Portland State University

From the SelectedWorks of Jeffrey S. Ovall

2007

FUNCTION, GRADIENT, AND HESSIAN RECOVERY USING QUADRATIC EDGE-BUMP FUNCTIONS

Jeffrey S. Ovall, *Portland State University*

Available at: https://works.bepress.com/jeffrey_ovall/23/

SIAM J. NUMER. ANAL.
Vol. 45, No. 3, pp. 1064–1080
© 2007 Society for Industrial and Applied Mathematics

FUNCTION, GRADIENT, AND HESSIAN RECOVERY USING QUADRATIC EDGE-BUMP FUNCTIONS*

JEFFREY S. OVALL†

Abstract. An approximate error function for the discretization error on a given mesh is obtained by projecting (via the energy inner product) the functional residual onto the space of continuous, piecewise quadratic functions which vanish on the vertices of the mesh. Conditions are given under which one can expect this hierarchical basis error estimator to give efficient and reliable function recovery, asymptotically exact gradient recovery, and convergent Hessian recovery in the square norms. One does not find similar function recovery results in the literature. The analysis given here is based on a certain superconvergence result which has been used elsewhere in the analysis of gradient recovery methods. Numerical experiments are provided which demonstrate the effectivity of the approximate error function in practice.

Key words. finite elements, a posteriori estimates, hierarchical bases, superconvergence, gradient recovery

AMS subject classifications. 65N15, 65N30, 65N50

DOI. 10.1137/060648908

1. Introduction.

Hierarchical basis a posteriori error estimators were introduced in the early 1980s [22], and a general framework for the analysis of their effectivity and computational cost has been given by Bank [5] and Bank and Smith [1]. The basic idea behind such methods is that the base space of functions V_h, in which we wish to find our finite element approximation u_h, is augmented by a complementary space $\tilde{V}_h$ such that the composite space $V_h \oplus \tilde{V}_h$ provides an improved finite element approximation $\bar{u}_h$. In symbols, we show this as $|\!|\!| u - \bar{u}_h |\!|\!| \le \beta |\!|\!| u - u_h |\!|\!|$ for some $\beta \in [0, 1)$, where $|\!|\!| \cdot |\!|\!|$ is the energy norm associated with the underlying bilinear form. This improved approximation assumption is referred to as a saturation assumption. An approximate error function $\varepsilon_h \approx u - u_h$ is computed in the space $\tilde{V}_h$. Using the saturation assumption and strengthened Cauchy inequalities between the spaces V_h and $\tilde{V}_h$, effectivity estimates of the form

$$c_1 \le \frac{|\!|\!| \varepsilon_h |\!|\!|}{|\!|\!| u - u_h |\!|\!|} \le c_2 \tag{1}$$

are proved.

In this paper a different sort of analysis, which yields stronger assertions, is given for the case where V_h is the space of continuous, piecewise linear functions on a given mesh and $\bar{V}_h$ is the space of continuous, piecewise quadratic functions on that same mesh. The augmenting space $\tilde{V}_h$ consists of quadratic "bump" functions which vanish on the vertices of the mesh. In particular, we show that the approximate error function, $\varepsilon_h \approx u - u_h$, provides efficient and reliable function recovery, asymptotically exact gradient recovery, and convergent Hessian recovery:

$$c_1 \le \frac{\|\varepsilon_h\|_{0,\Omega}}{\|u - u_h\|_{0,\Omega}} \le c_2, \qquad \frac{\|\varepsilon_h\|_{1,\Omega}}{\|u - u_h\|_{1,\Omega}} \to 1, \qquad \sum_{\tau \in \mathcal{T}_h} |\varepsilon_h|^2_{2,\tau} \to |u|^2_{2,\Omega}. \tag{2}$$

*Received by the editors January 3, 2006; accepted for publication (in revised form) January 8, 2007; published electronically May 7, 2007.
http://www.siam.org/journals/sinum/45-3/64890.html

†Max Planck Institute for Mathematics in the Sciences, Inselstraße 22-26, D-04103 Leipzig, Germany (ovall@mis.mpg.de).

Our analysis is based on a superconvergence result of Bank and Xu [3, 4], which also appears in a slightly more general form in [20]. This result was used in these papers to explain the success of a number of popular gradient recovery methods, but we use it here in the context of hierarchical basis error estimation to establish our key approximation results (2).

The rest of this paper is organized as follows. In section 2 we lay out the basic notation and assumptions for this paper. Section 3 contains a statement of the superconvergence result of Bank and Xu, which we then use to prove the above mentioned gradient and Hessian recovery results. In section 4 we prove the function recovery result and show why we cannot generally hope for asymptotic exactness in this case. Section 5 comprises almost half of the paper and consists of four examples, which are used to verify the effectivity of our estimator in practice, and a brief subsection on computational cost.

2. Notation and basic assumptions. Let $\Omega \subset \mathbb{R}^2$ be a bounded domain with Lipschitz boundary $\partial\Omega = \partial\Omega_D \cup \partial\Omega_N$, and define

$$\mathcal{H} \equiv \left\{ v \in H^1(\Omega) : v|_{\partial\Omega_D} = 0 \text{ in the trace sense} \right\}. \tag{3}$$

The usual spaces $W_p^k(\Omega)$ and $H^k(\Omega) \equiv W_2^k(\Omega)$ are equipped with their standard norms $\|\cdot\|_{k,p,\Omega}$ and $\|\cdot\|_{k,\Omega} \equiv \|\cdot\|_{k,2,\Omega}$, and seminorms $|\cdot|_{k,p,\Omega}$ and $|\cdot|_{k,\Omega}$, respectively. For simplicity in exposition, we will assume that $\partial\Omega$ is a polygon. Let data functions $a : \bar{\Omega} \to \mathbb{R}^{2\times 2}$, $\mathbf{b} : \bar{\Omega} \to \mathbb{R}^2$, $c, f : \bar{\Omega} \to \mathbb{R}$, and $g : \partial\Omega_N \to \mathbb{R}$ be given. The problem is to find $u \in \mathcal{H}$ such that

$$B(u, v) = F(v) \text{ for all } v \in \mathcal{H}, \tag{4}$$

$$B(u, v) \equiv \int_\Omega a\nabla u \cdot \nabla v + (\mathbf{b} \cdot \nabla u + cu)v \, dx, \tag{5}$$

$$F(v) \equiv \int_\Omega fv \, dx + \int_{\partial\Omega_N} gv \, ds. \tag{6}$$

We will assume that the data functions are sufficiently smooth, and that the matrix a is positive definite, with the smallest eigenvalue bounded below on Ω by some constant $\gamma > 0$. We make the following standard assumptions concerning the bilinear form B and linear functional F: There exist constants $\alpha, \nu, \mu > 0$, such that, for all $v, w \in \mathcal{H}$,

$$\begin{aligned} |F(v)| &\le \alpha \|v\|_{1,\Omega}, \\ |B(v, w)| &\le \nu \|v\|_{1,\Omega} \|w\|_{1,\Omega}, \\ B(v, v) &\ge \mu \|v\|_{1,\Omega}^2. \end{aligned}$$

Let $\mathcal{T}_h$ denote a shape-regular triangulation of Ω with mesh size $h \in (0, 1)$. Let $V_h \subset \mathcal{H}$ denote the space of continuous, piecewise-linear polynomials defined on $\mathcal{T}_h$, and let $\bar{V}_h \subset \mathcal{H}$ denote the continuous, piecewise-quadratic polynomials. We will think of $\bar{V}_h$ hierarchically as

$$\bar{V}_h = V_h \oplus \tilde{V}_h, \tag{7}$$

where $\tilde{V}_h$ is the space of quadratic "bump" functions, i.e., continuous piecewise-quadratic polynomials which vanish at all of the vertices of the triangulation. In what follows, $u_h \in V_h$ and $\bar{u}_h \in \bar{V}_h$ denote, respectively, the piecewise-linear and piecewise-quadratic approximate solutions of (4):

$$B(u_h, v) = F(v) \text{ for all } v \in V_h, \tag{8}$$

$$B(\bar{u}_h, v) = F(v) \text{ for all } v \in \bar{V}_h. \tag{9}$$

Let $u_\ell \in V_h$ and $u_q \in \bar{V}_h$ denote piecewise-linear and piecewise-quadratic interpolants of u on $\mathcal{T}_h$. We make the following standard assumptions about their asymptotic approximation quality:

$$\|u - u_\ell\|_{k,\Omega} \lesssim h^{2-k}\|u\|_{2,\Omega}, \tag{10}$$

$$\|u - u_q\|_{k,\Omega} \lesssim h^{3-k}\|u\|_{3,\Omega}, \tag{11}$$

for $0 \le k \le 1$.

3. Gradient and Hessian recovery. In this section we prove asymptotically exact gradient recovery and convergent Hessian recovery results,

$$\frac{\|\varepsilon_h\|_{1,\Omega}}{\|u - u_h\|_{1,\Omega}} \to 1, \qquad \sum_{\tau \in \mathcal{T}_h} |\varepsilon_h|_{2,\tau}^2 \to |u|_{2,\Omega}^2 \tag{12}$$

for the approximate error function $\varepsilon_h \approx u - u_h$ described below. We first describe the key assumption on the mesh that will play a role in these results. This mesh condition and a slight generalization of it can be found in [3, 20].

Let e denote an interior edge in $\mathcal{T}_h$ with adjacent triangles τ and τ'. We say that the quadrilateral formed by τ and τ' satisfies the *approximate $O(h^2)$-parallelogram property* provided that the lengths of opposite edges differ by only $O(h^2)$. The equivalent property at the boundary is as follows: Let e and e' denote adjacent boundary edges sharing the vertex x, and let τ and τ' be the triangles having the edges e and e', respectively. Let $\mathbf{t}$ and $\mathbf{t}'$ be the unit tangent vectors, corresponding to a counterclockwise orientation on τ and τ'. Starting with e for τ and e' for τ' we identify corresponding edges of τ and τ' by traversing their edges counterclockwise. We say that the triangles τ and τ' associated with the boundary vertex x satisfy the approximate $O(h^2)$-parallelogram property, provided that the lengths of corresponding edges in τ and τ' differ by only $O(h^2)$, and $|\mathbf{t} - \mathbf{t}'| = O(h)$. The key assumption on the triangulation follows.

ASSUMPTION 3.1 (an $O(h^{2\sigma})$-irregular triangulation).

1. *Let $\mathcal{E} = \mathcal{E}_1 \oplus \mathcal{E}_2$ denote the set of interior edges in $\mathcal{T}_h$. For each $e \in \mathcal{E}_1$, τ and τ' satisfy the approximate $\mathcal{O}(h^2)$-parallelogram property, while*

$$\sum_{e \in \mathcal{E}_2} |\tau| + |\tau'| = \mathcal{O}(h^{2\sigma}).$$

2. *Let $\mathcal{P} = \mathcal{P}_1 \oplus \mathcal{P}_2$ denote the set of boundary vertices. The elements associated with $x \in \mathcal{P}_1$ satisfy the approximate $\mathcal{O}(h^2)$-parallelogram property, and $|\mathcal{P}_2| = \kappa$, where κ is fixed independent of h.*

The second condition is necessary only in the case of Neumann boundary conditions, $\partial\Omega_N \neq \emptyset$. The following result, due to Bank and Xu [3], is the key lemma for the results in this paper.

LEMMA 3.2. *Under Assumption* 3.1*, we have*

$$\|u_h - u_\ell\|_{1,\Omega} \lesssim h^{1+\min(\sigma,1)}|\log h|^{1/2}\|u\|_{3,\infty,\Omega}. \tag{13}$$

We now present a new result based on Lemma 3.2 for computing a superconvergent approximation of the gradient. Suppose that we first solve for the linear finite element approximation, $u_h \in V_h$, and then augment this approximation by solving the residual equation on $\tilde{V}_h$, the space of quadratic bumps. In other words,

$$B(u_h, v) = F(v) \text{ for all } v \in V_h, \tag{14}$$

$$B(\varepsilon_h, v) = F(v) - B(u_h, v) \text{ for all } v \in \tilde{V}_h. \tag{15}$$

One can think of this as a projection of the residual error onto the space $\tilde{V}_h$. We have the following result.

THEOREM 3.3. *Under Assumption* 3.1*, we have*

$$\|u - (u_h + \varepsilon_h)\|_{1,\Omega} \lesssim h^{1+\min(\sigma,1)} |\log h|^{1/2} \|u\|_{3,\infty,\Omega}. \tag{16}$$

Proof. Using Galerkin orthogonality to replace $\varepsilon_h \in \tilde{V}_h$ with $u_b \in \tilde{V}_h$, the "bump" portion of the quadratic interpolant $u_q = u_\ell + u_b$, we get the following estimate:

$$\|u - (u_h + \varepsilon_h)\|_{1,\Omega}^2 \lesssim B(u - (u_h + \varepsilon_h), u - (u_h + \varepsilon_h)) \tag{17}$$

$$= B(u - (u_h + \varepsilon_h), u - (u_h + u_b)) \tag{18}$$

$$\lesssim \|u - (u_h + \varepsilon_h)\|_{1,\Omega} \|u - (u_h + u_b)\|_{1,\Omega}. \tag{19}$$

We bound the term $\|u - (u_h + u_b)\|_{1,\Omega}$ as follows:

$$\|u - (u_h + u_b)\|_{1,\Omega} \le \|u - u_q\|_{1,\Omega} + \|u_q - (u_h + u_b)\|_{1,\Omega} \tag{20}$$

$$= \|u - u_q\|_{1,\Omega} + \|u_\ell - u_h\|_{1,\Omega} \tag{21}$$

$$\lesssim h^2 \|u\|_{3,\Omega} + h^{1+\min(\sigma,1)} |\log h|^{1/2} \|u\|_{3,\infty,\Omega}. \tag{22}$$

This completes the proof. □

As an immediate corollary, we see conditions under which we can expect $\|\varepsilon_h\|_{1,\Omega}$ to be an asymptotically exact estimator of the true gradient error $\|u - u_h\|_{1,\Omega}$.

COROLLARY 3.4. *Suppose that there is some constant $c > 0$ such that $\|u - u_h\|_{1,\Omega} \ge ch$. Then under Assumption* 3.1*, we have*

$$\frac{\|\varepsilon_h\|_{1,\Omega}}{\|u - u_h\|_{1,\Omega}} \to 1. \tag{23}$$

Proof. It holds that

$$\left| \frac{\|\varepsilon_h\|_{1,\Omega}}{\|u - u_h\|_{1,\Omega}} - 1 \right| \le \frac{\|u - (u_h + \varepsilon_h)\|_{1,\Omega}}{\|u - u_h\|_{1,\Omega}}. \tag{24}$$

Combining this with the estimate from Theorem 3.3 completes the proof. □

Theorem 3.3 and Corollary 3.4 and their proofs have also appeared in [17, 18].

Recall that the quadratic interpolant $u_q \in \bar{V}_h$ of u is decomposed as the sum $u_q = u_\ell + u_b$ with $u_\ell \in V_h$ and $u_b \in \tilde{V}_h$. In the following lemma we compare the first and second derivatives of ε_h and u_b. The second of these results is used in the proof of Theorem 3.6 to establish the Hessian recovery result, and the first will play an important role in the next section, where we prove the function recovery result.

LEMMA 3.5. *Under Assumption* 3.1*, we have*

$$\|\varepsilon_h - u_b\|_{1,\Omega} \lesssim h^{1+\min(\sigma,1)} |\log h|^{1/2} \|u\|_{3,\infty,\Omega}, \tag{25}$$

$$\sum_{\tau \in \mathcal{T}_h} |\varepsilon_h - u_b|_{2,\tau}^2 \lesssim h^{2\min(\sigma,1)} |\log h| \|u\|_{3,\infty,\Omega}^2. \tag{26}$$

Proof. In the proof of Theorem 3.3, we saw that

$$\|u-(u_h+\varepsilon_h)\|_{1,\Omega},\ \|u-(u_h+u_b)\|_{1,\Omega} \lesssim h^{1+\min(\sigma,1)}|\log h|^{1/2}\|u\|_{3,\infty,\Omega}. \tag{27}$$

This gives us

$$\|\varepsilon_h-u_b\|_{1,\Omega} \le \|u-(u_h+\varepsilon_h)\|_{1,\Omega}+\|u-(u_h+u_b)\|_{1,\Omega} \tag{28}$$
$$\lesssim h^{1+\min(\sigma,1)}|\log h|^{1/2}\|u\|_{3,\infty,\Omega}. \tag{29}$$

Using a standard inverse estimate, we see that

$$\sum_{\tau\in\mathcal{T}_h}|\varepsilon_h-u_b|^2_{2,\tau} \lesssim h^{-2}\|\varepsilon_h-u_b\|^2_{1,\Omega} \lesssim h^{2\min(1,\sigma)}\,|\log h|\,\|u\|^2_{3,\infty,\Omega}, \tag{30}$$

so we have proven both results. ∎

The convergent Hessian recovery result follows.

THEOREM 3.6. *Under Assumption* 3.1*, we have*

$$\sum_{\tau\in\mathcal{T}_h}|u-\varepsilon_h|^2_{2,\tau} \lesssim h^{2\min(\sigma,1)}|\log h|\|u\|^2_{3,\infty,\Omega}. \tag{31}$$

Proof. We have $|u-\varepsilon_h|_{2,\tau} \le |u-u_b|_{2,\tau}+|u_b-\varepsilon_h|_{2,\tau}$, so

$$\sum_{\tau\in\mathcal{T}_h}|u-\varepsilon_h|^2_{2,\tau} \le 2\left(\sum_{\tau\in\mathcal{T}_h}|u-u_b|^2_{2,\tau}+\sum_{\tau\in\mathcal{T}_h}|u_b-\varepsilon_h|^2_{2,\tau}\right) \tag{32}$$
$$\lesssim h^2\|u\|^2_{3,\infty,\Omega}+\sum_{\tau\in\mathcal{T}_h}|u_b-\varepsilon_h|^2_{2,\tau}. \tag{33}$$

Combining this with the second estimate in Lemma 3.5 completes the proof. ∎

Provided that $\|u\|_{3,\infty,\Omega}<\infty$, the estimate in Theorem 3.5 is equivalent to

$$\sum_{\tau\in\mathcal{T}_h}|\varepsilon_h|^2_{2,\tau} \to |u|^2_{2,\Omega}. \tag{34}$$

4. Function recovery. In this section we prove that the approximate error function ε_h provides efficient and reliable approximation of the true error $u-u_h$ in the L_2-norm,

$$c_1 \le \frac{\|\varepsilon_h\|_{0,\Omega}}{\|u-u_h\|_{0,\Omega}} \le c_2. \tag{35}$$

We also explain why we cannot generally expect the same sort of asymptotic exactness result which we saw for the gradient error. In other words, we *cannot* generally expect that

$$\frac{\|\varepsilon_h\|_{0,\Omega}}{\|u-u_h\|_{0,\Omega}} \to 1, \tag{36}$$

although the constants c_1, c_2 may be near 1 in practice.

This first lemma will allow us to convert the gradient approximation result from Lemma 3.5 into the function (L_2) approximation results that follow.

LEMMA 4.1. *Let $\mathcal{T}_h$ be a shape-regular quasi-uniform mesh. For any $b \in \tilde{V}_h$, we have*

$$\|b\|_{0,\Omega} \lesssim h\|\nabla b\|_{0,\Omega}. \tag{37}$$

Proof. Let $\tau \in \mathcal{T}_h$ be given, and write b in terms of its three bump basis functions on τ, $b = c_1 b_1 + c_2 b_2 + c_3 b_3$. We denote the length of the edge on which b_k does *not* vanish by L_k, and without loss of generality take $L_1 \le L_2 \le L_3$. It holds that

$$\|b\|_{0,\tau}^2 = \frac{8|\tau|}{45}(c_1^2 + c_2^2 + c_3^2 + c_1c_2 + c_1c_3 + c_2c_3), \tag{38}$$

$$\|\nabla b\|_{0,\tau}^2 = \frac{1}{3|\tau|}\left((c_1 - c_2 - c_3)^2 L_1^2 + (c_2 - c_1 - c_3)^2 L_2^2 + (c_3 - c_1 - c_2)^2 L_3^2\right). \tag{39}$$

We bound $\|\nabla b\|_{0,\tau}^2$ from below as follows:

$$\begin{aligned}
\|\nabla b\|_{0,\tau}^2 &\ge \frac{L_1^2}{3|\tau|}\left((c_1 - c_2 - c_3)^2 + (c_2 - c_1 - c_3)^2 + (c_3 - c_1 - c_2)^2\right) && (40)\\
&= \frac{L_1^2}{3|\tau|}(3c_1^2 + 3c_2^2 + 3c_3^2 - 2c_1c_2 - 2c_1c_3 - 2c_2c_3) && (41)\\
&\ge \frac{L_1^2}{3|\tau|}\frac{1}{2}(c_1^2 + c_2^2 + c_3^2 + c_1c_2 + c_1c_3 + c_2c_3). && (42)
\end{aligned}$$

This gives us

$$\|b\|_{0,\tau}^2 \le \frac{48}{45}\frac{|\tau|^2}{L_1^2}\|\nabla b\|_{0,\tau}^2 \lesssim h^2\|\nabla b\|_{0,\tau}^2. \tag{43}$$

Summing over triangles completes the proof. ◻

LEMMA 4.2. *Under Assumption* 3.1, *we have*

$$\|\varepsilon_h - u_b\|_{0,\Omega} \lesssim h^{2+\min(\sigma,1)}|\log h|^{1/2}\|u\|_{3,\infty,\Omega}, \tag{44}$$

$$\|u - (u_\ell + \varepsilon_h)\|_{0,\Omega} \lesssim h^{2+\min(\sigma,1)}|\log h|^{1/2}\|u\|_{3,\infty,\Omega}. \tag{45}$$

Proof. Combining the first estimate from Lemma 3.5 with the result of Lemma 4.1 proves the first of these two estimates. We also have

$$\|u - (u_\ell + \varepsilon_h)\|_{0,\Omega} \le \|u - u_q\|_{0,\Omega} + \|\varepsilon_h - u_b\|_{0,\Omega} \lesssim h^3\|u\|_{3,\Omega} + \|\varepsilon_h - u_b\|_{0,\Omega}. \tag{46}$$

Combining this second estimate with the first completes the proof. ◻

We see from the estimate $\|u - (u_\ell + \varepsilon_h)\|_{0,\Omega} = o(h^2)$ that $\|\varepsilon_h\|_{0,\Omega}$ is an asymptotically exact estimator of the interpolation error $\|u - u_\ell\|_{0,\Omega}$, provided that $\|u - u_\ell\|_{0,\Omega} > m_1 h^2$ for some positive constant m_1. We are now ready to prove the main result of this section.

THEOREM 4.3. *Suppose that there are constants $m_1, m_2 > 0$, such that $\|u - u_\ell\|_{0,\Omega} \ge m_1 h^2$ and $\|u - u_h\|_{0,\Omega} \ge m_2 h^2$. Then, under Assumption* 3.1, *there are constants $c_1, c_2 > 0$, such that*

$$c_1 \le \frac{\|\varepsilon_h\|_{0,\Omega}}{\|u - u_h\|_{0,\Omega}} \le c_2. \tag{47}$$

Proof. It is certainly the case that there are constants $M_1, M_2 > 0$, such that $\|u - u_\ell\|_{0,\Omega} \le M_1 h^2$ and $\|u - u_h\|_{0,\Omega} \le M_2 h^2$. So we have

$$\frac{m_1}{M_2} \le \frac{\|u - u_\ell\|_{0,\Omega}}{\|u - u_h\|_{0,\Omega}} \le \frac{M_1}{m_2}. \tag{48}$$

The proof is completed by using the fact that $\|\varepsilon_h\|_{0,\Omega}$ is an asymptotically exact estimator of $\|u - u_\ell\|_{0,\Omega}$. □

Recall that the proof of the asymptotic exactness of $\|\varepsilon_h\|_{1,\Omega}$ as an estimator of $\|u-u_h\|_{1,\Omega}$ relied on the fact that $\|u_\ell-u_h\|_{1,\Omega} = o(h)$. We see in Lemma 4.4 below that we need $\|u_\ell - u_h\|_{0,\Omega} = o(h^2)$ to get asymptotic exactness of $\|\varepsilon_h\|_{0,\Omega}$ as an estimator of $\|u - u_h\|_{0,\Omega}$.

Lemma 4.4. *Under Assumption* 3.1*, we have*

$$\|u - (u_h + \varepsilon_h)\|_{0,\Omega} = o(h^2) \iff \|u_h - u_\ell\|_{0,\Omega} = o(h^2). \tag{49}$$

Proof. We have the inequalities

$$\|u - (u_h + \varepsilon_h)\|_{0,\Omega} \le \|u - (u_\ell + \varepsilon_h)\|_{0,\Omega} + \|u_h - u_\ell\|_{0,\Omega}, \tag{50}$$

$$\|u_h - u_\ell\|_{0,\Omega} \le \|u - (u_\ell + \varepsilon_h)\|_{0,\Omega} + \|u - (u_h + \varepsilon_h)\|_{0,\Omega}. \tag{51}$$

Lemma 4.2 completes the proof. □

The rest of this section is devoted to demonstrating by example that we *cannot* generally expect $\|u_\ell-u_h\|_{0,\Omega} = o(h^2)$ even in an ideal situation for which we can prove $\|u_\ell - u_h\|_{1,\Omega} \lesssim h^2|\log h|^{1/2}\|u\|_{3,\infty,\Omega}$. Thus, we cannot generally expect asymptotic exactness in the L_2-norm.

Consider the following simple problem on the unit square $\Omega = (0,1) \times (0,1)$:

$$\begin{aligned} -\Delta u = 2x(1-x) + 2y(1-y) \text{ in } \Omega,\\ u = 0 \text{ on } \partial\Omega. \end{aligned}$$

The exact solution is $u = x(1-x)y(1-y)$. We take the family of uniform meshes having mesh size $h = \frac{1}{n+1}$ and n^2 degrees of freedom located at $(x_i, y_j) = (ih, jh)$; see Figure 1. We will show that $h^2 \lesssim \|u_\ell - u_h\|_{0,\Omega}$.

Let $T \in \mathbb{R}^{n\times n}$ be the tridiagonal matrix with stencil $(-1, 2, -1)$. Under the standard ordering of unknowns (left to right, bottom to top) the stiffness matrix for this problem is given by

$$A = T \otimes I + I \otimes T = (V \otimes V)(D \otimes I + I \otimes D)(V \otimes V), \tag{52}$$

$$V_{ij} = \sqrt{\frac{2}{n+1}} \sin \frac{ij\pi}{n+1}\ , \ \ D_{ij} = \delta_{ij}\left(2 - 2\cos\frac{i\pi}{n+1}\right) = \delta_{ij} 4 \sin^2 \frac{i\pi}{2(n+1)}. \tag{53}$$

We note that $V = V^T = V^{-1}$. As a notational convenience, for $\mathbf{x} \in \mathbb{R}^{n^2}$ we use $\mathbf{x}_{(i,j)} \equiv \mathbf{x}_{(i-1)n+j}$. Similarly, we take $\phi_{(i,j)}$ to be the Lagrange nodal basis function associated with the grid point (x_i, y_j). We define $\mathbf{d}$ and $\mathbf{r}$ to be the error and residual, respectively, at the grid points

$$\mathbf{d}_{(i,j)} = u(x_i, y_j) - u_h(x_i, y_j) = u_\ell(x_i, y_j) - u_h(x_i, y_j), \tag{54}$$

$$\mathbf{r}_{(i,j)} = h^2 f(x_i, y_j) - \int_\Omega f\phi_{(i,j)}\, dxdy = \frac{2}{3}h^4. \tag{55}$$

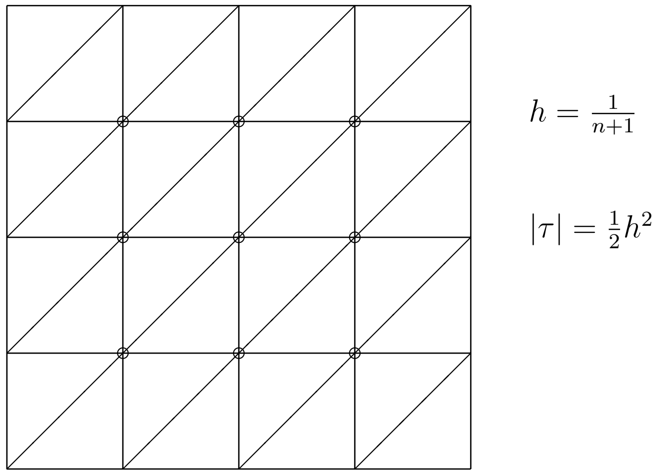

FIG. 1. *Uniform mesh with $n = 3$.*

We have $A\mathbf{d} = \mathbf{r}$. We first argue that $\|u_\ell - u_h\|_{0,\Omega} \geq \frac{h}{2}\|\mathbf{d}\|$, and then establish that $\|\mathbf{d}\| \geq Ch$, thereby proving that $h^2 \lesssim \|u_\ell - u_h\|_{0,\Omega}$. We begin by noting that for any linear function g on a triangle τ, given in terms of its three nodal basis functions, $g = c_1\ell_1 + c_2\ell_2 + c_3\ell_3$, we have

$$\|g\|^2_{0,\tau} = \frac{|\tau|}{6}(c_1^2 + c_2^2 + c_3^2 + c_1c_2 + c_1c_3 + c_2c_3) \geq \frac{|\tau|}{12}(c_1^2 + c_2^2 + c_3^2). \tag{56}$$

Therefore, if g is continuous and piecewise-linear on $\mathcal{T}$, we have

$$\|g\|^2_{0,\Omega} = \sum_{\tau\in\mathcal{T}_h} \|g\|^2_{0,\tau} \geq \frac{|\tau|}{2}\|\mathbf{c}\|^2 = \frac{h^2}{4}\|\mathbf{c}\|^2, \tag{57}$$

where $\mathbf{c}$ is the vector of coefficients. The factor of 6 comes from the fact that each coefficient appears in 6 of the summands $\|g\|^2_{0,\tau}$. This proves that

$$\|u_\ell - u_h\|_{0,\Omega} \geq \frac{h}{2}\|\mathbf{d}\|. \tag{58}$$

We now consider $\|\mathbf{d}\| = \|A^{-1}\mathbf{r}\| = \frac{2}{3}h^4\|A^{-1}(\mathbf{e}\otimes\mathbf{e})\|$, where $\mathbf{e} \in \mathbb{R}^n$ is the vector of ones. It holds that $\|A^{-1}(\mathbf{e}\otimes\mathbf{e})\| = \|(D\otimes I + I\otimes D)^{-1}(V\mathbf{e}\otimes V\mathbf{e})\|$, and

$$(V\mathbf{e})_i = \sqrt{\frac{2}{n+1}}\sum_{j=1}^{n}\sin\frac{ij\pi}{n+1} = \sqrt{\frac{2}{n+1}}\cot\frac{i\pi}{2(n+1)}\left|\sin\frac{i\pi}{2}\right|. \tag{59}$$

This gives us

$$\|A^{-1}(\mathbf{e}\otimes\mathbf{e})\|^2 = \frac{h^2}{4}\sum_{i=1}^{n}\sum_{j=1}^{n}\left|\sin\frac{i\pi}{2}\sin\frac{j\pi}{2}\right|\left(\frac{\cot\frac{i\pi}{2(n+1)}\cot\frac{j\pi}{2(n+1)}}{\sin^2\frac{i\pi}{2(n+1)} + \sin^2\frac{j\pi}{2(n+1)}}\right)^2 \tag{60}$$

$$> \frac{h^2}{4}\left(\frac{\cot\frac{\pi}{2(n+1)}\cot\frac{\pi}{2(n+1)}}{\sin^2\frac{\pi}{2(n+1)} + \sin^2\frac{\pi}{2(n+1)}}\right)^2 \tag{61}$$

$$= \frac{h^2}{16}\frac{\cos^4\frac{\pi}{2(n+1)}}{\sin^8\frac{\pi}{2(n+1)}} > \frac{h^2}{16}\frac{(\frac{1}{\sqrt{2}})^4}{(\frac{\pi}{2(n+1)})^8} = \frac{4}{\pi^8}h^{-6}. \tag{62}$$

Combining these results we have $\|u_\ell - u_h\|_{0,\Omega} > \frac{h}{2}\frac{2h^4}{3}\frac{2h^{-3}}{\pi^4} = \frac{2h^2}{3\pi^4}$, which completes the argument.

5. Experiments. In this section we offer four examples which illustrate the effectivity of our estimator and provide some comments on its computational cost. In particular, we wish to verify (2), the key results of this paper, in practice. The exact error for each of the examples solution is known, so we can judge the quality of our estimator directly. Throughout this section we use $e_h \equiv u - u_h$ for the exact error and the abbreviation EFF for each of the effectivity ratios

$$\frac{\|\varepsilon_h\|_{0,\Omega}}{\|e_h\|_{0,\Omega}}, \qquad \frac{|\varepsilon_h|_{1,\Omega}}{|e_h|_{1,\Omega}}, \qquad \frac{|\varepsilon_h|_{2,\Omega}}{|u|_{2,\Omega}}. \tag{63}$$

For the sake of convenience we abuse notation slightly by taking

$$|\varepsilon_h|_{2,\Omega} \equiv \sqrt{\sum_{\tau\in\mathcal{T}} |\varepsilon_h|^2_{2,\tau}}. \tag{64}$$

This is an abuse because $|v|_{2,\Omega}$ is infinite by its standard definition for functions such as ε_h, which have a gradient jump between elements in a mesh. Additionally, we abbreviate the standard scientific notation by placing the base 10 exponent as a subscript, for example, $3.54_{-2} \equiv 3.54 \times 10^{-2}$.

The quantity N appearing in the tables is the number of triangles in the mesh. For the larger values of N, this is roughly twice the number of vertices in the mesh. In the first four examples, for which the exact error is known, we use the error model $E = CN^{-p}$, derived from standard a priori estimates and $Nh^2 \sim 1$, to give a sense of the rate of convergence of error. In particular, we give the least-squares best fit for each of the normed errors. We note that $p = 1$ (resp., $p = 1/2$) corresponds to what is generally called *quadratic* (resp., *linear*) convergence—in terms of the mesh parameter h—and we use this language in the explanations below. The code used for the numerical experiments is PLTMG [6], with modifications necessary to implement our error estimation technique.

5.1. The simple problem. For our first experiment, we revisit the example from section 4 which was used to demonstrate that one cannot generally expect asymptotic exactness from our estimator in L_2. We will see, however, that the function recovery is very nearly exact in this case. Recall that the problem is to find u such that

$$-\Delta u = 2x(1-x) + 2y(1-y) \text{ in } \Omega,$$
$$u = 0 \text{ on } \partial\Omega.$$

Here Ω is the unit square, and the exact solution is $u = x(1-x)y(1-y)$. We provide the values of the various norms of u so that the relative errors can be readily assessed if desired:

$$\|u\|_{0,\Omega} = \sqrt{\frac{1}{900}} = 0.0\bar{3}, \qquad |u|_{1,\Omega} = \sqrt{\frac{1}{45}} \approx 0.149, \qquad |u|_{2,\Omega} = \sqrt{\frac{22}{45}} \approx 0.699.$$

This example is also used in the numerical experiments in [21, 23].

In Table 1 we see the predicted performance of the estimator in each of the square norms, with the L_2 error estimate having effectivity very near 1 on each mesh. Below, we give the approximate error models for the function and gradient errors:

$$\|e_h\|_{0,\Omega} \approx 0.159N^{-1.02}, \qquad |e_h|_{1,\Omega} \approx 0.307N^{-0.502}.$$

Table 1
Estimates, exact values, and effectivity for the simple problem.

N	88	441	1887	7765	31505	126919
$\|\varepsilon_h\|_{0,\Omega}$	1.71_{-3}	2.93_{-4}	6.98_{-5}	1.62_{-5}	3.90_{-6}	9.64_{-7}
$\|e_h\|_{0,\Omega}$	1.65_{-3}	3.09_{-4}	7.22_{-5}	1.63_{-5}	3.93_{-6}	9.76_{-7}
EFF	1.04	0.947	0.966	0.993	0.994	0.987
$\|\varepsilon_h\|_{1,\Omega}$	3.19_{-2}	1.36_{-2}	6.61_{-3}	3.14_{-3}	1.54_{-3}	7.67_{-4}
$\|e_h\|_{1,\Omega}$	3.14_{-2}	1.37_{-2}	6.61_{-3}	3.15_{-3}	1.55_{-3}	7.72_{-4}
EFF	1.01	0.998	1.00	0.997	0.996	0.996
$\|\varepsilon_h\|_{2,\Omega}$	0.726	0.713	0.709	0.705	0.703	0.703
$\|u\|_{2,\Omega}$	0.699	0.699	0.699	0.699	0.699	0.699
EFF	1.04	1.02	1.01	1.01	1.01	1.00

Table 2
Estimates, exact values, and effectivity for the oscillatory problem.

N	88	434	1888	7825	31679	127552
$\|\varepsilon_h\|_{0,\Omega}$	0.369	0.149	8.43_{-2}	1.50_{-2}	3.27_{-3}	7.99_{-4}
$\|e_h\|_{0,\Omega}$	0.499	0.172	9.60_{-2}	1.76_{-2}	3.89_{-3}	9.49_{-4}
EFF	0.738	0.865	0.878	0.853	0.846	0.842
$\|\varepsilon_h\|_{1,\Omega}$	15.1	8.43	6.56	3.04	1.46	0.716
$\|e_h\|_{1,\Omega}$	17.6	9.78	6.92	3.07	1.46	0.720
EFF	0.859	0.862	0.949	0.991	0.993	0.995
$\|\varepsilon_h\|_{2,\Omega}$	304	458	603	632	634	634
$\|u\|_{2,\Omega}$	632	632	632	632	632	632
EFF	0.481	0.693	0.954	1.00	1.00	1.00

We point out that we observe the predicted a priori quadratic convergence of $\|e_h\|_{0,\Omega}$ and linear convergence of $|e_h|_{1,\Omega}$.

5.2. The oscillatory problem. In this second example we consider the situation where the exact solution still possesses no singularities, but oscillates strongly. The problem is to find u such that

$$-\Delta u = 128\pi^2 \sin(8\pi x)\sin(8\pi y) \text{ in } \Omega,$$
$$u = 0 \text{ on } \partial\Omega.$$

Here Ω is again the unit square, and the exact solution is $u = \sin(8\pi x)\sin(8\pi y)$. The pertinent norms of u are given below:

$$\|u\|_{0,\Omega} = \sqrt{\frac{1}{4}} = 0.5, \qquad |u|_{1,\Omega} = \sqrt{32\pi^2} \approx 17.8, \qquad |u|_{2,\Omega} = \sqrt{4096\pi^4} \approx 632.$$

In Table 2 we again see effectivity approaching 1 for the gradient error and the Hessian in both norms. The effectivity of the function error estimate tends to stay in the 80–85% range. We see in the approximate error models below that the adaptive refinement seems to be producing suboptimal reduction of function and gradient error:

$$\|e_h\|_{0,\Omega} \approx 36.5N^{-0.873}, \qquad |e_h|_{1,\Omega} \approx 149N^{-0.443}.$$

This is due to the fact that the two coarsest meshes are just beginning to resolve the oscillatory behavior. When the error data from these two initial meshes is removed, we see the expected quadratic and linear convergence for the function and gradient errors, respectively. More precisely, the exponents for the L_2 and H^1 error models are $p = 1.09$ and 0.536, respectively.

Table 3
Estimates, exact values, and effectivity for the slit domain problem.

N	94	481	2031	8334	33704	135632
$\\|\varepsilon_h\\|_{0,\Omega}$	2.81_{-2}	3.20_{-3}	5.26_{-4}	1.39_{-4}	3.43_{-5}	8.51_{-6}
$\\|e_h\\|_{0,\Omega}$	0.122	3.92_{-2}	1.33_{-2}	3.57_{-3}	9.08_{-4}	1.78_{-4}
EFF	0.230	8.18_{-2}	3.96_{-2}	3.88_{-2}	3.78_{-2}	4.78_{-2}
$\|\varepsilon_h\|_{1,\Omega}$	0.419	0.231	0.132	6.93_{-2}	3.51_{-2}	1.62_{-2}
$\|e_h\|_{1,\Omega}$	0.590	0.331	0.189	9.91_{-2}	4.99_{-2}	2.25_{-2}
EFF	0.710	0.698	0.697	0.699	0.703	0.720
$\|\varepsilon_h\|_{2,\Omega_s}$	5.34	19.9	24.2	18.2	17.5	17.2
$\|u\|_{2,\Omega_s}$	17.2	17.2	17.2	17.2	17.2	17.2
EFF	0.310	1.16	1.40	1.06	1.02	1.00
N	94	481	2031	8334	33704	135632
$\\|\varepsilon_h\\|_{0,\Omega_s}$	2.81_{-2}	3.20_{-3}	5.04_{-4}	1.38_{-4}	3.43_{-5}	8.51_{-6}
$\\|e_h\\|_{0,\Omega_s}$	0.122	3.92_{-2}	1.32_{-2}	3.56_{-3}	9.04_{-4}	1.78_{-4}
EFF	0.230	8.18_{-2}	3.81_{-2}	3.89_{-2}	3.79_{-2}	4.80_{-2}
$\|\varepsilon_h\|_{1,\Omega_s}$	0.419	0.231	6.28_{-2}	2.69_{-2}	1.37_{-2}	6.89_{-3}
$\|e_h\|_{1,\Omega_s}$	0.590	0.331	0.119	3.47_{-2}	1.48_{-2}	6.99_{-3}
EFF	0.710	0.698	0.526	0.774	0.925	0.986

5.3. The slit domain problem. For our third example we consider a problem for which the boundary conditions force a singularity at the origin. Because of the infinite gradient at the origin, it is interesting to investigate the effectivity of the estimators. The problem is to find u such that

$$-\Delta u = 0 \text{ in } \Omega, \qquad u(r, 0^+) = 0, \qquad \nabla u \cdot \mathbf{n}\,(r, 2\pi^-) = 0, \qquad u(1,\theta) = \sin(\theta/4).$$

Here Ω is the unit disk with the positive x-axis removed, and the exact solution is $u = r^{1/4}\sin(\theta/4)$. Though the gradient of u is infinite at the origin, $|u|_{1,\Omega}$ is finite. However, this is not the case for $|u|_{2,\Omega}$—here we must avoid the origin to get a finite H^2 seminorm. Let Ω_s denote Ω with the disk of radius s about the origin removed. In the experiments, we take $s = 1/100$. The pertinent norms are given below:

$$\|u\|_{0,\Omega} = \sqrt{\frac{2\pi}{5}} \approx 1.12,\ |u|_{1,\Omega} = \sqrt{\frac{\pi}{4}} \approx 0.886,\ |u|_{2,\Omega_s} = \sqrt{\frac{3\pi}{32}(s^{-3/2} - 1)} \approx 17.2.$$

We note that the global smoothness condition $u \in W^3_\infty(\Omega)$ is certainly not satisfied here.

In Table 3 we see the clear effects of this singularity on the performance of the function error estimates and the gradient error. Here the function error estimates underestimate the true function error by roughly a factor of 26.5 at worst and a factor of 5 at best, and the gradient error estimate underestimates the true gradient error by 28% at best, though it is slowly improving. We also point out that the second derivatives are recovered quite well. Concerning Table 3, we mention finally that the performance of the gradient error estimate improves markedly if we restrict our attention to the error on the subdomain Ω_s, as is seen at the bottom of that table, but the performance of the function error estimate does not improve appreciably. The approximate error models given below, though showing subquadratic convergence of the function error and sublinear convergence of the gradient error, are actually quite encouraging for a problem with this sort of singularity, where we would expect $p \approx 1/8$ asymptotically for the gradient error $|e_h|_{1,\Omega}$ under uniform refinement:

$$\|e_h\|_{0,\Omega} \approx 9.31N^{-0.894}, \qquad |e_h|_{1,\Omega} \approx 5.11N^{-0.447}.$$

Table 4
Estimates, exact values, and effectivity for the jumping coefficient problem.

N	66	353	1530	6337	25734	103617
$\|\varepsilon_h\|_{0,\Omega}$	6.67	0.396	8.04_{-2}	1.86_{-2}	4.33_{-3}	1.22_{-3}
$\|e_h\|_{0,\Omega}$	11.1	0.811	0.116	2.29_{-2}	5.21_{-3}	1.49_{-3}
EFF	0.603	0.488	0.691	0.814	0.831	0.820
$\|\varepsilon_h\|_{1,\Omega}$	96.0	33.3	13.6	6.06	2.90	1.42
$\|e_h\|_{1,\Omega}$	108	36.1	14.1	6.16	2.92	1.43
EFF	0.886	0.923	0.960	0.985	0.994	0.997
$\|\varepsilon_h\|_{2,\Omega_s}$	1.17_3	2.49_3	2.50_3	2.40_3	2.35_3	2.33_3
$\|u\|_{2,\Omega_s}$	2.32_3	2.32_3	2.32_3	2.32_3	2.32_3	2.32_3
EFF	0.504	1.07	1.07	1.03	1.01	1.00

5.4. The jumping coefficient problem. The problem is to find u such that

$$-a_k \Delta u = 0 \text{ in } \Omega, u(r,0) = 0,$$
$$\nabla u \cdot \mathbf{n}\,(r,\pi) = 0, u(1,\theta) = b_k \sin(\alpha\theta) + c_k \cos(\alpha\theta).$$

Here Ω is the upper half of the unit disk, which is divided into two regions having differing coefficients of diffusion. In the first region, $0 < \theta < \frac{\pi}{4}$, we have $a_1 = 10^3$. In the second region, $\frac{\pi}{4} < \theta < \pi$, we have $a_2 = 1$. The exact solution is $u = r^\alpha(b_k \sin(\alpha\theta) + c_k \cos(\alpha\theta))$, where the values α, b_k, c_k are determined by the boundary conditions at $\theta = 0, \pi$ and the continuity of u and $a_k \nabla u \cdot \mathbf{n}$ along the interface $\theta = \frac{\pi}{4}$ between the two regions. The boundary condition at $r = 1$ is chosen to match the solution in the interior. The boundary conditions on the positive and negative x-axes and the continuity conditions at the interface provide four equations which are linear in b_1, c_1, b_2, c_2 (and trigonometric in α). It is clear that $b_1 = c_1 = b_2 = c_2 = 0$ trivially satisfies all of the specified conditions, so we must select α so that the resulting linear system is singular—therefore admitting nontrivial solutions. If there are any such choices of α, then there are infinitely many. We selected the following solution, with $\alpha \approx 0.666422$:

$$b_1 = 1, \qquad c_1 = 0, \qquad b_2 \approx 750.416, \qquad c_2 \approx -432.484,$$
$$\|u\|_{0,\Omega} \approx 515, \qquad |u|_{1,\Omega} \approx 767, \qquad |u|_{2,\Omega_s} \approx 2.32_3.$$

Again we take Ω_s to be Ω with the disk of radius $s = 1/100$ removed. Although $u \notin H^2(\Omega_s)$ because of the jump discontinuity of ∇u at the interface between the two regions, we abuse notation by taking

$$|u|^2_{2,\Omega_s} \equiv \sum_{\tau \in \mathcal{T}_s} |u|^2_{2,\tau} \qquad (65)$$

for Table 4. This sum is finite because the interface between the two regions does not pass through the interior of any of the triangles.

In Table 4, we see the data for this experiment. We point out that the performance of the various error estimates based on the approximate error function seem to be unaffected by the jump in the coefficient. In particular, we see effectivity ratios near or approaching 1 for the gradient error and the Hessian, and slightly better than 80% for the function values in each norm. The approximate error models given below show error convergence which is better than one would expect, with superquadratic convergence in function error and superlinear convergence in gradient error:

$$\|e_h\|_{0,\Omega} \approx 1.12_3 N^{-1.20}, \qquad |e_h|_{1,\Omega} \approx 1.58_3 N^{-0.589}.$$

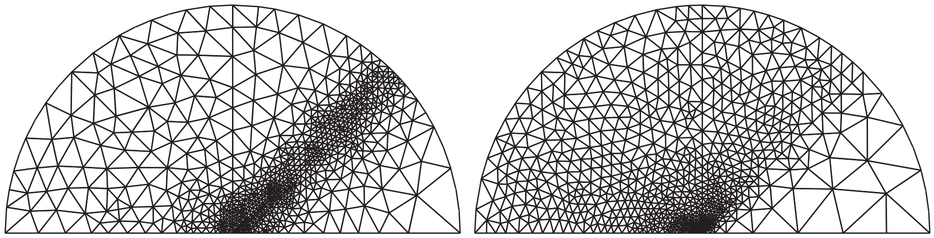

FIG. 2. *The meshes for the jumping coefficient problem after three stages of adaptive refinement, using Bank–Xu gradient recovery estimates (left) and bump function error estimates (right). The mesh on the left has* 804 *vertices and* 1534 *triangles, and the mesh on the right has* 808 *vertices and 1530 triangles.*

These convergence rates are elevated in the models because of the significant error reduction in the early stages of refinement. When we remove the error data from the first two meshes, the convergence rates drop to the more normal quadratic and linear levels.

In addition to having an r^α, $\alpha < 1$ singularity at the origin, the solution also has a jump discontinuity in its gradient at $\theta = \pi/4$. It is relevant at this point to consider which of these two types of singularities has the stronger (negative) effect on the performance of the estimator for problems of this sort. Considering that the slit domain problem possesses only an r^α singularity and that the α for that problem is smaller than the one for this problem, comparing the performance of the estimator in both cases suggests that r^α singularities are more influential than jump discontinuities in the gradient. In fact, a careful reading of either the Bank–Xu paper [3] or the Xu–Zhang paper [20] reveals that the key superconvergence result for this paper,

$$\|u_h - u_\ell\|_{1,\Omega} = o(h),$$

holds for u having a finite number of gradient jump discontinuities provided that u is sufficiently smooth in each of corresponding subdomains. So we see that, asymptotically, the effectivity of the estimator is affected by jumping coefficients only if they lead to singularities which are worse than gradient jump discontinuities.

We also mention that, for problems of the sort for which we can expect gradient jumps, a naive application of gradient recovery error estimators will lead to suboptimal and sometimes terrible performance. This is because of the fact that gradient recovery schemes involve some sort of local or global averaging. If care is not taken to avoid averaging across an interface where ∇u jumps, then the local error estimates near the interface will tend to overestimate the actual error there—particularly when u_h approximates u well. To illustrate this explicitly we give a brief summary of the result using the Bank–Xu recovery technique, which is a global recovery technique. In Figure 2, we see a clear qualitative difference between the sort of refinement produced by the bump estimator and the naive use of the Bank–Xu estimator—the sort of difference we might have guessed due to the overestimation of error near the interface for the latter. The error model for this refinement is $|e_h|_{1,\Omega} \approx 845\, N^{-0.487}$, with effectivity $EFF \approx 3$ as the mesh is refined. We are not trying to make the point that this sort of bad behavior is unavoidable for gradient recovery schemes—in practice it can be avoided by taking care to not average out a gradient jump where there should be one. Bank and Xu noted this in an example in [4], and performing their gradient recovery scheme for our problem on each subdomain separately restores the optimal performance. The point that we are trying to make with this discussion is that with

TABLE 5

Timing comparison: the ratio of the costs to compute ε_h and $\mathcal{R}\nabla u_h$. Ratios in the first three rows correspond to using SGS-CG to compute ε_h, and the bottom three rows correspond to using unpreconditioned CG.

Simple	3.17	3.37	3.53	3.25	2.66	2.24
Oscillatory	3.15	4.07	3.61	3.40	2.05	2.38
Slit domain	3.19	2.83	2.85	3.01	2.66	2.03
Simple	2.50	2.56	2.63	2.38	1.98	1.49
Oscillatory	2.45	2.53	2.74	2.59	1.44	1.53
Slit domain	2.55	2.09	2.15	2.12	1.94	1.40

the bump error estimator it is not necessary to treat subdomains differently. We think that this is an attractive feature of the estimator, particularly in cases where the number of jumps in the coefficient on the diffusion term (and hence the number of jumps in the gradient of the solution) is large, or where there are small or narrow regions in which the number of elements needed to get a good approximation of the true solution there is smaller than the number of elements needed to perform any of the standard gradient recovery techniques.

5.5. Computational cost. Although the linear system involved in the computation of ε_h can be expected to have roughly three times the number of unknowns as that for computing u_h, the system itself is readily solved because it is well-conditioned (see [5, p. 11], for example). But how does the cost compare with that of various gradient recovery schemes? We content ourselves with a direct comparison to the recovery scheme of Bank and Xu as it is currently implemented in PLTMG. In Table 5, we have the ratios of the times needed to compute ε_h and the recovered gradient $\mathcal{R}\nabla u_h$ for three of the four problems considered here—the jumping coefficient problem was omitted because it would have required a modification of the gradient recovery subroutines in PLTMG. We have used the symmetric Gauß–Seidel method as a preconditioner for CG in the computation of ε_h, as in all of the experiments above, and these data correspond to those experiments. For example, the ratio 3.17 for the simple problem corresponds to the coarsest mesh (88 triangles for both ε_h and $\mathcal{R}\nabla u_h$), and 2.24 corresponds to the finest mesh (126919 triangles for ε_h and 127020 for $\mathcal{R}\nabla u_h$).

For these three problems, unpreconditioned CG can be used instead with no loss in effectivity. When this is done, the timing ratios improve, as is shown in the bottom three rows of Table 5. We generally advocate using some sort of preconditioner for problems such as the jumping coefficient problem because otherwise one notices a drop in effectivity. We suggest that the greater computational cost, still quite small with respect to the total computational cost of the adaptive algorithm, may be worthwhile for this very robust and flexible error estimator. The robustness of the estimator is seen theoretically in that, even in situations where the assumptions taken here do not apply, we can fall back on the "old" analysis based on the milder saturation assumption and on the strengthened Cauchy inequality, which hold under quite general conditions (see [9] and [10, pp. 436–445]). The flexibility of the approximate error function $\varepsilon_h \approx u - u_h$ is clear in that it can be used to measure error in other norms or to approximate error in certain functionals of interest (see [18]), as well as for mesh smoothing procedures such as that proposed by Bank and Smith [2].

6. Final remarks. We have given proof and numerical evidence of the effectiveness of the hierarchical basis type bump function estimator $\varepsilon_h \approx u - u_h$ in recovering function values and first and second derivatives. The proofs offered here are based on the superconvergence result $\|u_h - u_\ell\|_{1,\Omega} = o(h)$, which is usually used in the proofs of the effectiveness of gradient recovery methods. In our proofs, we replace the standard saturation assumption and strengthened Cauchy inequality used in the analysis commonly given for hierarchical basis methods with relatively mild mesh symmetry conditions and relatively strong smoothness assumptions, which are sufficient but often not seen to be necessary in practice. We thereby obtain stronger theoretical results than are generally given for such estimators, and these results are borne out in practice. The approximation $\varepsilon_h \approx u - u_h$ is provably quite robust and can be used for error estimation and adaptivity in a variety of norms and other measures.

In terms of the asymptotically exact recovery of gradient error, our estimator $\|\nabla\varepsilon_h\|_{0,\Omega}$ has a lot of very good competition in the many gradient recovery procedures proposed in the literature. In addition to the recovery procedure of Bank and Xu, which is mentioned several times above, we also cite the local least-squares fitting of Zienkiewicz and Zhu [23, 24] (perhaps the most popular), the polynomial preserving method of Zhang and Naga [16, 21], and the method proposed by Wiberg and Li [15, 19], which has the most in common with our own in that it can be used directly to produce a locally quadratic (though not globally continuous) approximation of the error $u - u_h$. These methods should also be suitable for recovering second derivatives—Bank and Xu argue as much for their estimator—but not much has been written in the gradient recovery literature about estimating the function error. The notable exception in this regard is in the aforementioned works of Wiberg and Li, where numerical evidence of efficiency and reliability of their estimator are given, but no analysis is provided.

We now briefly consider a few straightforward generalizations of what has been presented here. The $\mathcal{O}(h^{2\sigma})$-irregular triangulation assumption is generalized in [20], where Xu and Zhang call it *Condition*(α, σ). We note that the σ in the Xu–Zhang paper plays the role of the 2σ used in both the Bank–Xu paper and our own, and an $\mathcal{O}(h^{1+\alpha})$-parallelogram property is used instead of an $\mathcal{O}(h^2)$-parallelogram property. In their paper, Xu and Zhang also use the less stringent regularity condition $u \in H^3(\Omega) \cap W^2_\infty(\Omega)$. Under these assumptions and a few natural assumptions on the bilinear form for the problem, they prove that

$$\|u_h - u_\ell\|_{1,\Omega} \lesssim h^{1+\min(\alpha,1/2,\sigma/2)}(\|u\|_{3,\Omega} + |u|_{2,\infty,\Omega}). \tag{66}$$

The results in this paper can be modified in the obvious way to incorporate the Xu–Zhang version of the mesh symmetry conditions and the weaker regularity assumption, with no change in the proofs.

We will mention two other ways in which the arguments given here can be readily generalized. The first is to consider linear simplicial elements in $\mathbb{R}^n$, $n > 2$. Recall that the key result from which all of the other estimates were proved was of the form

$$\|u_h - u_\ell\|_{1,\Omega} = o(h), \tag{67}$$

where u_h is the linear finite element approximation and u_ℓ is the linear Lagrange interpolant. Brandts and Křížek [7, 12] show that

$$\|u_h - u_\ell\|_{1,\Omega} \lesssim h^2\|u\|_{3,\Omega} \tag{68}$$

on very regular meshes for $u \in H^1_0(\Omega) \cap H^s(\Omega)$ and $s = 3$ for $n \leq 5$ and $s > n/2$ for $n \geq 6$. Any s greater than 3 is needed only to ensure that the nodal interpolant u_ℓ can be well-defined. Chen [8] generalizes the argument of [3] to mildly structured tetrahedral meshes in $\mathbb{R}^3$ to obtain

$$\|u_h - u_\ell\|_{1,\Omega} \lesssim h^{1+\min(1,\sigma)} \|u\|_{3,\infty,\Omega}, \tag{69}$$

where $u \in H_0^1(\Omega) \cap W_\infty^3(\Omega)$ and σ measures the violation of an $\mathcal{O}(h^2)$-parallelepiped property. With such superconvergence results, the extension of our results proceeds in the obvious fashion.

Another generalization would be to consider hierarchical error estimators for higher order elements. For example, let $\hat{V}_h = \bar{V}_h \oplus (\hat{V}_h \setminus \bar{V}_h)$ be the piecewise cubic finite element space, which we think of hierarchically. If $\bar{u}_h \in \bar{V}_h$ is the finite element solution, we might want to estimate the error $u - \bar{u}_h$ using a function in $\hat{V}_h \setminus \bar{V}_h$; call it $\bar{\varepsilon}_h$. Li [13, 14] has shown that Lagrange interpolation does not generally give the analogous superconvergence results for elements of degree 3 or higher in $\mathbb{R}^2$, but we are free to use some other appropriate interpolation scheme. Let $\Pi_q : C(\bar{\Omega}) \to \bar{V}_h$ and $\Pi_c : C(\bar{\Omega}) \to \hat{V}_h$ be defined by

$$\begin{aligned}
&\Pi_q u(v_i) = \Pi_c u(v_i) = u(v_i) \text{ for vertices } v_i,\\
&\int_{e_j} u - \Pi_q u \, ds = \int_{e_j} (u - \Pi_c u) v \, ds = 0 \text{ for edges } e_j \text{ and linear functions } v,\\
&\int_\tau u - \Pi_c u \, dx = 0 \text{ for triangles } \tau.
\end{aligned}$$

Huang and Xu [11] argue that

$$\|\bar{u}_h - \Pi_q u\|_{1,\Omega} \lesssim h^{2+\min(1,\sigma)/2} (\|u\|_{4,\Omega} + |u|_{3,\infty,\Omega}), \ \Pi_c u - \Pi_q u \in \hat{V}_h \setminus \bar{V}_h. \tag{70}$$

One might correctly infer from the statement of the result that a similar argument to those found in [3, 20] is used. With an estimate like this, the analogue of Theorem 3.3 can be proved in the obvious way. Using arguments like those given in Lemma 3.5 and Theorem 3.6, we see that our approximate error function $\bar{\varepsilon}_h \approx u - \bar{u}_h$ provides superconvergent approximation of $\|u - \bar{u}_h\|_{2,\Omega}$ and convergent approximation of $\|u\|_{3,\Omega}$. Finally, arguing along the same lines as in section 4 we get even better results than in the case of piecewise linears, because it actually does hold that $\|\bar{u}_h - \Pi_q u\|_{0,\Omega} = o(h^3)$. Huang and Xu have plans to extend their results to higher order elements as well, and the analogous results should be able to be plugged into our framework with little difficulty.

Acknowledgments. The author thanks Wolfgang Hackbusch and other colleagues in the scientific computation group at the Max Planck Institute in Leipzig for many helpful discussions, as well as Randy Bank at University of California, San Diego for assistance with PLTMG, and for useful comments concerning this manuscript. Additionally I thank the referees for their helpful input.

REFERENCES

[1] R. E. Bank and R. K. Smith, *A posteriori error estimates based on hierarchical bases*, SIAM J. Numer. Anal., 30 (1993), pp. 921–935.

[2] R. E. Bank and R. K. Smith, *Mesh smoothing using a posteriori error estimates*, SIAM J. Numer. Anal., 34 (1997), pp. 979–997.

[3] R. E. Bank and J. Xu, *Asymptotically exact a posteriori error estimators, Part* I. *Grids with superconvergence*, SIAM J. Numer. Anal., 41 (2003), pp. 2294–2312.

[4] R. E. Bank and J. Xu, *Asymptotically exact a posteriori error estimators, Part* II. *General unstructured grids*, SIAM J. Numer. Anal., 41 (2003), pp. 2313–2332.

[5] R. E. Bank, *Hierarchical bases and the finite element method*, in Acta Numerica, Volume 5, Cambridge University Press, Cambridge, UK, 1996, pp. 1–43.

[6] R. E. Bank, *Pltmg: A Software Package for Solving Elliptic Partial Differential Equations, Users' Guide* 9.0, Technical report, University of California, San Diego, 2004.

[7] J. Brandts and M. Křížek, *Gradient superconvergence on uniform simplicial partitions of polytopes*, IMA J. Numer. Anal., 23 (2003), pp. 489–505.

[8] L. Chen, *Superconvergence of tetrahedral linear finite elements*, Int. J. Numer. Anal. Model., 3 (2006), pp. 273–282.

[9] W. Dörfler and R. H. Nochetto, *Small data oscillation implies the saturation assumption*, Numer. Math., 91 (2002), pp. 1–12.

[10] A. Ern and J.-L. Guermond, *Theory and Practice of Finite Elements*, Appl. Math. Sci. 159, Springer-Verlag, New York, 2004.

[11] Y. Huang and J. Xu, *Superconvergence of Quadratic Finite Elements on Mildly Structured Grids*, Math. Comput., to appear.

[12] M. Křížek, *Superconvergence phenomena on three-dimensional meshes*, Int. J. Numer. Anal. Model., 2 (2005), pp. 43–56.

[13] B. Li, *Superconvergence for higher-order triangular finite elements*, Chinese J. Numer. Math. Appl., 12 (1990), pp. 75–79.

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

[15] X. D. Li and N.-E. Wiberg, *A posteriori error estimate by element patch post-processing, adaptive analysis in energy and L_2 norms*, Comput. & Structures, 53 (1994), pp. 907–919.

[16] A. Naga and Z. Zhang, *A posteriori error estimates based on the polynomial preserving recovery*, SIAM J. Numer. Anal., 42 (2004), pp. 1780–1800.

[17] J. S. Ovall, *Duality-Based Adaptive Refinement for Elliptic PDEs*, Ph.D. thesis, Department of Mathematics. University of California at San Diego, 2004.

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

[19] N.-E. Wiberg and X. D. Li, *Superconvergent patch recovery of finite-element solution and a posteriori L_2 norm error estimate*, Comm. Numer. Methods Engrg., 10 (1994), pp. 313–320.

[20] J. Xu and Z. Zhang, *Analysis of recovery type a posteriori error estimators for mildly structured grids*, Math. Comp., 73 (2004), pp. 1139–1152.

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

[22] O. C. Zienkiewicz, D. W. Kelly, J. Gago, and I. Babuška, *Hierarchical finite element approaches, error estimates and adaptive refinement*, in The Mathematics of Finite Elements and Applications, IV (Uxbridge, 1981), Academic Press, London, 1982, pp. 313–346.

[23] O. C. Zienkiewicz and J. Z. Zhu, *The superconvergent patch recovery and a posteriori error estimates.* I. *The recovery technique*, Internat. J. Numer. Methods Engrg., 33 (1992), pp. 1331–1364.

[24] O. C. Zienkiewicz and J. Z. Zhu, *The superconvergent patch recovery and a posteriori error estimates.* II. *Error estimates and adaptivity*, Internat. J. Numer. Methods Engrg., 33 (1992), pp. 1365–1382.