7-1-2002

Analysis of Recovery Type A Posteriori Error Estimators for Mildly Structured Grids

Jinchao Xu
Wayne State University

Zhimin Zhang
Wayne State University, zhimin.zhang@wayne.edu

ANALYSIS OF RECOVERY TYPE A POSTERIORI ERROR ESTIMATORS FOR MILDLY STRUCTURED GRIDS

JINCHAO XU and ZHIMIN ZHANG

WAYNE STATE UNIVERSITY

Detroit, MI 48202

Department of Mathematics
Research Report

2002 Series
#7

This research was partly supported by the National Science Foundation.

Analysis of Recovery Type A Posteriori Error Estimators for Mildly Structured Grids

Jinchao Xu*and Zhimin Zhang†

Abstract. Some recovery type error estimators for linear finite element method are analyzed under $O(h^{1+\alpha})$ ($\alpha > 0$) regular grids. Superconvergence is established for recovered gradients by three different methods when solving general non-self-adjoint second-order elliptic equations. As a consequence, *a posteriori* error estimators based on those recovery methods are asymptotically exact.

Key Words and Phrases. Gradient recovery, ZZ patch recovery, superconvergence, *a posteriori* error estimates.

2000 AMS Subject Classification. Primary 65N30, Secondary 65N50, 65N15, 65N12, 65D10, 74S05, 41A10, 41A25.

1. Introduction.

A posteriori error estimates have become standard in modern engineering and scientific computation. There are two types of popular error estimators. One is the residual type, see, e.g., [2, 4], and another is the recovery type, see, e.g., [20]. The most representative recovery type error estimator is the Zienkiewicz-Zhu error estimator [21]. A decade has passed since the first appearance of the Zienkiewicz-Zhu gradient patch recovery, which is based on a local discrete least-squares fitting. The method is now widely used in engineering practice for its robustness in *a posteriori* error estimates and its efficiency in computer implementation. It is a common belief that the robustness of the ZZ recovery is rooted in its superconvergence property under structured meshes. Superconvergence properties of the ZZ recovery are proven by Zhang [16] for all popular elements under rectangular mesh, by Li-Zhang [11] for linear element under strongly regular triangular meshes, and by Zhang-Victory [17] for tensor product element under strongly regular quadrilateral meshes.

While there is a sizable literature on theoretical investments for residual type error estimators (see, e.g., [1, 3, 10, 13] and reference therein), there have not been many theoretical

*The work of this author was supported in part by the National Science Foundation grants DMS-9706949 and Center for Computational Mathematics and Applications, Penn State University.

†The work of this author was supported in part by the National Science Foundation grants DMS-0074301, DMS-0079743, and INT-0196139.

results on recovery type error estimators. Nevertheless, the topic has recently attracted more and more attention in the scientific community, see, e.g., [5, 6, 7, 9, 15, 19], also see recent books [1, 3] for some general discussion and literature.

The recovery type error estimators perform astonishingly well even for unstructured grids. The current paper intends to explain this phenomenon. We observe that for an unstructured mesh, when adaptive is used, a mesh refinement will usually bring in some kind of structure locally. It is then reasonable to assume that for the most part of the domain, every two adjacent triangles form an $O(h^{1+\alpha})$ approximate parallelogram. Under this assumption, we are able to establish superconvergence property of the gradient recovery operator for three popular methods, weighed averaging, local L^2-projection, and the ZZ patch recovery. Furthermore, by utilizing an integral identity for linear element on one triangular element developed by Bank and Xu [5], we are able to generalize their superconvergence result between the finite element solution and the linear interpolation from $O(h^2)$ regular grid to $O(h^{1+\alpha})$ regular grid. Finally, we are able to prove global asymptotic exactness of the three recovery error estimators.

The literature regarding finite element superconvergence theory can be found in the following books [8, 10, 12, 14, 18].

2. Geometry Identities of A Triangle. In this section, we shall generalize the result in [5] for $\alpha = 1$ to all $\alpha > 0$. We follow closely the argument there. We consider in Figure 1, a triangle τ with vertices $\boldsymbol{p}_k^t = (x_k, y_k)$, $1 \leq k \leq 3$, oriented counterclockwise, and corresponding nodal basis functions (barycentric coordinates) $\{\phi_k\}_{k=1}^3$. Let $\{e_k\}_{k=1}^3$ denote the edges of element τ, $\{\theta_k\}_{k=1}^3$ the angles, $\{\boldsymbol{n}_k\}_{k=1}^3$ the unit outward normal vectors, $\{\boldsymbol{t}_k\}_{k=1}^3$ the unit tangent vectors with counterclockwise orientation, $\{\ell_k\}_{k=1}^3$ the edge lengths, and $\{d_k\}_{k=1}^3$ the perpendicular heights. Let $\tilde{\boldsymbol{p}}$ be the point of intersection for the perpendicular bisectors of the three sides of τ. Let $|s_k|$ denote the distance between $\tilde{\boldsymbol{p}}$ and side k. If τ has no obtuse angles, then the s_k will be nonnegative; otherwise, the distance to the side opposite the obtuse angle will be negative.

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

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

The important special case $\mathcal{D}_\tau = I$ corresponds to $-\Delta$, and in this case $\xi_k = \cos\theta_k$. Let $q_k = \phi_{k+1}\phi_{k-1}$ denote the quadratic bump function associated with edge e_k and let

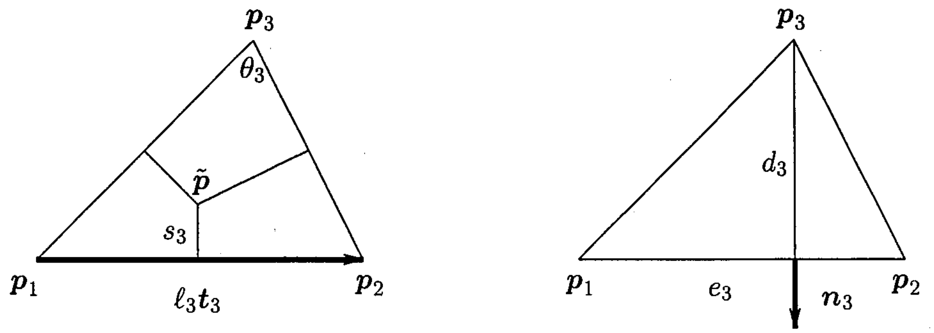

Figure 1: Parameters associated with the triangle τ

$\psi_k = \phi_k(1-\phi_k)$.

The following fundamental identity is proved in [5]:

$$\int_\tau \nabla(u-u_I)\cdot\mathcal{D}_\tau\nabla v_h = \sum_{k=1}^{3}\int_{e_k}\frac{\xi_k q_k}{2\sin\theta_k}\left\{(\ell_{k+1}^2-\ell_{k-1}^2)\frac{\partial^2 u}{\partial \boldsymbol{t}_k^2}+4|\tau|\frac{\partial^2 u}{\partial \boldsymbol{t}_k\partial \boldsymbol{n}_k}\right\}\frac{\partial v_h}{\partial \boldsymbol{t}_k}$$
$$-\int_\tau\sum_{k=1}^{3}\frac{\ell_k\xi_k}{2\sin^2\theta_k}\left\{\ell_{k+1}\psi_{k-1}\frac{\partial^3 u}{\partial^2 \boldsymbol{t}_{k+1}\partial \boldsymbol{t}_{k-1}}+\ell_{k-1}\psi_{k+1}\frac{\partial^3 u}{\partial^2 \boldsymbol{t}_{k-1}\partial \boldsymbol{t}_{k+1}}\right\}\frac{\partial v_h}{\partial \boldsymbol{t}_k}. \tag{2.1}$$

We say that two adjacent triangles (sharing a common edge) form an $O(h^{1+\alpha})$ $(\alpha>0)$ approximate parallelogram if the lengths of any two opposite edges differ only by $O(h^{1+\alpha})$.

Definition: The triangulation $\mathcal{T}_h = \mathcal{T}_{1,h}\cup\mathcal{T}_{2,h}$ is said to satisfy *Condition* (α,σ) if there exists positive constants α and σ such that: Every two adjacent triangles inside $\mathcal{T}_{1,h}$ form an $O(h^{1+\alpha})$ parallelogram and

$$|\Omega_{2,h}| = O(h^\sigma), \quad \bar{\Omega}_{2,h}\equiv\bigcup_{\tau\in\mathcal{T}_{2,h}}\bar{\tau}.$$

Remark. There are two important ingredients in an automatic mesh generation code: one is called *swap diagonal* which changes directions of some diagonal edges in order to have near parallel directions for adjacent element edges and to make as many nodes as possible to have six triangles attached; another is called *Lagrange smoothing* which iteratively relocate nodes to make each node near a mesh symmetry center (see condition (3.1) in Section 3).

Clearly, both swap diagonal and Lagrange smoothing intend to make every two adjacent triangles form an $O(h^{1+\alpha})$ parallelogram. Eventually, only small portion of elements (including boundary elements) do not satisfy this condition and then belong to $\Omega_{2,h}$, which has a small measure. Therefore, *Condition* (α, σ) is a very reasonable condition in practice and can be satisfied by most of meshes constructed from automatic mesh generation codes.

Lemma 2.1. Assume that $\mathcal{T}_h$ satisfy *Condition* (α, σ). Let $\mathcal{D}_\tau$ be a piecewise constant matrix function defined on $\mathcal{T}_h$, whose elements $\mathcal{D}_{\tau ij}$ satisfy

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

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

$$\left| \sum_{\tau \in \mathcal{T}_h} \int_\tau \nabla(u - u_I) \cdot \mathcal{D}_\tau \nabla v_h \right| \lesssim h^{1+\rho}(\|u\|_{3,\Omega} + |u|_{2,\infty,\Omega})|v|_{1,\Omega}, \quad \rho = \min(\alpha, \frac{\sigma}{2}, \frac{1}{2}). \tag{2.2}$$

Proof: Applying (2.1),

$$\sum_{\tau \in \mathcal{T}_h} \int_\tau \nabla(u - u_I) \cdot \mathcal{D}_\tau \nabla v_h = I_1 + I_2 \tag{2.3}$$

where

$$\begin{aligned} I_1 &= \sum_{\tau \in \mathcal{T}_h} \sum_{k=1}^{3} \int_{e_k} \frac{\xi_k q_k}{2 \sin \theta_k} \left\{ (\ell_{k+1}^2 - \ell_{k-1}^2) \frac{\partial^2 u}{\partial t_k^2} + 4|\tau| \frac{\partial^2 u}{\partial t_k \partial n_k} \right\} \frac{\partial v_h}{\partial t_k} \\ I_2 &= -\sum_{\tau \in \mathcal{T}_h} \int_\tau \sum_{k=1}^{3} \frac{\ell_k \xi_k}{2 \sin^2 \theta_k} \left\{ \ell_{k+1} \psi_{k-1} \frac{\partial^3 u}{\partial^2 t_{k+1} \partial t_{k-1}} + \ell_{k-1} \psi_{k+1} \frac{\partial^3 u}{\partial^2 t_{k-1} \partial t_{k+1}} \right\} \frac{\partial v_h}{\partial t_k} \end{aligned}$$

I_2 is easily estimated by

$$|I_2| \lesssim h^2 \|u\|_{3,\Omega} |v_h|_{1,\Omega}. \tag{2.4}$$

To estimate I_1, we separate all interior edges into two different groups. $\mathcal{E}_1$ is the set of edges e such that the two adjacent triangles sharing e form an $O(h^{1+\alpha})$ approximate parallelogram and $\mathcal{E}_2$ is the set of the remaining interior edges. The set of all interior edges is given by $\mathcal{E} = \mathcal{E}_1 + \mathcal{E}_2$.

For each $e \in \mathcal{E}$, we have two triangles, say τ and τ', that share e as a common edge. Denote, with respect to τ,

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

and with respect to τ',

$$\alpha'_e = \frac{\xi_{k'}}{2\sin\theta_{k'}}(\ell^2_{k'+1} - \ell^2_{k'-1}), \quad \beta'_e = \frac{\xi_{k'}}{2\sin\theta_{k'}}4|\tau'|.$$

Taking $\boldsymbol{n}$ and $\boldsymbol{t}$ to correspond to τ, we can write

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

where

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

for $j = 1, 2$, and

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

It is easy to see that, if $v_h = 0$ on $\partial\Omega$, then $I_{13} = 0$. Otherwise, we have the following estimate:

$$|I_{13}| \lesssim h^{3/2}|u|_{2,\infty,\partial\Omega}|v_h|_{1,\Omega}. \tag{2.5}$$

Setting $\boldsymbol{z} = \boldsymbol{t}$ and $\boldsymbol{z} = \boldsymbol{n}$, we estimate

$$\left| \int_e q_e \frac{\partial^2 u}{\partial \boldsymbol{t}\partial \boldsymbol{z}}\frac{\partial v_h}{\partial \boldsymbol{t}} \right| \lesssim h^{-1}|u|_{2,\infty,\Omega}\int_\tau |\nabla v_h|. \tag{2.6}$$

By definition, for $e \in \mathcal{E}_1$, $\alpha'_e = \alpha_e(1 + O(h^\alpha))$ and $\beta'_e = \beta_e(1 + O(h^\alpha))$. Therefore

$$|\alpha_e - \alpha'_e| \lesssim h^{2+\alpha}, \quad |\beta_e - \beta'_e| \lesssim h^{2+\alpha}.$$

Combining this with (2.6), we have

$$|I_{11}| \lesssim h^{1+\alpha}|u|_{2,\infty,\Omega}\int_\Omega |\nabla v_h| \lesssim h^{1+\alpha}|u|_{2,\infty,\Omega}|v_h|_{1,\Omega}. \tag{2.7}$$

Now we turn to the estimate for I_{12}. Since adjacent elements in $\Omega_{2,h}$ do not form an $O(h^{1+\alpha})$ approximate parallelogram, we simply estimate

$$|\alpha_e - \alpha'_e| \le |\alpha_e| + |\alpha'_e| \lesssim h^2, \quad |\beta_e - \beta'_e| \le |\beta_e| + |\beta'_e| \lesssim h^2.$$

Similar to (2.7), this leads to

$$|I_{12}| \lesssim h|u|_{2,\infty,\Omega}\sum_{\tau\in\mathcal{T}_{2,h}}\int_\tau |\nabla v_h| \lesssim h|u|_{0,\infty,\Omega}\|\nabla v_h\|_{0,\Omega_{2,h}}h^{\sigma/2}.$$

Combining this with (2.5) and (2.7) leads to

$$|I_1| \lesssim h^{1+\rho}|u|_{2,\infty,\Omega}|v_h|_{1,\Omega}. \tag{2.8}$$

Finally, applying (2.4) and (2.8) to (2.3), we obtain (2.2). □

3. Gradient Recovery Operators. We define $\mathcal{N}_h$ as the nodal set of a quasi-uniform triangulation $\mathcal{T}_h$. Given $z \in \mathcal{N}_h$, we consider an element patch ω around z where we choose as the origin of a local coordinates. Let (x_j, y_j) be the barycentric center of a triangle $\tau_j \subset \omega$, $j = 1, 2, \dots, m$. We require one of the following two geometric conditions be satisfied for $\alpha \geq 0$:

$$\frac{1}{m}\sum_{j=1}^{m}(x_j, y_j) = O(h^{1+\alpha})(1,1). \tag{3.1}$$

$$\sum_{j=1}^{m}\frac{|\tau_j|}{|\omega|}(x_j, y_j) = O(h^{1+\alpha})(1,1). \tag{3.2}$$

Here we use (x_j, y_j) to represent a vector in conditions (3.1) and (3.2).

Remark. Condition (α, σ) implies both conditions (3.1) and (3.2) for $z \in \Omega_{1,h} \cap \mathcal{N}_h$. Indeed, conditions (3.1) and (3.2) are trivially (with $\alpha = \infty$) satisfied by uniform meshes of the regular pattern, the union-jack pattern, and the criss-cross pattern, and allow an $O(h^{1+\alpha})$ deviation from those meshes. For example, a strongly regular mesh is an $O(h^2)$ deviation from a uniform mesh of the regular pattern. Note that the condition (3.1) depends only on relative positions of the barycentric centers of the triangles and is independent of the shapes, sizes, and numbers of those triangles.

A boundary node z usually leads to $\alpha = 0$. However, if z is an interior node with $\alpha = 0$, then there are no restrictions and we have a completely unstructured mesh around z.

Let $\mathcal{V}_h$ be the space of piecewise linear functions on the triangulation $\mathcal{T}_h$, and let $u_I \in \mathcal{V}_h$ be the linear interpolation of a given function u. We shall discuss a gradient recovery operator G_h and prove a superconvergent property between ∇u and $G_h u_I$.

The value of $G_h u_I$ is first determined at a vertex, and then linearly interpolated over the whole domain. There are three popular ways to generate $G_h u_I$ at a vertex z.

a) Weighted averaging.

$$G_h u_I(z) = \sum_{j=1}^{m}\frac{|\tau_j|}{|\omega|}\nabla u_I(x_j, y_j). \tag{3.3}$$

b) Local L^2-projection. We seek linear functions $p_l \in P_1(\omega)$ $(l = 1, 2)$, such that

$$\int_\omega [p_l(x,y) - \partial_l u_I(x,y)]^2 dxdy = \min_{q \in P_1(\omega)} \int_\omega [q(x,y) - \partial_l u_I(x,y)]^2 dxdy. \tag{3.4}$$

Then we define $G_h u_I(z) = (p_1(0,0), p_2(0,0))$.

c) Local discrete least-squares fitting proposed by Zienkiewicz-Zhu [21]. We seek linear functions $p_l \in P_1(\omega)$ $(l = 1, 2)$, such that

$$\sum_{j=1}^{m} [p_l(x_j, y_j) - \partial_l u_I(x_j, y_j)]^2 = \min_{q \in P_1(\omega)} \sum_{j=1}^{m} [q(x_j, y_j) - \partial_l u_I(x_j, y_j)]^2, \tag{3.5}$$

Then we define $G_h u_I(z) = (p_1(0,0), p_2(0,0))$. Note that c) is a discrete version of b).

The existence and uniqueness of the minimizers in b) and c) can be found in [11, Lemma 1]. Note that the characteristic equations for minimizers in b) and c) are

$$\int_\omega [p_l(x,y) - \partial_l u_I(x,y)] q(x,y) dxdy = 0, \quad \forall q \in P_1(\omega), \quad l = 1, 2, \tag{3.6}$$

$$\sum_{j=1}^{m} [p_l(x_j, y_j) - \partial_l u_I(x_j, y_j)] q(x_j, y_j) = 0, \quad \forall q \in P_1(\omega), \quad l = 1, 2, \tag{3.7}$$

respectively.

Remark. Given a finite element solution $u_h \in \mathcal{V}_h$, we can always obtain $G_h u_h(z)$ at an interior node z by the above three recovery methods. For a boundary node, there are many practical issues ivolved. Here we simply assume that the recovery exists. The worst case is to use the original finite element gradient. Since the measure of boundary elements set is $O(h)$ which can be included in $\Omega_{2,h}$, the boundary issue will not influence our analysis.

Theorem 3.1. Let ω be an element patch around a node $z \in \mathcal{N}_h$, let $u \in W^3_\infty(\omega)$, and let $G_h u_I(z)$ be produced by either the local L^2-projection or weighted averaging under condition (3.2), or by the local discrete least-squares fitting under condition (3.1). Then

$$|G_h u_I(z) - \nabla u(z)| \lesssim h^{1+\min(1,\alpha)} \|u\|_{3,\infty,\omega}.$$

Proof: a) For the weighted averaging, we have

$$
\begin{aligned}
&\sum_{j=1}^{m} \frac{|\tau_j|}{|\omega|} \partial_l u_I(x_j, y_j) - \partial_l u(0,0) \\
= &\sum_{j=1}^{m} \frac{|\tau_j|}{|\omega|} \partial_l (u_I - u)(x_j, y_j) + \sum_{j=1}^{m} \frac{|\tau_j|}{|\omega|} [\partial_l u(x_j, y_j) - \partial_l u(0,0)] \\
= &\sum_{j=1}^{m} \frac{|\tau_j|}{|\omega|} \partial_l (u_I - u)(x_j, y_j) + \nabla \partial_l u(0,0) \cdot \sum_{j=1}^{m} \frac{|\tau_j|}{|\omega|} (x_j, y_j) + R_1(u),
\end{aligned}
$$

where, by the Taylor expansion,

$$|R_1(u)| \lesssim h^2 |u|_{3,\infty,\omega}.$$

Since the barycentric center is the derivative superconvergent point for the linear interpolation, then

$$|\partial_l (u_I - u)(x_j, y_j)| \lesssim h^2 |u|_{3,\infty,\omega}, \quad j = 1, 2, \ldots, m.$$

Recall the condition (3.2), and we derive

$$|\nabla \partial_l u(0,0) \cdot \sum_{j=1}^{m} \frac{|\tau_j|}{|\omega|} (x_j, y_j)| \lesssim h^{1+\alpha} |u|_{2,\infty,\omega}.$$

Therefore,

$$|\sum_{j=1}^{m} \frac{|\tau_j|}{|\omega|} \partial_l u_I(x_j, y_j) - \partial_l u(0,0)| \lesssim h^{1+\min(1,\alpha)} \|u\|_{3,\infty,\omega}. \tag{3.8}$$

b) For the local L^2-projection, we set $q = 1$ in (3.6) to obtain

$$\sum_{j=1}^{m} |\tau_j| p_l(x_j, y_j) = \sum_{j=1}^{m} |\tau_j| \partial_l u_I(x_j, y_j).$$

Therefore,

$$p_l(0,0) - \sum_{j=1}^{m} \frac{|\tau_j|}{|\omega|} \partial_l u_I(x_j, y_j) = p_l(0,0) - \sum_{j=1}^{m} \frac{|\tau_j|}{|\omega|} p_l(x_j, y_j) = -\nabla p_l(0,0) \cdot \sum_{j=1}^{m} \frac{|\tau_j|}{|\omega|} (x_j, y_j).$$

Using (see [11, Lemma 2])

$$|\nabla p_l(0,0)| \lesssim \|u\|_{3,\infty,\omega}, \tag{3.9}$$

and condition (3.2), we obtain,

$$|p_l(0,0) - \sum_{j=1}^{m} \frac{|\tau_j|}{|\omega|} \partial_l u_I(x_j, y_j)| \lesssim h^{1+\alpha} \|u\|_{3,\infty,\omega}. \tag{3.10}$$

Combining (3.8) and (3.10), we have proved

$$|p_l(0,0) - \partial_l u(0,0)| \lesssim h^{1+\min(1,\alpha)} \|u\|_{3,\infty,\omega}. \tag{3.11}$$

c) For the local discrete least-squares fitting, we set $q = 1$ in (3.7) to obtain

$$\sum_{j=1}^{m} p_l(x_j, y_j) = \sum_{j=1}^{m} \partial_l u_I(x_j, y_j).$$

Therefore,

$$p_l(0,0) - \frac{1}{m}\sum_{j=1}^{m} \partial_l u_I(x_j, y_j) = p_l(0,0) - \frac{1}{m}\sum_{j=1}^{m} p_l(x_j, y_j) = -\frac{1}{m}\nabla p_l(0,0) \cdot \sum_{j=1}^{m}(x_j, y_j).$$

Using (3.9) and condition (3.1), we obtain

$$|p_l(0,0) - \frac{1}{m}\sum_{j=1}^{m} \partial_l u_I(x_j, y_j)| \lesssim h^{1+\alpha} \|u\|_{3,\infty,\omega}. \tag{3.12}$$

Next,

$$\begin{aligned}
& \frac{1}{m}\sum_{j=1}^{m} \partial_l u_I(x_j, y_j) - \partial_l u(0,0) \\
= \quad & \frac{1}{m}\sum_{j=1}^{m} \partial_l (u_I - u)(x_j, y_j) + \frac{1}{m}\sum_{j=1}^{m}[\partial_l u(x_j, y_j) - \partial_l u(0,0)] \\
= \quad & \frac{1}{m}\sum_{j=1}^{m} \partial_l (u_I - u)(x_j, y_j) + \frac{1}{m}\nabla \partial_l u(0,0) \cdot \sum_{j=1}^{m}(x_j, y_j) + R_2(u),
\end{aligned}$$

with $|R_2(u)| \lesssim h^2 |u|_{3,\infty,\omega}$. Therefore,

$$|\frac{1}{m}\sum_{j=1}^{m} \partial_l u_I(x_j, y_j) - \partial_l u(0,0)| \lesssim h^{1+\min(1,\alpha)} \|u\|_{3,\infty,\omega}. \tag{3.13}$$

Combining (3.12) and (3.13), we obtain (3.11) for the current case. $\square$

Theorem 3.2. The recovery operator G_h satisfies

$$G_h v(z) = \sum_{j=1}^{m} c_j \nabla v(x_j, y_j), \quad \sum_{j=1}^{m} c_j = 1,$$

in all three cases unconditionally. Furthermore, $c_j > 0$ for

a) the weighted averaging unconditionally;

b) the local L^2-projection under the condition (3.2);

c) the local discrete least-squares fitting under the condition (3.1).

Proof: The assertion is obvious for the weighted averaging case.

Choose $v = x + y$, then the minimizer $p_1 = 1$ and $p_2 = 1$ in both cases b) and c). Therefore,

$$G_h v(z) = (1,1) = \sum_{j=1}^{m} c_j \nabla(x+y) = \sum_{j=1}^{m} c_j(1,1).$$

Now we let $p_l(x,y) = a_0 + a_1 x + a_2 y$. Then for the local discrete least-squares fitting, a_i's are given by

$$\begin{pmatrix} m & \sum_j x_j & \sum_j y_j \\ \sum_j x_j & \sum_j x_j^2 & \sum_j x_j y_j \\ \sum_j y_j & \sum_j x_j y_j & \sum_j y_j^2 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} \sum_j \partial_l u_h(x_j, y_j) \\ \sum_j x_j \partial_l u_h(x_j, y_j) \\ \sum_j y_j \partial_l u_h(x_j, y_j) \end{pmatrix}, \tag{3.14}$$

Note that

$$\sum_j x_j^2 = O(h^2), \quad \sum_j x_j y_j = O(h^2), \quad \sum_j y_j^2 = O(h^2);$$

and under condition (3.1),

$$\sum_j x_j = O(h^{1+\alpha}), \quad \sum_j y_j = O(h^{1+\alpha}).$$

By scaling argument we see that

$$a_1 = O(h^{\alpha-1}), \quad a_2 = O(h^{\alpha-1}).$$

Therefore,

$$\begin{aligned} a_0 &= \frac{1}{m}\sum_j \partial_l u_h(x_j, y_j) - \frac{a_1}{m}\sum_j x_j - \frac{a_2}{m}\sum_j y_j \\ &= \sum_j c_j \partial_l u_h(x_j, y_j) \end{aligned}$$

with

$$c_j = \frac{1}{m} + O(h^{2\alpha}) > 0.$$

A similar argument shows that

$$c_j = \frac{|\tau_j|}{|\omega|} + O(h^{2\alpha}) > 0$$

for the local L^2-projection when condition (3.2) is satisfied. $\square$

Under the given condition, the recovered gradient at a vertex z is a convex combination of gradient values on the element patch surrounding z.

4. Superconvergence of the Recovery Operators. We consider the non-self-adjoint problem: find $u \in H^1(\Omega)$ such that

$$B(u,v) = \int_\Omega [(\mathcal{D}\nabla u + \boldsymbol{b}u)\cdot\nabla v + cuv] = f(v), \quad \forall v \in H^1(\Omega). \tag{4.1}$$

Here $\mathcal{D}$ is a 2×2 symmetric, positive definite matrix, and $f(\cdot)$ is a linear functional. We assume that all the coefficient functions are smooth, and the bilinear form $B(\cdot,\cdot)$ is continuous and satisfies the inf-sup condition on $H^1(\Omega)$. These conditions insure that (4.1) has a unique solution.

Let $\mathcal{V}_h \subset H^1(\Omega)$ be the C^0 linear finite element space associated with the triangulation $\mathcal{T}_h$. The finite element solution $u_h \in \mathcal{V}_h$ satisfies

$$B(u_h, v_h) = f(v_h) \quad \forall v_h \in \mathcal{V}_h. \tag{4.2}$$

To insure a unique solution for (4.2), we further assume the inf-sup condition of B be satisfied on $\mathcal{V}_h$.

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

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

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

Theorem 4.1. Let the solution of (4.1) satisfy $u \in H^3(\Omega)\cap W^2_\infty(\Omega)$, let u_h be the solution of (4.2) and $u_I \in \mathcal{V}_h$ be the linear interpolation of u. Assume that the triangulation $\mathcal{T}_h$ satisfies *Condition* (α,σ). Then

$$\|u_h - u_I\|_{1,\Omega} \lesssim h^{1+\rho}(\|u\|_{3,\Omega} + |u|_{2,\infty,\Omega}), \quad \rho = \min(\alpha, \frac{1}{2}, \frac{\sigma}{2}).$$

Proof: We begin with the identity

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

The first term I_1 is estimated using Lemma 2.1. I_2 and I_3 can be easily estimated by

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

Thus

$$|B(u - u_I, v_h)| \lesssim h^{1+\rho} \left(\|u\|_{3,\Omega} + |u|_{2,\infty,\Omega} \right) \|v_h\|_{1,\Omega}.$$

We complete the proof using the inf-sup condition in

$$\begin{aligned} \|u_h - u_I\|_{1,\Omega} &\lesssim \sup_{v_h \in \mathcal{V}_h} \frac{B(u_h - u_I, v_h)}{\|v_h\|_{1,\Omega}} = \sup_{v_h \in \mathcal{V}_h} \frac{B(u - u_I, v_h)}{\|v_h\|_{1,\Omega}} \\ &\lesssim h^{1+\rho} \left(\|u\|_{3,\Omega} + |u|_{2,\infty,\Omega} \right). \quad \square \end{aligned}$$

Theorem 4.2. Let the solution of (4.1) satisfy $u \in W^3_\infty(\Omega)$, let u_h be the solution of (4.2), and let G_h be a recovery operator defined by one of the three: a) the weighted averaging, b) the local L^2-projection, and c) the local discrete least-squares fitting. Assume that the triangulation $\mathcal{T}_h$ satisfies *Condition* (α, σ). Then

$$\|\nabla u - G_h u_h\|_{0,\Omega} \lesssim h^{1+\rho} \|u\|_{3,\infty,\Omega}.$$

Proof: We decompose

$$\nabla u - G_h u_h = (\nabla u - (\nabla u)_I) + ((\nabla u)_I - G_h u_I) + G_h(u_I - u_h), \tag{4.3}$$

where $(\nabla u)_I \in \mathcal{V}_h \times \mathcal{V}_h$ is the linear interpolation of ∇u. By the standard approximation theory,

$$\|\nabla u - (\nabla u)_I\|_{0,\Omega} \lesssim h^2 |u|_{3,\Omega}. \tag{4.4}$$

We observe that when we pick an element patch on $\mathcal{T}_{1,h}$, both conditions (3.1) and (3.2) are satisfied. Therefore, using Theorem 3.1, we have

$$\begin{aligned} \|(\nabla u)_I - G_h u_I\|_{0,\Omega_{1,h}} &\leq \left(\sum_{\tau \in \Omega_{1,h}} |\tau| \sum_{z \in \mathcal{N}_h \cap \bar{\tau}} |G_h u_I(z) - \nabla u(z)|^2 \right)^{1/2} \\ &\lesssim h^{1+\min(1,\alpha)} \|u\|_{3,\infty,\Omega} |\Omega_{1,h}|^{1/2} \lesssim h^{1+\min(1,\alpha)} \|u\|_{3,\infty,\Omega}. \end{aligned} \tag{4.5}$$

On the other hand,

$$\|(\nabla u)_I - G_h u_I\|_{0,\Omega_{2,h}} \lesssim h \|u\|_{3,\infty,\Omega} |\Omega_{2,h}|^{1/2} \lesssim h^{1+\sigma/2} \|u\|_{3,\infty,\Omega} \tag{4.6}$$

by *Condition* (α, σ). Combining (4.5) with (4.6), we have

$$\|(\nabla u)_I - G_h u_I\|_{0,\Omega} \lesssim h^{1+\min(\alpha,\sigma/2)} \|u\|_{3,\infty,\Omega}. \tag{4.7}$$

Similar as in (4.5), we have, by using the fact proved in Theorem 3.2, that $G_h v(z)$ is a convex combination of $\nabla v|_{\tau_z}$s,

$$\begin{aligned}
\|G_h(u_I - u_h)\|_{0,\Omega_{1,h}} &\leq \left(\sum_{\tau \in \mathcal{T}_{1,h}} |\tau| \sum_{z \in \mathcal{N}_h \cap \bar{\tau}} |G_h(u_I - u_h)(z)|^2 \right)^{1/2} \\
&\lesssim \left(\sum_{\tau \in \mathcal{T}_{1,h}} |\tau| \|\nabla(u_I - u_h)|_\tau|^2 \right)^{1/2} = \|\nabla(u_I - u_h)\|_{0,\Omega_{1,h}} \\
&\lesssim h^{1+\rho} \|u\|_{3,\infty,\Omega},
\end{aligned} \tag{4.8}$$

by Theorem 4.1. In addition,

$$\begin{aligned}
\|G_h(u_I - u_h)\|_{0,\Omega_{2,h}} &\leq \left(\sum_{\tau \in \mathcal{T}_{2,h}} |\tau| \sum_{z \in \mathcal{N}_h \cap \bar{\tau}} |G_h(u_I - u_h)(z)|^2 \right)^{1/2} \\
&\lesssim h \|u\|_{3,\infty,\Omega} \left(\sum_{\tau \in \mathcal{T}_{2,h}} |\tau| \right)^{1/2} \\
&\lesssim h^{1+\sigma/2} \|u\|_{3,\infty,\Omega}.
\end{aligned} \tag{4.9}$$

Combining (4.8) and (4.9) yields

$$\|G_h(u_I - u_h)\|_{0,\Omega} \lesssim h^{1+\rho} \|u\|_{3,\infty,\Omega}. \tag{4.10}$$

The conclusion follows by applying (4.4), (4.7), and (4.10) to the right hand side of (4.3).
□

5. Asymptotic Exactness of the Recovery Type Error Estimators. With preparation in the previous sections, it is now straightforward to prove the asymptotic exactness of error estimators based on the recovery operator G_h. The global error estimator is naturally defined by

$$\eta_h = \|G_h u_h - \nabla u_h\|_{0,\Omega}. \tag{5.1}$$

Theorem 5.1. Assume the hypotheses of Theorem 4.2. Furthermore, assume that there exists a constant $c(u) > 0$ such that

$$\|\nabla(u - u_h)\| \geq c(u)h. \tag{5.2}$$

then

$$\left| \frac{\eta_h}{\|\nabla(u - u_h)\|_{0,\Omega}} - 1 \right| \lesssim h^{\rho}.$$

Proof: By Theorem 4.2 and hypothesis (5.2), we have

$$\left| \frac{\eta_h}{\|\nabla(u - u_h)\|_{0,\Omega}} - 1 \right| \leq \frac{\|G_h u_h - \nabla u\|_{0,\Omega}}{\|\nabla(u - u_h)\|_{0,\Omega}} \lesssim \frac{h^{1+\rho}\|u\|_{3,\infty,\Omega}}{c(u)h} \lesssim h^{\rho}.$$

□

We see that the error estimator (5.1) based on the gradient recovery operator is asymptotically exact under *Condition* (α, σ). As we mentioned above, this condition is not a very restrictive condition in practice. An automatic mesh generator usually produces some grids which are mildly structured. In practice, a completely unstructured mesh is seldomly seen. Our analysis explains in part the good performance of the ZZ error estimator based on the local discrete least-squares fitting for general grids.

References

[1] M. Ainsworth and J.T. Oden, *A Posteriori Error Estimation in Finite Element Analysis*, Wiley Interscience, New York, 2000.

[2] I. Babuška and W.C. Rheinboldt, A Posteriori Error Error Estimates for the Finite Element Method, Internat. J. Numer. Methods Engrg., 12 (1978), pp.1597-1615.

[3] I. Babuška and T. Strouboulis, *The Finite Element Method and its Reliability*, Oxford University Press, London, 2001.

[4] R.E. Bank and A. Weiser, Some *a posteriori* error estimators for elliptic partial differential equations, Math. Comp., 44 (1985), pp.283-301.

[5] R.E. Bank and J. Xu, Asymptotically exact *a posteriori* error estimators, Part I: Grid with superconvergence, preprint.

[6] R.E. Bank and J. Xu, Asymptotically exact *a posteriori* error estimators, Part II: General Unstructured Grids, preprint.

[7] C. Carstensen and S. Bartels, Each averaging technique yields reliable *a posteriori* error control in FEM on unstructured grids. Part I: Low order conforming, nonconforming, and mixed FEM, Math. Comp. 71 (2002), pp.945-969.

[8] C.M. Chen and Y.Q. Huang, *High Accuracy Theory of Finite Element Methods.* Hunan Science Press, Hunan, China, 1995 (in Chinese).

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

[10] M. Křížek, P. Neittaanmäki, and R. Stenberg (Eds.), *Finite Element Methods: Superconvergence, Post-processing, and A Posteriori Estimates*, Lecture Notes in Pure and Applied Mathematics Series, Vol.196, Marcel Dekker, New York, 1997.

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

[12] Q. Lin and N. Yan, *Construction and Analysis of High Efficient Finite Elements* (in Chinese), Hebei University Press, P.R. China, 1996.

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

[14] L.B. Wahlbin, *Superconvergence in Galerkin Finite Element Methods*, Lecture Notes in Mathematics, Vol.1605, Springer, Berlin, 1995.

[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, Ultraconvergence of the patch recovery technique II, Math. Comp., 69 (2000), pp.141-158.

[17] Z. Zhang and H.D. Victory Jr., Mathematical Analysis of Zienkiewicz-Zhu's derivative patch recovery techniques, Numerical Methods for Partial Differential Equations, 12 (1996), pp.507-524.

[18] Q.D. Zhu and Q. Lin, *Superconvergence Theory of the Finite Element Method*, Hunan Science Press, China, 1989 (in Chinese).

[19] J.Z. Zhu and Z. Zhang, The relationship of some *a posteriori* error estimators, Comput. Methods Appl. Mech. Engrg., 176 (1999), pp.463-475.

[20] O.C. Zienkiewicz and J.Z. Zhu, A simple error estimator and adaptive procedure for practical engineering analysis, Internat. J. Numer. Methods Engrg., 24 (1987), pp.337-357.

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

Center for Computational Mathematics and Applications, Department of Mathematics, The Pennsylvania State University, University Park, PA 16802.

Email address: xu@math.psu.edu.

Department of Mathematics, Wayne State University, Detroit, MI 48202.

Email address: zzhang@math.wayne.edu.