precision arithmetic. We reprise some experiments given in [3] for the case of continuous piecewise linear finite elements.

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

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

where f and g are chosen such that $u = \mathrm{e}^{x+y}$ 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.

In Tables 4.1–4.2, 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
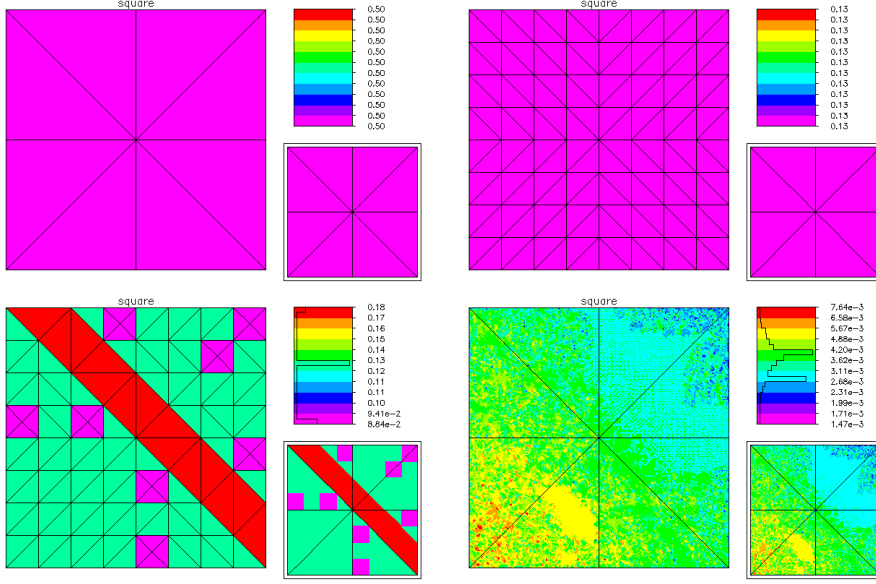
FIG. 4.1. *Top left: 3×3 initial mesh. Top right: uniform refinement with $nt = 128$. Bottom left: adaptive refinement with $nt = 137$. Bottom right: adaptive refinement with $nt = 131105$. Elements are colored according to size.*

procedure to produce adaptive meshes with similar numbers of elements to the uniform refinement case. Note that the dimension of the quadratic finite element space is approximately $2\,nt$, where nt is the number of elements reported in the tables. Other values are defined as follows:

$$L2 = \|u - u_h\|_{0,\Omega},$$

$$\widetilde{L2} = \|\epsilon_h\|_{0,\Omega},$$

$$EF_0 = \frac{\|\epsilon_h\|_{0,\Omega}}{\|u - u_h\|_{0,\Omega}},$$

$$H1 = |u - u_h|_{1,\Omega},$$

$$\widetilde{H1} = |\epsilon_h|_{1,\Omega},$$

$$EF_1 = \frac{|\epsilon_h|_{1,\Omega}}{|u - u_h|_{1,\Omega}},$$

$$H2 = |u - u_h|_{2,\Omega},$$

$$\widetilde{H2} = |u - R(u_h)|_{2,\Omega},$$

$$EF_2 = \frac{|R(u_h) - u_h|_{2,\Omega}}{|u - u_h|_{2,\Omega}}.$$

For each type of norm, 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.

We note here the superconvergence of the second derivatives and effectivity ratios that are close to one. Despite lack of a complete theory, error estimates $\widetilde{L2}$ and $\widetilde{H1}$ are also quite accurate, and the orders of convergence are optimal in all three norms

TABLE 4.1
Error estimates for uniform refinement.

nt	$L2$	$\widetilde{L2}$	EF_0	$H1$	$\widetilde{H1}$	EF_1	$H2$	$\widetilde{H2}$	EF_2
8	8.8e-3	1.0e-2	1.1	0.1	0.2	1.6	1.3	2.1	1.7
32	1.0e-3	1.8e-3	1.8	3.0e-2	0.1	1.8	0.7	1.2	2.0
128	1.2e-4	2.0e-4	1.6	7.5e-3	1.2e-2	1.6	0.3	0.5	1.7
512	1.6e-5	2.4e-5	1.6	1.9e-3	2.9e-3	1.5	0.2	0.2	1.5
2048	1.9e-6	2.7e-6	1.4	4.7e-4	6.5e-4	1.4	0.1	0.1	1.4
8192	2.4e-7	3.1e-7	1.3	1.2e-4	1.5e-4	1.3	4.2e-2	3.4e-2	1.3
32768	3.0e-8	3.5e-8	1.2	3.0e-5	3.4e-5	1.2	2.1e-2	1.3e-2	1.2
131072	3.8e-9	4.1e-9	1.1	7.4e-6	8.0e-6	1.1	1.0e-2	4.7e-3	1.1
order	3.04	3.15		2.02	2.13		1.01	1.43	

TABLE 4.2
Error estimates for adaptive refinement.

nt	$L2$	$\widetilde{L2}$	EF_0	$H1$	$\widetilde{H1}$	EF_1	$H2$	$\widetilde{H2}$	EF_2
8	6.9e-4	3.7e-4	0.5	1.0e-2	5.6e-3	0.6	0.2	0.2	0.5
33	2.5e-4	1.8e-4	0.7	5.1e-3	4.8e-3	0.9	0.1	0.2	1.0
137	1.6e-5	2.2e-5	1.4	8.9e-4	1.5e-3	1.6	0.1	0.1	1.7
523	1.8e-6	2.2e-6	1.2	1.8e-4	2.6e-4	1.4	2.2e-2	3.1e-2	1.6
2063	2.0e-7	2.0e-7	1.0	3.7e-5	4.4e-5	1.2	1.0e-2	1.0e-2	1.3
8207	1.8e-8	1.6e-8	0.9	7.9e-6	7.9e-6	1.0	4.7e-3	2.6e-3	1.1
32775	2.2e-9	1.7e-9	0.8	1.9e-6	1.7e-6	0.9	2.3e-3	7.3e-4	1.0
131105	2.6e-1	2.0e-1	0.8	4.5e-7	4.1e-7	0.9	1.1e-3	2.1e-4	1.0
order	3.15	3.24		2.12	2.20		1.06	1.83	

(and superconvergent for the recovered second derivatives).

In our second example, we consider the nonlinear problem

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

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 Tables 4.3–4.4.

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 larger derivatives than the first example. Nonetheless, we see a similar behavior of the gradient recovery scheme and a posteriori error estimate.
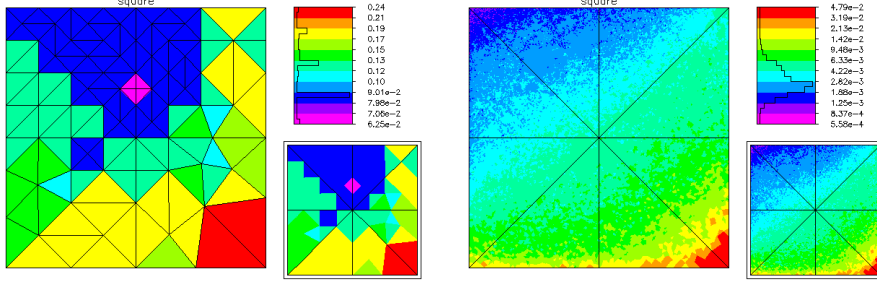
FIG. 4.2. *Left: adaptive refinement with nt = 135. Right: adaptive refinement with nt = 129345. Elements are colored according to size.*

TABLE 4.3
Error estimates for uniform refinement.

nt	L2	$\widetilde{L2}$	EF_0	H1	$\widetilde{H1}$	EF_1	H2	$\widetilde{H2}$	EF_2
8	6.9e-4	3.7e-4	0.5	1.0e-2	5.6e-3	0.6	0.2	0.2	0.5
32	2.2e-4	1.3e-4	0.6	4.7e-3	3.9e-3	0.8	0.1	0.2	1.0
128	4.3e-5	3.1e-5	0.7	2.0e-3	2.0e-3	1.0	0.1	0.1	1.4
512	6.1e-6	5.3e-6	0.9	6.2e-4	7.0e-4	1.1	4.4e-2	0.1	1.5
2048	6.7e-7	7.4e-7	1.1	1.5e-4	2.0e-4	1.3	2.1e-2	3.7e-2	1.6
8192	6.4e-8	8.9e-8	1.4	3.0e-5	4.7e-5	1.6	1.0e-2	1.6e-2	1.7
32768	6.4e-9	1.0e-8	1.5	6.4e-6	1.0e-5	1.6	4.5e-3	6.3e-3	1.6
131072	7.1e-1	1.0e-9	1.5	1.5e-6	2.2e-6	1.5	2.2e-3	2.4e-3	1.4
order	3.26	3.21		2.18	2.20		1.08	1.35	

In our third example, we consider the problem

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

where Ω is a circle of radius one centered at the origin, and with a crack along the positive x-axis $0 \leq x \leq 1$. The boundary $\partial\Omega_2$ is the bottom edge of the crack, and $\partial\Omega_1 = \partial\Omega \setminus \partial\Omega_2$. 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 Tables 4.5–4.6. The solution u is not smooth in this case ($u \in H^{5/4-\epsilon}(\Omega)$), and this is reflected in the results. In particular, both $H2$ and $\widetilde{H2}$ should be infinite and are computed as finite only because of numerical quadrature. Despite this singularity, the resulting error approximations ϵ_τ still provided useful information and formed a reliable basis for adaptive refinement.

For the case of uniform refinement, the 0.25 order of convergence of the gradient coincides with the smoothness of the solution. The effectivity ratios for all cases show lack of asymptotic exactness. EF_1, although not approaching unity, still seems under control, reflecting the analysis (3.5)–(3.7). On the other hand, EF_0 seems very poorly controlled, indicating that $\|\epsilon_\tau\|_{0,\Omega}$ is not a very reliable estimate for $\|u - u_h\|_{0,\Omega}$.

TABLE 4.4
Error estimates for adaptive refinement.

| nt	$L2$	$\widetilde{L2}$	EF_0	$H1$	$\widetilde{H1}$	EF_1	$H2$	$\widetilde{H2}$	EF_2
8	6.9e-4	2.0e-4	0.3	1.0e-2	3.0e-3	0.3	0.2	0.2	0.5
32	2.2e-4	8.2e-5	0.4	4.7e-3	2.5e-3	0.5	0.1	0.2	1.0
135	1.5e-5	1.9e-5	1.2	8.6e-4	1.5e-3	1.7	0.1	0.1	1.8
524	2.0e-6	2.8e-6	1.4	1.7e-4	5.2e-4	3.0	2.1e-2	3.0e-2	1.6
2060	3.4e-7	4.1e-7	1.2	4.5e-5	1.4e-4	3.0	1.0e-2	9.1e-3	1.3
8119	4.6e-8	4.8e-8	1.1	1.2e-5	3.2e-5	2.6	5.2e-3	2.2e-3	1.1
32333	5.4e-9	5.2e-9	1.0	2.9e-6	6.4e-6	2.2	2.7e-3	4.9e-4	1.0
129345	5.8e-1	5.4e-1	0.9	7.0e-7	1.3e-6	1.8	1.3e-3	1.5e-4	1.0
order	3.16	3.25		2.07	2.30		0.99	1.86	
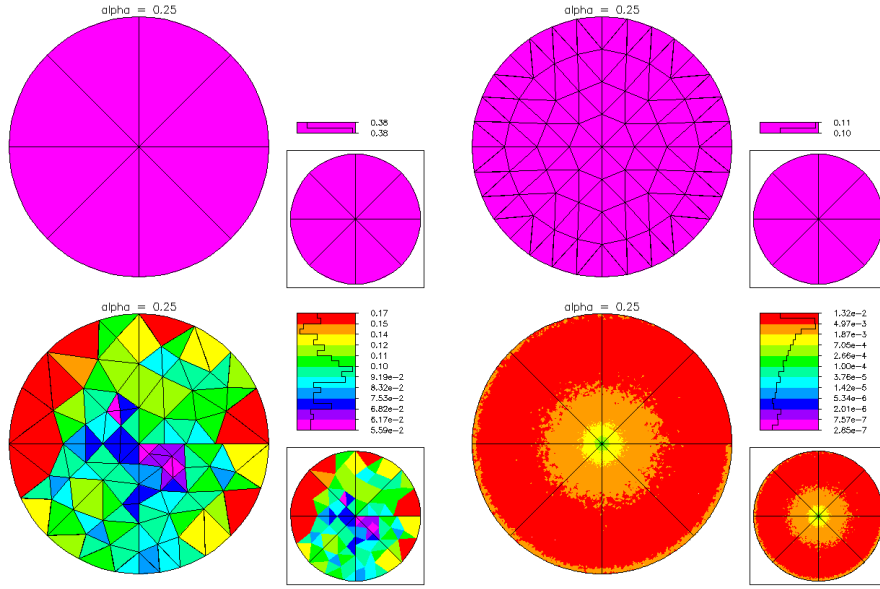
FIG. 4.3. *Top left: initial mesh with $nt = 8$. Top right: uniform refinement with $nt = 128$. Bottom left: adaptive refinement with $nt = 133$. Bottom right: adaptive refinement with $nt = 133890$. Elements are colored according to size.*

TABLE 4.5
Error estimates for uniform refinement.

nt	$L2$	$\widetilde{L2}$	EF_0	$H1$	$\widetilde{H1}$	EF_1	$H2$	$\widetilde{H2}$	EF_2
8	0.1	3.9e-2	0.3	0.6	0.3	0.5	4.1	4.3	0.4
32	0.1	8.6e-3	0.1	0.5	0.1	0.3	6.6	6.7	0.5
128	0.1	2.9e-3	4.9e-2	0.4	0.1	0.3	11.0	11.0	0.5
512	4.0e-2	1.2e-3	2.9e-2	0.3	0.1	0.3	18.0	19.0	0.5
2048	2.7e-2	4.7e-4	1.7e-2	0.2	0.1	0.3	31.0	32.0	0.5
8192	1.9e-2	2.0e-4	1.0e-2	0.2	0.1	0.3	52.0	53.0	0.5
32768	1.3e-2	8.1e-5	6.1e-3	0.2	4.6e-2	0.3	87.0	90.0	0.5
131072	9.3e-3	3.4e-5	3.7e-3	0.1	3.8e-2	0.3	1.5e2	1.5e2	0.5
order	0.53	1.29		0.27	0.27		-0.76	-0.76	

TABLE 4.6
Error estimates for adaptive refinement.

nt	$L2$	$\widetilde{L2}$	EF_0	$H1$	$\widetilde{H1}$	EF_1	$H2$	$\widetilde{H2}$	EF_2
8	0.1	3.9e-2	0.3	0.6	0.3	0.5	4.1	4.3	0.4
31	0.1	8.9e-3	0.1	0.5	0.1	0.3	6.4	6.5	0.5
133	4.5e-2	2.3e-3	5.0e-2	0.3	0.1	0.3	14.0	14.0	0.5
533	2.3e-2	4.4e-4	1.9e-2	0.2	0.1	0.3	34.0	35.0	0.5
2078	8.5e-3	5.7e-5	6.7e-3	0.1	0.1	0.4	1.5e2	1.6e2	0.5
8237	2.3e-3	4.6e-6	2.0e-3	0.1	2.7e-2	0.4	1.0e3	1.1e3	0.5
32796	6.0e-4	5.8e-7	1.0e-3	3.5e-2	1.2e-2	0.3	7.7e3	8.1e3	0.5
130890	1.1e-4	7.4e-8	6.6e-4	1.5e-2	7.6e-3	0.5	9.8e4	1.1e5	0.6
order	2.18	3.07		1.12	0.83		-3.29	-3.30	

For the adaptive meshes, the order of convergence improves and seems to be approaching order one for the gradient and order two for the solution. This is suboptimal for quadratic elements. This was not due to poor error indicators, but rather to an overly aggressive refinement strategy. In the adaptive refinement procedure implemented in PLTMG, all elements are placed on a heap, with the element having largest error at the root. The root element is then selected for refinement. When an element is refined, it is removed from the heap, and its child elements are added to the heap. This of course requires the child elements to have error indicators. These are constructed using derivative values inherited from the parent, and their own geometry information. Thus a single element might undergo several levels of refinement during a given adaptive step. Using old derivative information for new elements will generally fail to be optimal after sufficiently many levels of refinement.

In PLTMG, the amount of refinement allowed in a given refinement step is governed by the user, by specifying a target number of vertices in the refined mesh. In this example, we chose a strategy that increased the number of vertices by roughly a factor of four in each refinement step, in order to closely match the size of problems generated by uniform refinement. If there are too many levels of refinement of individual elements before the problem is resolved, the resulting mesh might be lower quality, as was the case in this example. On the other hand, frequently assembling and resolving the global finite element equations results in higher quality adaptive meshes, but at a much greater cost. Since the appropriate compromise is likely to be highly problem dependent, in PLTMG it is up to the user to choose the proper balance.

For this example, we solved this problem adaptively a second time, this time specifying the number of vertices should be increased by a factor of roughly two between resolves, rather than four. Here we see near optimal rate of convergence for both $L2$ and $H1$. In other respects, the data are quite similar to Table 4.6. Meshes corresponding to the adaptive meshes in Figure 4.3 are shown in Figure 4.4.

It is interesting to note that a refinement factor of four caused no problems in the case of a similar experiment performed in the case $p = 1$ in [3]. For $p = 2$ the refinement is much sharper in the region of the singularity, and it was this increase in sharpness that required a less aggressive refinement strategy. Even higher order elements are likely to require even more conservative refinement strategies to create meshes corresponding to optimal orders of convergence. Perhaps this adds another dimension to already complex general discussions evaluating the relative merits of
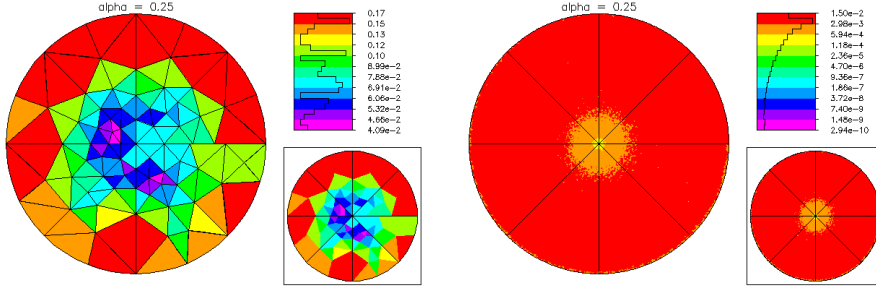
FIG. 4.4. *Left: adaptive refinement with nt = 136. Right: adaptive refinement with nt = 130586. Elements are colored according to size.*

TABLE 4.7
Error estimates for adaptive refinement.

nt	$L2$	$\widetilde{L2}$	EF_0	$H1$	$\widetilde{H1}$	EF_1	$H2$	$\widetilde{H2}$	EF_2
8	0.1	3.9e-2	0.3	0.6	0.3	0.5	4.1	4.3	0.4
31	0.1	8.9e-3	0.1	0.5	0.1	0.3	6.4	6.5	0.5
74	0.1	4.9e-3	0.1	0.4	0.1	0.4	9.5	9.8	0.5
136	4.5e-2	1.8e-3	4.1e-2	0.3	0.1	0.3	14.0	15.0	0.5
302	2.5e-2	5.6e-4	2.2e-2	0.3	0.1	0.3	34.0	35.0	0.4
534	1.3e-2	2.7e-4	2.1e-2	0.2	0.1	0.4	87.0	89.0	0.4
1179	5.8e-3	7.8e-5	1.4e-2	0.1	4.1e-2	0.3	2.5e2	2.6e2	0.5
2077	2.8e-3	3.5e-5	1.3e-2	0.1	2.7e-2	0.3	7.0e2	7.3e2	0.5
4617	1.1e-3	1.2e-5	1.1e-2	5.0e-2	1.9e-2	0.4	3.1e3	3.3e3	0.5
8204	3.7e-4	5.1e-6	1.4e-2	3.0e-2	1.1e-2	0.4	1.4e4	1.4e4	0.5
18388	1.4e-4	1.5e-6	1.1e-2	1.8e-2	6.2e-3	0.4	6.4e4	6.7e4	0.5
32669	4.0e-5	6.7e-7	1.7e-2	9.5e-3	4.2e-3	0.4	3.8e5	4.0e5	0.6
73439	1.1e-5	2.0e-7	1.8e-2	5.0e-3	1.9e-3	0.4	2.6e6	2.7e6	0.5
130586	2.6e-6	8.5e-8	3.2e-2	2.5e-3	1.0e-3	0.4	2.0e7	2.2e7	0.5
order	3.56	2.96		1.77	1.73		-5.21	-5.22	

higher order methods.

REFERENCES

[1] R. E. Bank, *PLTMG: A software package for solving elliptic partial differential equations, users' guide 9.0*, Technical report, Department of Mathematics, University of California at San Diego, 2004.

[2] R.E. Bank and J. Xu, *Asymptotic exact a posteriori error estimators, Part I: Grids with super-convergence*, SIAM J. Numer. Anal., 41 (2003), pp. 2294–2312.

[3] R.E. Bank and J. Xu, *Asymptotic exact a posteriori error estimators, Part II: General Unstructured Grids*, SIAM J. Numer. Anal., 41 (2003), pp. 2313–2332.

[4] J.H. Bramble, J.A. Nitsche, and A.H. Schatz, *Maximum-norm interior estimates for Ritz-Galerkin methods*, Math. Comp., 29 (1975), pp. 677–688.

[5] 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.

[6] E. Hinton and J.S. Campbell, *Local and global smoothing of discontinuous finite element functions using a least square method*, Internat. J. Numer. Methods Engrg., 8 (1974), pp. 461–480.

[7] I. Hlaváček, M. Křížek, V. Pištora, *How to recover the gradient of linear elements on nonuniform triangulations*, Appl. Math. 41 (1996), pp. 241–267.

[8] Y.Q. Huang and J. Xu, *Superconvergence for quadratic triangular finite elements on mildly structured grids*, Preprint, 2005.

[9] B. Li, *Lagrange interpolation and finite element superconvergence*, Numer. Methods. Partial Differ. Equations, 20 (2004), pp. 33–59.

[10] J.T. Oden and H.J. Brauchli, *On the calculation of consistent stress distributions in finite element applications*, Internat. J. Numer. Methods Engrg., 3 (1971), pp. 317–325.

[11] J.S. Ovall, *Asymptotically exact functional error estimators based on superconvergent gradient recovery*, Nmuer. Math., 102 (2006), pp. 543–558.

[12] N.-E. Wiberg and F. Abdulwahab, *Patch recovery based on superconvergent derivatives and equilibrium*, Internat. J. Numer. Methods Engrg., 36 (1993), pp. 2703–2724.

[13] N.-E. Wiberg, F. Abdulwahab, and S. Ziukas, *Enhanced superconvergence patch recovery incoorporating equilibrium and boundary conditions*, Internat. J. Numer. Methods Engrg., 37 (1994), pp. 3417–3440.

[14] J. Xu and Z. Zhang, *Analysis of recovery type a posteriori error estimators for mildly structured grids*, Math. Comp., 73 (2003), pp. 1139–1152.

[15] 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.

[16] Z. Zhang and A. Naga, *A new finite element gradient recovery method: superconvergence property*, SIAM J. Sci. Comput., 26 (2005), pp. 1192–1213.

[17] O.C. Zienkiewicz and J.Z. Zhu, *The superconvergence patch recovery and a posteriori error estimates part 1: the recovery technique*, Internat. J. Numer. Methods Engrg, 33 (1992), pp. 1331–1364.

[18] O.C. Zienkiewicz and J.Z. Zhu, *The superconvergence patch recovery and a posteriori error estimates part 2: error estimates and adaptivity*, Internat. J. Numer. Methods Engrg., 33 (1992), pp. 1365–1382.