<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Superconvergence for the gradient of finite element approximations by L2-projections</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Heimsund, Bjorn-Ove; Tai, Xue Cheng; Wang, Junping</td>
</tr>
<tr>
<td><strong>Date</strong></td>
<td>2002</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10220/6049">http://hdl.handle.net/10220/6049</a></td>
</tr>
</tbody>
</table>
SUPERCONVERGENCE FOR THE GRADIENT OF FINITE ELEMENT APPROXIMATIONS BY $L^2$ PROJECTIONS

BJØRN-OVE HEIMSUND,† XUE-CHENG TAI,† AND JUNPING WANG‡

Abstract. A gradient recovery technique is proposed and analyzed for finite element solutions which provides new gradient approximations with high order of accuracy. The recovery technique is based on the method of least-squares surface fitting in a finite-dimensional space corresponding to a coarse mesh. It is proved that the recovered gradient has a high order of superconvergence for appropriately chosen surface fitting spaces. The recovery technique is robust, efficient, and applicable to a wide class of problems such as the Stokes and elasticity equations.

Key words. finite element methods, superconvergence, error estimates, adaptive refinement

AMS subject classifications. 65N30, 65N15, 65F10

PII. S003614290037410X

1. Introduction. It has been known for a long time that finite element solutions of partial differential equations can have superconvergence in some subregions of the domain [26, 23, 8, 1]. Superconvergence is a phenomenon that the numerical solution converges to the exact solution at a rate higher than the optimal order error estimate. To exploit superconvergence in the finite element method, several methods have been proposed in the literature in the last 30 years. The method of local averaging has turned out to be a common and useful technique in the study of superconvergence in most of the existing results; see, for example, [23, 8, 1, 7, 35, 19, 18, 20, 17, 21, 25, 26, 10, 13] and the references therein. In theory, all the existing results require the underlying finite element mesh to have some special properties such as uniformity [23, 7, 21], local point symmetry [25, 26], local translation invariance [1, 26], or orthogonality (e.g., rectangular partition) [8, 10, 13, 19, 18, 20, 28, 34].

The Zienkiewicz and Zhu (ZZ) method [32, 33] is a procedure which postprocesses the gradient of the finite element solution by using a discrete least-squares fitting on a local patch with high order polynomials. Due to its high efficiency and robustness, the ZZ postprocessing has been widely used for mesh adaptivity and error control in finite element methods [32, 33, 5, 6]. For appropriately chosen discrete norms, this procedure has been computationally justified to yield some superconvergence for the gradient. If the underlying finite element partition is uniform or rectangular, one can provide a theoretical proof for the ZZ method [31, 34, 30] by using some existing superconvergent estimates [8, 23, 17, 35, 18].

Our objective of this paper is twofold. First, we modify the ZZ method by applying a global least-squares fitting to the gradient of the finite element approximation. The surface fitting space consists of continuous or discontinuous piecewise polynomials of high order on a coarse partition. Second, we provide a theoretical analysis for
the modified ZZ (MZZ) method by establishing a superconvergence estimate for the recovered gradient/flux on general quasi-uniform meshes. To the authors’ knowledge, our result is the first that gives a theoretical proof for the superconvergence of the ZZ method with some modifications under general assumptions for the finite element partition. The essential idea behind the approach is the use of a coarser mesh and a higher order of polynomials which can be translated to the method of “long” and “accurate” finite difference quotients. The same idea has been applied in [15] to yield asymptotically exact a posteriori estimators for the pointwise gradient error.

Our presentation follows a framework established in Wang [27] (see also [29]), where the least-squares surface fitting (the projection method) was applied to the finite element solution \(u_h\) in order to produce a new and better approximation for the original unknown function \(u = u(x)\) and its gradient \(\nabla u\). The approach of this paper is different in that the projection will be applied directly to the numerical gradient \(\nabla u_h\) in order to provide a superconvergent numerical solution for \(\nabla u\). Like all the existing results in superconvergence, our results are based on a certain regularity assumption for the exact solution of the underlying model problem.

For simplicity of discussion, our superconvergence result will be presented only for Dirichlet boundary value problems. The results can be extended to Neumann and Robin boundary conditions without any difficulty.

The paper is organized as follows. In section 2, we introduce a model problem for which the required regularity condition is satisfied. In section 3, we present an extension of the ZZ method by using a global least-squares fitting in a high order finite element space corresponding to a coarse mesh. Some error estimates for the new gradient approximation will be derived in section 4. In sections 5 and 6 we apply the error estimate to show that the projected gradient is superconvergent if the fitting space is properly chosen. Section 7 applies the gradient recovery scheme to mesh adaption, and section 8 gives numerical results and comparison for various adaptive schemes.

2. A model problem. To illustrate the idea, we consider boundary value problems for the second order elliptic equation. Let \(\Omega\) be an open bounded domain in \(\mathbb{R}^d, d = 2, 3\). Denote by \(x = (x_1, \ldots, x_d)\) the points in \(\Omega\). Let \(\partial_i = \frac{\partial}{\partial x_i}\) be the partial derivative operator in the direction of \(x_i, i = 1, \ldots, d\). The Dirichlet boundary value problem seeks a function \(u = u(x)\) such that \(u(x) = g(x)\) for any \(x \in \partial \Omega\) and

\[
\sum_{i,j=1}^{d} \partial_j (a_{ij} \partial_i u) + \sum_{i=1}^{d} b_i \partial_i u + cu = f \quad \text{in } \Omega, \tag{1}
\]

where \(a = (a_{ij})_{i,j=1}^d\) is the coefficient tensor which is symmetric, bounded, and uniformly positive definite in the domain \(\Omega\) with measurable entries \(a_{ij} = a_{ij}(x)\). The other coefficients \(b = (b_i(x))_{i=1}^d\) and \(c = c(x)\) are assumed to ensure a uniqueness of solutions for (1).

Standard notations for Sobolev spaces and norms are adopted in this paper. For an \(s \geq 0\), which may not be an integer, and a given domain \(\Omega\), \(H^s(\Omega)\) denotes the Sobolev space with norm \(\| \cdot \|_s\) as defined in [14]. The space \(H^0_0(\Omega)\) is a closed subspace of \(H^s(\Omega)\) that is the closure of \(C^0_0(\Omega)\) (the set of compact-supported \(C^0\) functions) in the norm of \(H^s(\Omega)\). For \(s < 0\), \(H^s(\Omega)\) is defined to be the dual space of \(H^{-s}(\Omega)\); see [14] for details. The Sobolev space \(H^1(\Omega)\) coincides with \(L^2(\Omega)\), in which case the norm and inner product are denoted by \(\| \cdot \|\) and \((\cdot, \cdot)\), respectively.
Let
\[ a(u, v) = \sum_{i,j=1}^{d} \int_{\Omega} a_{ij} \partial_i u \partial_j v \, dx + \sum_{i=1}^{d} \int_{\Omega} b_i \partial_i u v \, dx + \int_{\Omega} cu v \, dx \]
be a bilinear form defined in \( H^1(\Omega) \times H^1(\Omega) \). A weak form for the problem (1) seeks a function \( u \in H^1(\Omega) \) such that \( u = g \) on \( \partial \Omega \) and
\[
(2) \quad a(u, v) = (f, v) \quad \forall v \in H^1_0(\Omega).
\]
Here we have assumed that the boundary data \( g \in H^{1/2}(\partial \Omega) \).

Let \( s \geq 1 \) be a positive real number. Assume that the dual problem of (2) has \( H^s \) regularity in the sense that, for any given \( f \in H^{s-2}(\Omega) \), the problem
\[
(3) \quad a(v, w) = (f, v) \quad \forall v \in H^1(\Omega)
\]
has a unique solution \( w \in H^1(\Omega) \cap H^s(\Omega) \) such that
\[
\| w \|_s + \left\| \frac{\partial w}{\partial n_a} \right\|_{s-rac{3}{2}, \partial \Omega} \leq C \| f \|_{s-2},
\]
where \( n \) is the unit outward normal vector of \( \partial \Omega \) and \( \frac{\partial w}{\partial n_a} = (a \nabla w) \cdot n \) denotes the normal component of the flux variable on the boundary \( \partial \Omega \) for the dual solution \( w \).

It is well known that the bilinear form \( a(\cdot, \cdot) \) is bounded in \( H^1(\Omega) \). In other words, there exists a constant \( C \) such that
\[
|a(u, v)| \leq C \| u \|_1 \| v \|_1 \quad \forall u, v \in H^1(\Omega).
\]

The finite element solution of (2) is a function \( u_h = u_h(x) \) from a finite element space \( S_h \subset H^1(\Omega) \) associated with a prescribed finite element partition \( \Omega_h \) such that \( u_h(x) = g_h(x) \) for all \( x \in \partial \Omega \) and
\[
(4) \quad a(u_h, v) = (f, v) \quad \forall v \in S_h^0.
\]
Here \( S_h^0 = H^1_0(\Omega) \cap S_h \) and \( g_h \) is a certain approximation of the Dirichlet boundary data \( g \). Let \( \Lambda_h \) be the restriction of the finite element space \( S_h \) on the boundary of \( \Omega \). For simplicity, we shall deal with polygonal or polyhedral domain \( \Omega \) so that the boundary \( \partial \Omega \) is exactly represented by the finite element partition \( \Omega_h \). Among many possibilities, we are particularly interested in two cases for the approximate boundary data:
- \( g_h \) is the standard nodal interpolation of \( g \) in \( \Lambda_h \) for sufficiently smooth \( g \).
- \( g_h \) is the \( L^2 \) projection of \( g \) in \( \Lambda_h \).

We recall that the \( L^2 \) projection of \( g \) in \( \Lambda_h \) is given by solving the following system of linear equations:
\[
(5) \quad \langle g_h, v \rangle = \langle g, v \rangle \quad \forall v \in \Lambda_h,
\]
where \( \langle \cdot, \cdot \rangle \) is the standard \( L^2 \)-inner product on \( \partial \Omega \).

Assume that \( S_h \) consists of continuous piecewise polynomials of order \( k \geq 1 \). Let \( h \) be the mesh parameter for the finite element partition \( \Omega_h \). The finite element space \( S_h \) is assumed to have the following approximation property:
\[
\inf_{v \in S_h} (\| w - v \|_1 + h \| w - v \|) \leq C h^m \| w \|_m \quad \forall w \in H^m(\Omega)
\]
for any \( 0 \leq m \leq k + 1 \).
3. Gradient recovery by projections. To obtain an approximate gradient and flux with superconvergence, we consider a new finite-dimensional space $\mathcal{L}_\tau$ with parameter $\tau \gg h$ and a higher order approximation property than $S_h$ [15]. The functions in $\mathcal{L}_\tau$ are vector valued and will be employed to approximate the exact gradient/flux variable $q = a \nabla u$. In practice, the mesh parameter $\tau$ is proportional to $h^\alpha$ for some $\alpha \in (0, 1)$ in order to obtain a superconvergent approximation from the projection space $\mathcal{L}_\tau$. Details can be found from Wang [27].

For simplicity, assume that $\mathcal{L}_\tau$ is a finite element space associated with another finite element partition $\Omega_\tau$ and consists of piecewise polynomials of order $r \geq 0$. The finite element space $\mathcal{L}_\tau$ is required to satisfy the following properties:

- Inverse property.
  \[ \|v_\tau\|_{H^m(K)} \leq C \tau^{-m} \|v_\tau\|_{L^2(K)} \quad \forall v_\tau \in \mathcal{L}_\tau, \quad \forall K \in \Omega_\tau \]
  for all nonnegative integer $m \geq 0$.

- Approximation property.
  \[ \inf_{v_\tau \in \mathcal{L}_\tau} \|v - v_\tau\|_0 \leq C \tau^m \|v\|_m \quad \forall v \in [H^m(\Omega)]^d, \quad 0 \leq m \leq r + 1. \]

- Smoothness property.
  \[ \mathcal{L}_\tau \subset [H^{s-1}(\Omega)]^d. \]

Here $s \geq 1$ is associated with the regularity of the dual problem as indicated in (3).

We emphasize that the space $\mathcal{L}_\tau$ can be replaced by other finite-dimensional spaces as trigonometric functions, B-splines, and any special functions if the domain is of special type. In such cases the approximation property and the inverse inequality will be different, and the forthcoming analysis must be modified accordingly.

3.1. Recovery based on a mixed formulation. Our objective here is to provide a very accurate approximation for the flux variable $q$ by using the finite element solution $u_h$. The relation between the flux $q = q(x)$ and the original function $u = u(x)$ can be rewritten as follows:

\[ a^{-1}q = \nabla u. \]

Let

\[ H(\text{div}; \Omega) = \{ v : v \in [L^2(\Omega)]^d, \nabla \cdot v \in L^2(\Omega) \} \]

be equipped with the norm

\[ \|v\|_H = \left( \|v\|^2 + \|\nabla \cdot v\|^2 \right)^{1/2}. \]

By testing (9) against any $v \in H(\text{div}; \Omega)$ we arrive at

\[ (a^{-1}q, v) = -(u, \nabla \cdot v) + \langle g, v \cdot n \rangle, \]

where we have employed the integration by parts to the term on the right-hand side.

Equation (10) can be employed to provide a new flux/gradient recovery $\tilde{q}_\tau$ defined as follows:

\[ (a^{-1}\tilde{q}_\tau, v) = -(u_h, \nabla \cdot v) + \langle g, v \cdot n \rangle \quad \forall v \in \mathcal{L}_\tau. \]
The new flux approximation $\hat{q}_r$ will be denoted by

$$\hat{q}_r = \hat{Q}_r q_h.$$  

It is clear that $\hat{Q}_r$ can be regarded as a linear operator onto the fitting space $L_r$.

Since (11) was obtained by using test functions in the space $H(\text{div};\Omega)$, the fitting space $L_r$ has to be constructed as a finite element subspace of $H(\text{div};\Omega)$. In practical computation, the standard mixed finite element spaces of Raviart and Thomas [24], Brezzi et al. [2, 3], Brezzi, Douglas, and Marini [4], and Douglas and Wang [11] can be employed to accomplish the goal. Of course, one can also use continuous finite element spaces in the place of $L_r$. The well-known inf-sup condition is no longer an issue in this procedure because the flux computed is based on a Galerkin approximation of the scalar variable.

### 3.2. Recovery based on $L^2$ projections.

Let $Q_r$ be the weighted $L^2$ projection onto the fitting space $L_r$ with respect to the weighted inner product $(a^{-1} \cdot, \cdot)$. More precisely, for any $v \in [L^2(\Omega)]^d$, the projection $Q_r v$ is a function in $L_r$ such that

$$\langle a^{-1} Q_r v, \phi \rangle = \langle a^{-1} v, \phi \rangle \quad \forall \phi \in L_r.$$  

It follows from the definition of the Galerkin approximation $u_h$ that $q_h = a \nabla u_h$ is an approximate solution of the exact flux variable $q$. In addition, it is not hard to derive the following error estimate:

$$\|q - q_h\| \leq C\|u - u_h\|$$  

for some constant $C$.

With the $L^2$-projection operator $Q_r$, we can provide a new flux approximation given as follows:

$$q = a \nabla u = Q_r q_h.$$  

From the definition of $Q_r$, we see that the new flux approximation $Q_r q_h$ satisfies the following system of equations:

$$\langle a^{-1} Q_r q_h, \phi \rangle = \langle \nabla u_h, \phi \rangle \quad \forall \phi \in L_r.$$  

When the fitting space $L_r$ consists of discontinuous piecewise polynomials of order $k$ on each element of $\Omega_r$, our flux recovery method is closely related to the ZZ [32, 33] patch recovery technique. The difference lies on the selection of the fitting space $L_r$ and the way that the projection was defined. The ZZ method uses a discrete version of the $L^2$-inner product, and the fitting space is based on a patch of elements from the original finite element partition $\Omega_h$.

In practical computation, the recovery space contains polynomials of higher order than the original finite element space. In other words, the value of the parameter $r$ is normally larger than $k$.

### 4. Error estimates.

The objective of this section is to analyze the approximation formulas (13) and (11). The accuracy of the approximations is given in Theorem 4.1 for the case when the boundary data $g$ is approximated by its $L^2$ projection in $\Lambda_h$. In case that $g_h$ is the nodal point interpolation or other approximations satisfying (29), the corresponding superconvergence estimate is given in Theorem 4.2.
For simplicity of notation, we use the following element-wise Sobolev norms:

$$
\|v\|_{m,h} = \left( \sum_{K \in \Omega_h} |v|^2_{H^m(K)} \right)^{1/2}, \quad \|v\|_{m,\tau} = \left( \sum_{K \in \Omega_{\tau}} |v|^2_{H^m(K)} \right)^{1/2},
$$

where

$$
|v|_{H^m(K)} = \left( \sum_{K \in \Omega_h} \sum_{|\alpha|=m} \int_K |D^\alpha v|^2 dx \right)^{1/2}
$$
is the seminorm of $v \in H^m(K)$. Here $\alpha = (\alpha_1, \ldots, \alpha_d), \alpha_i \geq 0$ is a multi-index and $D^\alpha v = \partial_1^{\alpha_1} \cdots \partial_d^{\alpha_d} v$. Thus, our function $v$ above needs only to be in the Sobolev space $H^m(K)$ over each element $K$ from $\Omega_h$ or $\Omega_{\tau}$ in order to guarantee that the norms exist. If (6) and (7) are satisfied, then it is not hard to prove the following estimate:

$$
(15) \quad \|v - Q_{\tau} v\|_0 \leq C\tau^m \|v\|_{m,\tau}, \quad 0 \leq m \leq r + 1.
$$

4.1. Projection space $L_{\tau}$ of class $H(\text{div}; \Omega)$. Let us analyze the approximation schemes (11) and (13) by assuming that the projection space $L_{\tau}$ is of class $H(\text{div}; \Omega)$. The following theorem is concerned with the case when the Dirichlet boundary data is approximated by $L^2$ projections.

**Theorem 4.1.** Let $u$ be the exact solution of (2) and $u_h$ be its finite element approximation given by (4). Let $q = a \nabla u$ be the flux/gradient with the obvious approximation $q_h = a \nabla u_h$ and let $\mathcal{G}_{\tau}$ be a postprocessing operator given by either $Q_{\tau}$ or $\tilde{Q}_{\tau}$ as in the previous section. Assume that the approximate boundary value $g_h$ is taken to be the $L^2(\partial \Omega)$ projection of $g$ and (3) and (8) are valid for an $s \in [1, k + 1]$. Assume that the fitting (projection) space $L_{\tau}$ is constructed so that $L_{\tau} \subset H(\text{div}; \Omega)$. Then

$$
(16) \quad \|q - \mathcal{G}_{\tau} q_h\|_0 \leq C\tau^{r+1} \|q\|_{r+1,\tau} + C(h\tau^{-1})^{s-1} \|u - u_h\|_1.
$$

**Proof.** First, we provide a proof for $\mathcal{G}_{\tau} = \tilde{Q}_{\tau}$. To this end, we observe that

$$
(17) \quad \|\tilde{Q}_{\tau} q_h - Q_{\tau} q\|_0 \leq C \sup_{\phi \in L_{\tau}, \|\phi\|_{L_{\tau}} = 1} (a^{-1} \tilde{Q}_{\tau} q_h - a^{-1} Q_{\tau} q, \phi).
$$

For a given $\phi \in L_{\tau}$, it is true that

$$
(a^{-1} \tilde{Q}_{\tau} q_h - a^{-1} Q_{\tau} q, \phi) = (a^{-1} \tilde{Q}_{\tau} (a \nabla u_h) - a^{-1} Q_{\tau} (a \nabla u), \phi)
$$

$$
= -(u_h, \nabla \cdot \phi) + \langle g, \phi \cdot n \rangle - (\nabla u, \phi) = (u - u_h, \nabla \cdot \phi).
$$

Define $w \in H^1_0(\Omega)$ to be the solution of

$$
(18) \quad a(v, w) = (v, \nabla \cdot \phi) \quad \forall v \in H^1_0(\Omega).
$$

Applying the theory of distributions, it can be proved that

$$
(a(v, w) - \left< \frac{\partial w}{\partial n_{\alpha}}, v \right>) = (\nabla \cdot \phi, v) \quad \forall v \in H^1(\Omega).
$$
Using (2), (4), (5), and the above equation, it is easy to see that, for any \(v \in S_h^0\) and \(\xi \in \Lambda_h\),

\[
(u - u_h, \nabla \phi) = a(u - u_h, w) - \left< \frac{\partial w}{\partial n_a}, u - u_h \right>
= a(u - u_h, w - v) - \left< \frac{\partial w}{\partial n_a} - \xi, g - g_h \right>.
\]

It follows that

\[
|(u - u_h, \nabla \phi)| \leq C\|u - u_h\|_1\|w - v\|_1 + \|\frac{\partial w}{\partial n_a} - \xi\|_{\frac{1}{2}, \partial \Omega} \|u - u_h\|_{\frac{1}{2}, \partial \Omega}.
\]

Using the trace inequality

\[
\|v\|_{\frac{1}{2}, \partial \Omega} \leq C\|v\|_1 \quad \forall v \in H^1(\Omega)
\]

and the interpolation estimates

\[
\begin{align}
\inf_{v \in S_h} \|w - v\|_1 & \leq Ch^{s-1}\|w\|_s, \\
\inf_{\xi \in \Lambda_h} \left\| \frac{\partial w}{\partial n_a} - \xi \right\|_{\frac{1}{2}, \partial \Omega} & \leq Ch^{s-1}\|w\|_s,
\end{align}
\]

we obtain

\[
|(u - u_h, \nabla \phi)| \leq Ch^{s-1}\|u - u_h\|_1\|w\|_s
\]

for an \(s \in [1, k + 1]\). Next, we use the \(H^s\) regularity assumption (3) to obtain

\[
|(u - u_h, \nabla \phi)| \leq Ch^{s-1}\|u - u_h\|_1\|\nabla \phi\|_{s-2} \leq Ch^{s-1}r^{1-s}\|u_h - u\|_1\|
\]

where we have also used the inverse property (6) in the last inequality. Collecting all the estimates we obtain

\[
\|Q(T q_h - Q_T q)\|_0 \leq Ch^{s-1}r^{1-s}\|u_h - u\|_1,
\]

which, together with (15), gives the desired error estimate for \(G_T = \tilde{Q}_T\).

To analyze the case \(G_T = Q_T\), it suffices to estimate \(\|Q_T(q - q_h)\|_0\). Since

\[
\|Q_T(q - q_h)\|_0 \leq C\sup_{\phi \in L_r, \|
\phi\|_0 = 1} (a^{-1}Q_T(q - q_h), \phi)
\]

and

\[
(a^{-1}Q_T(q - q_h), \phi) = (\nabla(u - u_h), \phi),
\]

then it is sufficient to estimate \(\|\nabla(u - u_h), \phi\|\). Recall that, by assumption, we have \(L_r \subset H(\text{div}; \Omega)\). Thus, it follows from the integration by parts that

\[
(\nabla(u_h - u), \phi) = (u - u_h, \nabla \phi) + (u_h - u, \phi \cdot n).
\]

Let \(w \in H^1_0(\Omega)\) be defined as the solution of the following problem:

\[
a(v, w) = -\langle \nabla \phi, v \rangle \quad \forall v \in H^1_0(\Omega).
\]
It follows from (2), (4), and (5) that, for any \( v \in S^0_h \) and \( \xi \in \Lambda_h \),

\[
\begin{align*}
(\nabla(u_h - u), \phi) &= (u - u_h, \nabla \cdot \phi) + (u_h - u, \phi \cdot n) \\
&= a(u - u_h, w) + \left( \frac{\partial w}{\partial n_a} u_h - u \right) + (\phi \cdot n, u_h - u) \\
&= a(u - u_h, w - v) + \left( \frac{\partial w}{\partial n_a} + \phi \cdot n - \xi, u_h - u \right).
\end{align*}
\]

(27)

Using the standard approximation property of \( \Lambda_h \) and the trace inequality in Sobolev spaces as in (20)–(21), we obtain

\[
\inf_{\xi \in S_h} \left\| \frac{\partial w}{\partial n_a} + \phi \cdot n - \xi \right\|_{-\frac{1}{2}, \partial \Omega} \leq Ch^{s-1} \tau_{s-1} \| \phi \|_0.
\]

It is also not hard to see that

\[
\inf_{v \in S^0_h} |a(u - u_h, w - v)| \leq Ch^{s-1} \tau_{s-1} \| u - u_h \|_1 \| \phi \|_0.
\]

Substituting the above two estimates into (27), we obtain

\[
| (\nabla(u_h - u), \phi) | \leq Ch^{s-1} \tau_{s-1} \| u - u_h \|_1 \| \phi \|_0,
\]

which implies that

\[
(28) \quad \| Q_r(q - q_h) \|_0 \leq Ch^{s-1} \tau_{s-1} \| u - u_h \|_1.
\]

This completes the proof of the theorem.

If the exact solution is sufficiently smooth, then we have from the estimate (16) that

\[
\| q - G_r q_h \|_0 \leq C(u, q) \left( \tau^{r+1} + \tau^{k+s-1} \right).
\]

Assume that the model problem has the \( H^{k+1} \) regularity (i.e., \( s = k + 1 \)). Thus,

\[
\| q - G_r q_h \|_0 \leq C(u, q) \left( \tau^{r+1} + \tau^{k} h^{2k} \right).
\]

By choosing \( \tau = h^\alpha \), the above estimate becomes

\[
\| q - G_r q_h \|_0 \leq C(u, q) \left( h^{\alpha(r+1)} + h^{(2-\alpha)k} \right),
\]

which is optimized when

\[
\alpha(r+1) = (2-\alpha)k \iff \alpha = \frac{2k}{r+k+1}.
\]

The corresponding error estimate is given by

\[
\| q - G_r q_h \|_0 \leq C(u, q) h^{\frac{2k(r+1)}{r+k+1}}.
\]

With \( k = 3 \) and \( r = 3 \), the above estimate implies an accuracy of order \( h^{\frac{2}{5}} \) which is much better than the optimal order \( h^2 \).

In practical computation, the Dirichlet boundary data is often approximated by a scheme different from the \( L^2 \) projection. Thus, the estimate in Theorem 4.1 is
no longer valid for such problems. Our next goal of this section is to derive some superconvergence for general approximation schemes of the Dirichlet boundary data.

**Theorem 4.2.** Assume that (3) and (8) are valid for an \( s \in [3/2, k + 1] \) and that \( g_h \) is an approximation of the Dirichlet data \( g \) on the boundary such that

\[
\|g - g_h\|_{0, \partial \Omega} \leq C h^{k+1}\|g\|_{k+1, \partial \Omega}.
\]

Assume that the projection space \( \mathcal{L}_\tau \subset H(\text{div}; \Omega) \). Then there exists a constant \( C \) such that

\[
\|q - \tilde{Q}_\tau q_h\|_0 \leq C r^{r+1}\|q\|_{r+1, \tau} + Ch^{r-1}\tau^{1-s}\|u - u_h\|_1 + Ch^{k+1}\tau^{1/2}\|g\|_{k+1, \partial \Omega}.
\]

**Proof.** The proof is similar to that of Theorem 4.1. The only modification is on the treatment of \((u - u_h, \nabla \cdot \phi)\). To this end, we observe that

\[
(u - u_h, \nabla \cdot \phi) = a(u - u_h, w - v) - \left( \frac{\partial w}{\partial n}, g - g_h \right).
\]

Thus,

\[
\|(u - u_h, \nabla \cdot \phi)\| \leq Ch^{s-1}\|u - u_h\|_1\|w\|_s + \left\| \frac{\partial w}{\partial n} \right\| \|g - g_h\|_{0, \partial \Omega}
\]

\[
\leq Ch^{s-1}\|u - u_h\|_1\|\nabla \cdot \phi\|_{s-2} + Ch^{k+1}\|u\|_2\|g\|_{k+1, \partial \Omega}
\]

\[
\leq Ch^{s-1}\tau^{1-s}\|u_h - u\|_1\|\phi\|_0 + Ch^{k+1}\tau^{1/2}\|g\|_{k+1, \partial \Omega}\|\phi\|_0.
\]

The rest of the proof is similar to that of Theorem 4.1 and is omitted. \( \square \)

Theorem 4.2 shows that if the Dirichlet boundary data is not approximated by the \( L^2 \) projection, then the superconvergence estimate for the recovered flux/gradient approximation will suffer. In fact, our estimate of Theorem 4.2 ensures only a superconvergence of order \( O(h^{k+1}) \) for sufficiently smooth solution \( u \) and the projection space \( \mathcal{L}_\tau \).

**4.2. Discontinuous projection space \( \mathcal{L}_\tau \).** The flux approximation scheme (13) or (14) is well defined for discontinuous projection space \( \mathcal{L}_\tau \). When discontinuous finite elements are employed in the projection method, the computation of the recovered flux/gradient can be implemented locally on each element \( K \in \Omega_\tau \), which results in a great saving of computer time and efficiency. However, due to the use of integration by parts in (25), the superconvergence established in Theorems 4.1 and 4.2 is no longer applicable to discontinuous projection space. Our objective of this section is to provide a superconvergent theory for the approximation scheme (13) when \( \mathcal{L}_\tau \) contains discontinuous finite element functions.

Let \( K \) be any element from the partition \( \Omega_\tau \). It is not hard to show that there exists a constant \( C \) independent of \( K \) and \( v \) such that

\[
\int_{\partial K} v^2 ds \leq C \left( \tau^{-1} + \epsilon^{-1} \right) \int_K v^2 dx + \epsilon \int_K |\nabla v|^2 dx,
\]

where \( \epsilon > 0 \) is any real number.

**Theorem 4.3.** Let \( u \) be the exact solution of (2) and let \( u_h \) be its finite element approximation given by (4). Let \( q = a \nabla u \) be the flux/gradient with the obvious approximation \( q_h = a \nabla u_h \) and let \( Q_\tau \) be a postprocessing operator given by \( Q_\tau \). Then,
for any projection space $\mathcal{L}_r$ which is a piecewise polynomial of order $r$, we have for any $\epsilon > 0$

\[(31) \quad \|q - G_\tau q_h\|_0 \leq C \tau^{r+1} \|q\|_{r+1, \tau} + C \tau^{-\frac{1}{2}} ((\tau^{-\frac{1}{2}} + \epsilon^{-\frac{1}{2}})) \|u - u_h\|_0 + C \epsilon^{\frac{1}{2}} \|u - u_h\|_1.\]

Proof. Since $G_\tau = Q_\tau$, then

\[\|q - G_\tau q_h\|_0 \leq \|q - Q_\tau q\|_0 + \|Q_\tau q - Q_\tau q_h\|_0.\]

The error $\|q - Q_\tau q\|_0$ can be estimated by using (15). To estimate $\|Q_\tau (q - q_h)\|_0$, we see from (23) and (24) that it suffices to deal with $(\nabla (u - u_h), \phi)$ for any $\phi \in \mathcal{L}_r$ such that $\|\phi\| = 1$. To this end, using the integration by parts we obtain

\[(32) \quad (\nabla (u - u_h), \phi) = \sum_{K \in \Omega_r} \int_K (u_h - u) \nabla \cdot \phi dx + \sum_{K \in \Omega_r} \int_{\partial K} (u - u_h) \phi \cdot n_K ds.\]

The first term on the right-hand side of (32) can be bounded as follows:

\[(33) \quad \left| \sum_{K \in \Omega_r} \int_K (u_h - u) \nabla \cdot \phi dx \right| \leq \|u - u_h\|_0 \|\nabla \cdot \phi\|_0 \leq C \tau^{-\frac{1}{2}} \|u - u_h\|_0,\]

where we have used the standard inverse estimate for $\|\nabla \cdot \phi\|$. To estimate the second term on the right-hand side of (32), we use the Schwarz inequality to obtain

\[(34) \quad \left| \sum_{K \in \Omega_r} \int_{\partial K} (u - u_h) \phi \cdot n_K ds \right| \leq \sum_{K \in \Omega_r} \left( \|u - u_h\|_{0, \partial K} \|\phi\|_{0, \partial K} \right) \leq \left( \sum_{K \in \Omega_r} \|u - u_h\|^2_{0, \partial K} \right)^{\frac{1}{2}} \left( \sum_{K \in \Omega_r} \|\phi\|^2_{0, \partial K} \right)^{\frac{1}{2}}.\]

It follows from (30) that

\[\|u - u_h\|^2_{0, \partial K} \leq C \left( (\tau^{-1} + \epsilon^{-1}) \|u - u_h\|^2_{0, K} + \epsilon \|\nabla (u - u_h)\|^2_{0, K}.\right)\]

Similarly, from (30) with $\epsilon = \tau$ we have

\[\|\phi\|^2_{0, \partial K} \leq C (\tau^{-1} \|\phi\|^2_{1, K} + \tau \|\nabla \phi\|^2_{0, K}).\]

Substituting the above two estimates into (34) yields

\[\left| \sum_{K \in \Omega_r} \int_{\partial K} (u - u_h) \phi \cdot n_K ds \right| \leq C \left( (\tau^{-1} + \epsilon^{-1}) \|u - u_h\|^2_{0, K} + \epsilon \|\nabla (u - u_h)\|^2_{0, K} \right)^{\frac{1}{2}} \left( \tau^{-1} \|\phi\|^2_{0, K} + \tau \|\nabla \phi\|^2_{0, K} \right)^{\frac{1}{2}}.\]

Now using the standard inverse inequality and the fact that $\|\phi\|_0 = 1$, we obtain

\[\tau^{-1} \|\phi\|^2_{0} + \tau \|\nabla \phi\|^2_{0} \leq C \tau^{-1}.\]
Thus,

\begin{equation}
(35) \quad \left| \sum_{K \in \Omega_r} \int_{\partial K} (u - u_h) \phi \cdot n_K \, ds \right| \\
\leq C \tau^{-\frac{1}{2}} \left( (\tau^{-\frac{1}{2}} + \epsilon^{-\frac{1}{2}}) \|u - u_h\|_0 + \epsilon^{\frac{3}{2}} \|\nabla (u - u_h)\|_0 \right).
\end{equation}

The combination of (32) with (33) and (35) gives

\[ \|\nabla (u - u_h)\|_0 \leq C \tau^{-\frac{1}{2}} \left( (\tau^{-\frac{1}{2}} + \epsilon^{-\frac{1}{2}}) \|u - u_h\|_0 + \epsilon^{\frac{3}{2}} \|\nabla (u - u_h)\|_0 \right), \]

which completes the proof. \(\square\)

The discontinuous projection space \(L_r\) has many distinguished features in theory and application. In practical implementation, it allows a local and parallel computation of the projected flux \(Q \cdot q_h\). In addition, one does not need to worry about any special treatment of the boundary condition \(u = g\). From the analysis of Theorem 4.3, we see that the estimate (31) does not require the regularity/smoothness assumptions (3) and (8). However, in order to get a superconvergence from the estimate (31), the \(L^2\) norm of the error must have a higher order of convergence than the \(H^1\) norm. This is often accomplished via a duality argument which requires a certain regularity for the dual problem.

For illustration, we consider a model problem where the flux \(q\) and the solution \(u\) satisfy

\begin{equation}
(36) \quad C(u) = \|u\|_{k+1,h} < \infty, \quad C(q) = \|q\|_{r+1,\tau} < \infty.
\end{equation}

Assume that the \(H^2\) regularity is satisfied for the dual problem. Then the following error estimate is well known:

\[ \|u - u_h\|_0 + h \|\nabla (u - u_h)\|_0 \leq C(u) h^{k+1}. \]

Substituting the above with \(\tau = h^\alpha\) and \(\epsilon = h\) into (31) yields

\[ \|q - \mathcal{G}_\tau q_h\|_0 \leq C(u,q)(h^{\alpha(r+1)} + h^{k + 0.5 - 0.5\alpha}). \]

The above estimate is optimized when

\[ \alpha(r+1) = k + 0.5 - 0.5\alpha \iff \alpha = \frac{k + 0.5}{r+1.5}, \]

which gives the following error estimate:

\begin{equation}
(37) \quad \|q - \mathcal{G}_\tau q_h\|_0 \leq C(u,q) h^{\left(\frac{k+0.5(r+1)}{r+1.5}\right)}.
\end{equation}

With \(k = 2\) and \(r = 3\) we obtain

\[ \|q - \mathcal{G}_\tau q_h\|_0 \leq C(u,q) h^{\frac{3}{2}}, \]

which is much better than the optimal order error estimate \(O(h^2)\) for the straightforward gradient approximation \(q_h = a \nabla u_h\).

For problems with reentrant corners in the domain or discontinuous data in the coefficient tensor \(\{a_{ij}\}\), the \(H^2\) regularity for the dual problem is not satisfied. The dual problem, however, has the \(H^{1+\sigma}(\Omega)\) regularity for some \(\sigma \in (0,1)\). For sufficiently smooth solution \(u\) and the flux \(q\), it is possible to show that

\[ \|u - u_h\|_0 + h^\sigma \|\nabla (u - u_h)\|_0 \leq C(u) h^{k+\sigma}. \]

By choosing \(\epsilon\) properly in Theorem 4.3, one is able to determine a value of \(\alpha\) in \(\tau = h^\alpha\) which gives a superconvergence for the gradient. Details of this analysis are omitted.
5. A relation with ZZ patch recovery. For a given element \( K \) of \( S_h \), the ZZ method projects \( \nabla u_h \) to a covering patch \( \tilde{K} \) consisting of \( K \) and several neighboring elements \([32, 33]\). The projection space in the ZZ method is the restriction of \( S_h \) on each patch \( \tilde{K} \). The superconvergence of the original ZZ patch recovery has been observed only through numerical experiments with specially defined discrete \( L^2 \)-inner products.

We shall modify the ZZ patch recovery method as follows. First, we replace the ZZ projection space by the space of polynomials of order \( r \geq 0 \) on each patch. Second, we assume that each patch is of size \( \tau \) which is larger than the original mesh size \( h \). By adjusting the size \( \tau \) and the fitting polynomial order \( r \), we are able to obtain a superconvergence for the MZZ patch recovery method.

We now present a detailed discussion on the MZZ method. Based on the finite element partition for \( S_h \), we shall first divide the mesh domain \( \Omega \) into many nonoverlapping and simply connected subdomains \( \Omega_i \). Each subdomain is a union of finite elements of \( \Omega_h \). Assume that the partition \( \{ \Omega_i \} \) is regular in the sense that each \( \Omega_i \) is of diameter proportional to \( \tau \) and contains a ball of diameter also proportional to \( \tau \). The finite element space \( L_\tau \) is defined as

\[
L_\tau = \{ v = (v_1, \ldots, v_d) : v_i|_{\Omega_i} \in P_r \forall i \}.
\]

In other words, for \( v \in L_\tau \), each component of \( v \) in \( \Omega_i \) is the restriction of a polynomial of order \( r \). Notice that \( v \) can be discontinuous on the interface between the subdomains. Here are some points of why discontinuous fitting functions are preferable:

- Each subdomain (element) \( \Omega_i \) is constructed from the elements of \( \Omega_h \) by regrouping. Thus, the implementation of the projection operator \( Q_\tau \) is computationally feasible, since the corresponding numerical integration for the matrix problem of \( Q_\tau \) is easy to compute.
- As the functions \( v \) can be totally discontinuous on the interfaces, the boundary of the element \( \Omega_i \) does not need to be straight lines.
- The projection operator \( Q_\tau \) can be computed on each subdomain in parallel, and the projections over the subdomains do not interact with each other.

An error estimate can be established for the MZZ scheme by using Theorem 4.3. In fact, from the estimate (31) with \( \epsilon = h \) we have

\[
\| q - G_\tau q_h \|_0 \leq C \tau^{r+1} \| q \|_{r+1, \tau} + \tau^{-\frac{1}{2}} (h^{-\frac{1}{2}} \| u - u_h \|_0 + h^\frac{1}{2} \| u - u_h \|_1),
\]

where \( G_\tau \) is given by \( Q_\tau \). If the exact solution \( u \) is sufficiently smooth, then

\[
\| q - G_\tau q_h \|_0 \leq C (\tau^{r+1} \| q \|_{r+1, \tau} + \tau^{-\frac{1}{2}} h^{k+0.5} \| u \|_{k+1,h}).
\]

For a given \( \tau \), we can choose the order of polynomials of \( L_\tau \) properly such that \( \tau^{r+1} \leq \sqrt{\frac{h}{\tau}} \). Hence, the gain on the convergence for the flux/gradient is of a factor \( \sqrt{\frac{h}{\tau}} \).

In case that \( u \) has only a limited regularity, we need to choose the mesh size \( \tau \) properly to obtain the best possible superconvergence. For example, if \( q \in [H^2(\Omega)]^d \) and \( u \in H^2(\Omega) \) and \( S_h \) contains continuous piecewise linear functions, we take the projection space \( L_\tau \) to be the restriction of linear functions on each subdomain (or patch) \( \Omega_i \). For any \( v \in L_\tau \), we have

\[
v|_{\Omega_i} = a_0^i + \sum_{j=1}^d a_j^i x_j.
\]
In computing the projection of the flux over each $\Omega_i$, we need only to compute the coefficients $a_i^0, a_i^1, \ldots, a_i^r$ on each $\Omega_i$ by using the standard least-squares method. As $k = r = 1$, we have that

$$
\|q - G_\tau q_h\|_0 \leq C\tau^2 \|q\|_{2,\tau} + C(h/\tau)^{0.5} h \|u\|_{2,h}.
$$

By choosing $\tau = h^{\frac{3}{5}}$, we arrive at the following superconvergence:

$$
\|q - G_\tau q_h\|_0 \leq C h^{\frac{5}{2}} (\|q\|_{2,\tau} + \|u\|_{2,h}).
$$

The gain for the convergence order for the flux or gradient is then $1$. In case that $r = 3$, the gain of the convergence order can be $1/3$.

For the original ZZ method, it is typical that $r = k$ and $\tau = Lh$ for some fixed value $L \geq 1$. Correspondingly, our estimate implies that

$$
\|q - G_\tau q_h\|_0 \leq C L^{k+1} h^{k+1} \|q\|_{k+1,\tau} + C\sqrt{L^{-1}} h^k \|u\|_{k+1,h},
$$

which does not claim any superconvergence for the recovered flux or gradient approximation.

6. A remark on continuous least-squares surface fitting. Locality and parallelization are the main features in using discontinuous finite elements to fit the approximate flux. However, as indicated by (38), the maximum gain on the order of convergence with discontinuous projection space is $h^{\frac{5}{2}}$ over the optimal order error estimate. In fact, our convergence analysis in previous sections suggests that continuous finite element fitting spaces should be used in order to achieve a high order of superconvergence for the recovered flux/gradient approximation.

Let $\mathcal{L}_\tau$ be a finite element space of class $C^0$ consisting of continuous piecewise polynomials of order $r$ over each element. Recall that the finite element partition for $\Omega_h$ does not need to be a refinement of the elements of $\Omega_\tau$. It is well known that, for any $\epsilon \in (0, \frac{1}{2})$, $\mathcal{L}_\tau \subset H^{\frac{5}{2}-\epsilon}(\Omega)$. In other words, assumption (8) is satisfied with $s = 2.5 - \epsilon$. Assume that (3) is also valid with $s = 2.5 - \epsilon$. An application of Theorem 4.1 shows that the convergence for the recovered flux approximations is given by

$$
\|q - G_\tau q_h\|_0 \leq C\tau^{r+1} \|q\|_{r+1,\tau} + C(h/\tau)^{1.5-\epsilon} \|u - u_h\|_1,
$$

where $G_\tau$ is either $\bar{Q}_\tau$ or $Q_\tau$. If $u$ is sufficiently smooth, we can choose the order of polynomials of $\mathcal{L}_\tau$ properly such that $\tau^{r+1} \leq C(h/\tau)^{1.5-\epsilon}$. In such a case, the gain of the convergence for the flux and gradient is of a factor $(h/\tau)^{1.5-\epsilon}$. For simplicity of discussion, we shall assume $\epsilon = 0$ in the rest of this section. In case that $u$ has only a limited regularity, we need to choose $r$ according to the regularity of $u$ and choose $\tau$ such that $O(\tau^{r+1}) = (h/\tau)^{1.5} h^k$. In Table 1, we show some theoretical gain of the convergence order for the flux/gradient with different values of $r$ and $k$. In theory, the computational result can only be better than this.

In a similar manner, the improvement on the convergence of the flux/gradient would be of a factor $O(h/\tau)^{(\ell+1.5)}$ if $\mathcal{L}_\tau$ is a finite element space of class $C^\ell$ for $0 \leq \ell \leq k - 1.5$. In case that $u$ has only a limited regularity, we need to choose the mesh size $\tau$ properly to get the best possible superconvergence.
Table 1
The value of $\beta$ in $\|q - G_\tau q_h\|_0 \leq Ch^{k+\beta}$.

<table>
<thead>
<tr>
<th>$r$</th>
<th>$k = 1$</th>
<th>$k = 2$</th>
<th>$k = 3$</th>
<th>$k = 4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.4286</td>
<td>0.3394</td>
<td>0.3333</td>
<td>0.2727</td>
</tr>
<tr>
<td>2</td>
<td>0.6667</td>
<td>0.5455</td>
<td>0.4615</td>
<td>0.4000</td>
</tr>
<tr>
<td>3</td>
<td>0.8182</td>
<td>0.6923</td>
<td>0.6000</td>
<td>0.5000</td>
</tr>
<tr>
<td>4</td>
<td>0.9231</td>
<td>0.8000</td>
<td>0.7071</td>
<td>0.6000</td>
</tr>
<tr>
<td>5</td>
<td>1.0000</td>
<td>0.8000</td>
<td>0.8000</td>
<td>0.7071</td>
</tr>
</tbody>
</table>

7. An application to mesh adaptivity. The superconvergence estimates can be used to refine the finite element mesh adaptively. Let us note that

$$\|q - q_h\|_0 \leq \|q - G_\tau q\|_0 + \|G_\tau (q - q_h)\|_0 + \|G_\tau q_h - q_h\|_0.$$  

Assume that the finite element solution $u_h$ is nontrivial in the sense that

$$\|q - q_h\|_0 \approx h^k \|q\|_{k,h}.$$  

From (15), (22), and (28), and assuming that $q$ has the needed regularity, we obtain

$$\|q - G_\tau q\|_0 + \|G_\tau (q - q_h)\|_0 \leq C r^r \|q\|_{r+1,\tau} + \|G_\tau (q - q_h)\|_0 \leq C r^r \|q\|_{r+1,\tau} + \alpha \|q - q_h\|_0,$$

where

$$\alpha = C r^r \|q\|_{r+1,\tau} + \|q - q_h\|_0.$$

The mesh parameters $\tau$ and $r$ can be chosen properly to ensure that $\alpha \to 0$ when $\tau \to 0$, $h \to 0$. For simplicity, let us take $\tau = \kappa h$. Thus, 

$$\alpha = C (q) \kappa^{r+1} h^{r+1-k} + C \kappa^{-(s-1)}.$$  

By letting $\kappa \to \infty$ and $\kappa \leq o(h^{k/(r+1)-1})$, we have $\alpha \to 0$. In fact, the choices we have discussed for $k, r, h, \tau$ in sections 5 and 6 will all guarantee that $\alpha \to 0$. Thus, 

$$(1 - \alpha) \|q - q_h\|_0 \leq \|G_\tau q_h - q_h\|_0.$$  

We emphasize that the right-hand side of the above estimate is computable. In practical computations, the value of $\alpha$ is small but not known exactly. We can produce a mesh to guarantee that

$$\|G_\tau q_h - q_h\|_0 \leq \varepsilon,$$

where $\varepsilon$ stands for a prescribed tolerance. To this end, for a given coarse mesh $\mathcal{L}_\tau$, we compute the maximum value of the error indicator over all the coarse mesh elements:

$$\eta_\tau = \max_{K \in \mathcal{L}_\tau} \|G_\tau q_h - q_h\|_{0,K}.$$
We choose a parameter $\theta \in (0, 1)$. For a given coarse mesh element $K \in \Omega_r$, we refine $K$ if
\[ \| \mathcal{G}_r q_h - q_h \|_{0,K} \geq \theta \eta_r. \]
The refinement process is stopped if either (40) is satisfied or the memory limit has been reached, or the change of the computed solution in the energy norm is less than a given tolerance.

Under assumption (39), the error indicator $\| \mathcal{G}_r q_h - q_h \|_0$ is in fact equivalent to the true error due to the fact that
\[ \| \mathcal{G}_r q_h - q_h \|_0 \leq \| \mathcal{G}_r (q_h - q) \|_0 + \| \mathcal{G}_r q - q \|_0 + \| q - q_h \|_0 \leq (1 + \alpha) \| q - q_h \|_0. \]

Note that we refine the coarse mesh $\Omega_r$ instead of the fine mesh $\Omega_h$. The fine mesh $\Omega_h$ is always produced from $\Omega_r$ by refining each coarse mesh element into several smaller elements. See [5] for some results about using averaging-type error estimators and the ZZ method for mesh refinement for general unstructured meshes.

8. Numerical experiments. Two meshes $\Omega_r$ and $\Omega_h$ are needed in the computation. The coarse mesh $\Omega_r$ is produced by the adaptive strategy of section 7. The fine mesh $\Omega_h$ is always produced from $\Omega_r$. To produce $\Omega_h$, each coarse mesh element is refined into 4 elements by connecting the edge middle points or refined uniformly twice to produce 16 elements for two-dimensional problems. Continuous piecewise linear finite element functions over $\Omega_r$ and $\Omega_h$ are used for the projection $\mathcal{G}_r$ and for the solution of the finite element approximation $u_h$, respectively.

The proposed algorithms are tested for
\[ -\nabla \cdot (a \nabla u) = f \quad \text{on} \quad \Omega, \quad u = g \quad \text{on} \quad \partial \Omega \]
with $\Omega = (0, 1) \times (0, 1), f = 2\pi^2 \sin(\pi x) \sin(\pi y), g = 0, a = 1$. The exact solution is easily seen to be $u = \sin(\pi x) \sin(\pi y)$.

The global coarse mesh recovery is as described earlier, and the element-wise coarse mesh recovery works by projecting the gradient in $S_h$ to the fitting space $\mathcal{L}_r$ consisting of piecewise linear functions over $\Omega_r$. Equation (14) is thus solved for each element in $\Omega_r$ independently, and these local solutions are combined to a global solution by an averaging procedure. As we shall see, this gives a worse convergence rate than the global projection, but it is still superconvergent.

We use Figures 1 and 2 to show the computational results. In the plot, the $x$-axis represents the degree of freedom of the mesh. The $y$-axis represents the $L^2$ error of the gradient. Note that both axes are scaled using $\log_{10}$. Figure 2 compares the mesh quality produced by the superconvergence error estimator and the error estimator of Johnson [16] and Eriksson and Johnson [12]. The error for the finite element solution over $S_h$, i.e., the mesh produced by the superconvergence error estimator, is slightly better than the error for the finite element solution for the mesh produced by the error estimator of [16, 12]. To reach the same accuracy, we need a much smaller degree of freedom in our new method; see also Figure 1. The projected gradient over $\Omega_r$ has a better convergence order, as can be seen from Figure 1. The convergence rate for the different errors are plotted in Figure 1. The convergence rate and the accuracy of the MZZ and the ZZ methods are nearly the same.
Fig. 1. Comparison of the mesh quality produced by different recovery methods compared to the error estimator of Johnson [16] and Eriksson and Johnson [12]. All the errors are measured in $L^2$ for the gradients. FEM refers to the error $\|\nabla u_h - \nabla u\|_0$, where $u_h$ is the finite element solution over $S_h$.

Fig. 2. The left is the mesh produced by the superconvergence estimator of this paper, and the right is the mesh of the residual error estimator of Johnson [16] and Eriksson and Johnson [12]. The $L^2$ error for the left mesh is 0.0094 with 2279 nodes, and the right mesh has an $L^2$ error of 0.061 with 2497 nodes.

Our computational experiment reveals that the convergence rate of standard Galerkin finite element solution over $S_h$ is 1.2. For ZZ recovery the order of convergence is 1.72. The MZZ method proposed earlier has a convergence order of 1.74.
Johnson’s method gives a rate of 1.08. Finally, the $L^2$ projection has a convergence of order 1.84. We clearly see that the $L^2$ projection has superior performance. The convergence order is evaluated by using the formula $\log_{10} \| \nabla u_h - \nabla u \|_0 / \log_{10}(\sqrt{\text{DoF}})$ as in [6, 22], where DoF stands for the total number of nodal points in the finite element partition. The continuous least-squares surface fitting is easy to implement and has the best accuracy.

REFERENCES


