

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

*Received by the editors June 21, 2000; accepted for publication (in revised form) February 25, 2002; published electronically September 27, 2002.

<http://www.siam.org/journals/sinum/40-4/37410.html>

[†]Department of Mathematics, University of Bergen, Johannes Brunsgt. 12, 5007 Bergen, Norway (bjornoh@math.uib.no, Xue-Cheng.Tai@math.uib.no). The research of the second author was partially supported by the Research Council of Norway (NFR) under grant 128224/431.

[‡]Department of Mathematical and Computer Sciences, Colorado School of Mines, Golden, CO 80401 (jwang@mines.edu). The research of this author was supported in part by the NSF grant DMS-9706985.

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 ∇u . The approach of this paper is different in that the projection will be applied directly to the numerical gradient ∇u_h in order to provide a superconvergent numerical solution for ∇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 Ω be an open bounded domain in \mathbb{R}^d , $d = 2, 3$. Denote by $x = (x_1, \dots, x_d)$ the points in Ω . Let $\partial_i = \frac{\partial}{\partial x_i}$ be the partial derivative operator in the direction of x_i , $i = 1, \dots, d$. The Dirichlet boundary value problem seeks a function $u = u(x)$ such that $u(x) = g(x)$ for any $x \in \partial\Omega$ and

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

where $\mathbf{a} = (a_{ij})_{i,j=1}^d$ is the coefficient tensor which is symmetric, bounded, and uniformly positive definite in the domain Ω with measurable entries $a_{ij} = a_{ij}(x)$. The other coefficients $\mathbf{b} = (b_i(x))_{i=1}^d$ 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 Ω , $H^s(\Omega)$ denotes the Sobolev space with norm $\|\cdot\|_s$ as defined in [14]. The space $H_0^s(\Omega)$ is a closed subspace of $H^s(\Omega)$ that is the closure of $C_0^s(\Omega)$ (the set of compact-supported C^s 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 $H^0(\Omega)$ coincides with $L^2(\Omega)$, in which case the norm and inner product are denoted by $\|\cdot\|$ and (\cdot, \cdot) , respectively.

Let

$$a(u, v) = \sum_{i,j=1}^d \int_{\Omega} a_{ij} \partial_i u \partial_j v dx + \sum_{i=1}^d \int_{\Omega} b_i \partial_i u v dx + \int_{\Omega} c u v dx$$

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

$$(2) \quad a(u, v) = (f, v) \quad \forall v \in H_0^1(\Omega).$$

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

$$a(v, w) = (f, v) \quad \forall v \in H_0^1(\Omega)$$

has a unique solution $w \in H_0^1(\Omega) \cap H^s(\Omega)$ such that

$$(3) \quad \|w\|_s + \left\| \frac{\partial w}{\partial \mathbf{n}_a} \right\|_{s-\frac{3}{2}, \partial\Omega} \leq C \|f\|_{s-2},$$

where \mathbf{n} is the unit outward normal vector of $\partial\Omega$ and $\frac{\partial w}{\partial \mathbf{n}_a} = (\mathbf{a} \nabla w) \cdot \mathbf{n}$ 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 Ω_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 $S_h^0 = H_0^1(\Omega) \cap S_h$ and g_h is a certain approximation of the Dirichlet boundary data g . Let Λ_h be the restriction of the finite element space S_h on the boundary of Ω . For simplicity, we shall deal with polygonal or polyhedral domain Ω so that the boundary $\partial\Omega$ is exactly represented by the finite element partition Ω_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 Λ_h for sufficiently smooth g .
- g_h is the L^2 projection of g in Λ_h .

We recall that the L^2 projection of g in Λ_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 Ω_h . The finite element space S_h is assumed to have the following approximation property:

$$\inf_{v \in S_h} (\|w - v\| + h \|w - v\|_1) \leq Ch^m \|w\|_m \quad \forall w \in H^m(\Omega)$$

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}_τ with parameter $\tau \gg h$ and a higher order approximation property than S_h [15]. The functions in \mathcal{L}_τ are vector valued and will be employed to approximate the exact gradient/flux variable $\mathbf{q} = \mathbf{a}\nabla u$. In practice, the mesh parameter τ is proportional to h^α for some $\alpha \in (0, 1)$ in order to obtain a superconvergent approximation from the projection space \mathcal{L}_τ . Details can be found from Wang [27].

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

- Inverse property.

$$(6) \quad \|\mathbf{v}_\tau\|_{[H^m(K)]^d} \leq C\tau^{-m} \|\mathbf{v}_\tau\|_{[L^2(K)]^d} \quad \forall \mathbf{v}_\tau \in \mathcal{L}_\tau, \quad \forall K \in \Omega_\tau$$

for all nonnegative integer $m \geq 0$.

- Approximation property.

$$(7) \quad \inf_{\mathbf{v}_\tau \in \mathcal{L}_\tau} \|\mathbf{v} - \mathbf{v}_\tau\|_0 \leq C\tau^m \|\mathbf{v}\|_m \quad \forall \mathbf{v} \in [H^m(\Omega)]^d, \quad 0 \leq m \leq r + 1.$$

- Smoothness property.

$$(8) \quad \mathcal{L}_\tau \subset [H^{s-1}(\Omega)]^d.$$

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

We emphasize that the space \mathcal{L}_τ 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 \mathbf{q} by using the finite element solution u_h . The relation between the flux $\mathbf{q} = \mathbf{q}(x)$ and the original function $u = u(x)$ can be rewritten as follows:

$$(9) \quad \mathbf{a}^{-1}\mathbf{q} = \nabla u.$$

Let

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

be equipped with the norm

$$\|\mathbf{v}\|_H = (\|\mathbf{v}\|^2 + \|\nabla \cdot \mathbf{v}\|^2)^{1/2}.$$

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

$$(10) \quad (\mathbf{a}^{-1}\mathbf{q}, \mathbf{v}) = -(u, \nabla \cdot \mathbf{v}) + \langle g, \mathbf{v} \cdot \mathbf{n} \rangle,$$

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 $\tilde{\mathbf{q}}_\tau$ defined as follows:

$$(11) \quad (\mathbf{a}^{-1}\tilde{\mathbf{q}}_\tau, \mathbf{v}) = -(u_h, \nabla \cdot \mathbf{v}) + \langle g, \mathbf{v} \cdot \mathbf{n} \rangle \quad \forall \mathbf{v} \in \mathcal{L}_\tau.$$

The new flux approximation $\tilde{\mathbf{q}}_\tau$ will be denoted by

$$\tilde{\mathbf{q}}_\tau = \tilde{\mathbf{Q}}_\tau \mathbf{q}_h.$$

It is clear that $\tilde{\mathbf{Q}}_\tau$ can be regarded as a linear operator onto the fitting space \mathcal{L}_τ .

Since (11) was obtained by using test functions in the space $H(\text{div}; \Omega)$, the fitting space \mathcal{L}_τ 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 \mathcal{L}_τ . 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 \mathbf{Q}_τ be the weighted L^2 projection onto the fitting space \mathcal{L}_τ with respect to the weighted inner product $(\mathbf{a}^{-1}\cdot, \cdot)$. More precisely, for any $\mathbf{v} \in [L^2(\Omega)]^d$, the projection $\mathbf{Q}_\tau \mathbf{v}$ is a function in \mathcal{L}_τ such that

$$(12) \quad (\mathbf{a}^{-1} \mathbf{Q}_\tau \mathbf{v}, \phi) = (\mathbf{a}^{-1} \mathbf{v}, \phi) \quad \forall \phi \in \mathcal{L}_\tau.$$

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

$$\|\mathbf{q} - \mathbf{q}_h\| \leq C \|u - u_h\|_1$$

for some constant C .

With the L^2 -projection operator \mathbf{Q}_τ , we can provide a new flux approximation given as follows:

$$(13) \quad \mathbf{q} = \mathbf{a} \nabla u \approx \mathbf{Q}_\tau \mathbf{q}_h.$$

From the definition of \mathbf{Q}_τ , we see that the new flux approximation $\mathbf{Q}_\tau \mathbf{q}_h$ satisfies the following system of equations:

$$(14) \quad (\mathbf{a}^{-1} \mathbf{Q}_\tau \mathbf{q}_h, \phi) = (\nabla u_h, \phi) \quad \forall \phi \in \mathcal{L}_\tau.$$

When the fitting space \mathcal{L}_τ consists of discontinuous piecewise polynomials of order k on each element of Ω_τ , 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 \mathcal{L}_τ 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 Ω_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 Λ_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:

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

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 $\alpha = (\alpha_1, \dots, \alpha_d), \alpha_i \geq 0$ is a multi-index and $D^\alpha v = \partial_1^{\alpha_1} \dots \partial_d^{\alpha_d} v$. Thus, our function v above needs only to be in the Sobolev space $H^m(K)$ over each element K from Ω_h or Ω_τ in order to guarantee that the norms exist. If (6) and (7) are satisfied, then it is not hard to prove the following estimate:

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

4.1. Projection space \mathcal{L}_τ of class $H(\text{div}; \Omega)$. Let us analyze the approximation schemes (11) and (13) by assuming that the projection space \mathcal{L}_τ 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 $\mathbf{q} = \mathbf{a}\nabla u$ be the flux/gradient with the obvious approximation $\mathbf{q}_h = \mathbf{a}\nabla u_h$ and let \mathcal{G}_τ be a postprocessing operator given by either \mathbf{Q}_τ or $\tilde{\mathbf{Q}}_\tau$ 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 \mathcal{L}_τ is constructed so that $\mathcal{L}_\tau \subset H(\text{div}; \Omega)$. Then*

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

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

$$(17) \quad \|\tilde{\mathbf{Q}}_\tau \mathbf{q}_h - \mathbf{Q}_\tau \mathbf{q}\|_0 \leq C \sup_{\phi \in \mathcal{L}_\tau, \|\phi\|_{L^2} = 1} (\mathbf{a}^{-1} \tilde{\mathbf{Q}}_\tau \mathbf{q}_h - \mathbf{a}^{-1} \mathbf{Q}_\tau \mathbf{q}, \phi).$$

For a given $\phi \in \mathcal{L}_\tau$, it is true that

$$\begin{aligned} (\mathbf{a}^{-1} \tilde{\mathbf{Q}}_\tau \mathbf{q}_h - \mathbf{a}^{-1} \mathbf{Q}_\tau \mathbf{q}, \phi) &= (\mathbf{a}^{-1} \tilde{\mathbf{Q}}_\tau (\mathbf{a}\nabla u_h) - \mathbf{a}^{-1} \mathbf{Q}_\tau (\mathbf{a}\nabla u), \phi) \\ &= -(u_h, \nabla \cdot \phi) + \langle g, \phi \cdot \mathbf{n} \rangle - (\nabla u, \phi) = (u - u_h, \nabla \cdot \phi). \end{aligned}$$

Define $w \in H_0^1(\Omega)$ to be the solution of

$$(18) \quad a(v, w) = (v, \nabla \cdot \phi) \quad \forall v \in H_0^1(\Omega).$$

Applying the theory of distributions, it can be proved that

$$a(v, w) - \left\langle \frac{\partial w}{\partial \mathbf{n}_\mathbf{a}}, v \right\rangle = (\nabla \cdot \phi, v) \quad \forall v \in H^1(\Omega).$$

Using (2), (4), (5), and the above equation, it is easy to see that, for any $v \in S_h^0$ and $\xi \in \Lambda_h$,

$$\begin{aligned} (u - u_h, \nabla \cdot \phi) &= a(u - u_h, w) - \left\langle \frac{\partial w}{\partial \mathbf{n}_a}, u - u_h \right\rangle \\ &= a(u - u_h, w - v) - \left\langle \frac{\partial w}{\partial \mathbf{n}_a} - \xi, g - g_h \right\rangle. \end{aligned}$$

It follows that

$$|(u - u_h, \nabla \cdot \phi)| \leq C \|u - u_h\|_1 \|w - v\|_1 + \left\| \frac{\partial w}{\partial \mathbf{n}_a} - \xi \right\|_{-\frac{1}{2}, \partial\Omega} \|u - u_h\|_{\frac{1}{2}, \partial\Omega}.$$

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

$$(19) \quad \inf_{v \in S_h} \|w - v\|_1 \leq Ch^{s-1} \|w\|_s,$$

$$(20) \quad \inf_{\xi \in \Lambda_h} \left\| \frac{\partial w}{\partial \mathbf{n}_a} - \xi \right\|_{-\frac{1}{2}, \partial\Omega} \leq Ch^{s-1} \|w\|_s,$$

we obtain

$$(21) \quad |(u - u_h, \nabla \cdot \phi)| \leq Ch^{s-1} \|u - u_h\|_1 \|w\|_s$$

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

$$|(u - u_h, \nabla \cdot \phi)| \leq Ch^{s-1} \|u - u_h\|_1 \|\nabla \cdot \phi\|_{s-2} \leq Ch^{s-1} \tau^{1-s} \|u_h - u\|_1 \|\phi\|_0,$$

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

$$(22) \quad \|\tilde{Q}_\tau \mathbf{q}_h - \mathbf{Q}_\tau \mathbf{q}\|_0 \leq Ch^{s-1} \tau^{1-s} \|u_h - u\|_1,$$

which, together with (15), gives the desired error estimate for $\mathcal{G}_\tau = \tilde{Q}_\tau$.

To analyze the case $\mathcal{G}_\tau = \mathbf{Q}_\tau$, it suffices to estimate $\|\mathbf{Q}_\tau(\mathbf{q} - \mathbf{q}_h)\|_0$. Since

$$(23) \quad \|\mathbf{Q}_\tau(\mathbf{q} - \mathbf{q}_h)\|_0 \leq C \sup_{\phi \in \mathcal{L}_\tau, \|\phi\|_0=1} (\mathbf{a}^{-1} \mathbf{Q}_\tau(\mathbf{q} - \mathbf{q}_h), \phi)$$

and

$$(24) \quad (\mathbf{a}^{-1} \mathbf{Q}_\tau(\mathbf{q} - \mathbf{q}_h), \phi) = (\nabla(u - u_h), \phi),$$

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

$$(25) \quad (\nabla(u_h - u), \phi) = (u - u_h, \nabla \cdot \phi) + \langle u_h - u, \phi \cdot \mathbf{n} \rangle.$$

Let $w \in H_0^1(\Omega)$ be defined as the solution of the following problem:

$$(26) \quad a(v, w) = -(\nabla \cdot \phi, v) \quad \forall v \in H_0^1(\Omega).$$

It follows from (2), (4), and (5) that, for any $v \in S_h^0$ and $\xi \in \Lambda_h$,

$$\begin{aligned}
 (\nabla(u_h - u), \boldsymbol{\phi}) &= \langle u - u_h, \nabla \cdot \boldsymbol{\phi} \rangle + \langle u_h - u, \boldsymbol{\phi} \cdot \mathbf{n} \rangle \\
 (27) \quad &= a(u - u_h, w) + \left\langle \frac{\partial w}{\partial \mathbf{n}_a}, u_h - u \right\rangle + \langle \boldsymbol{\phi} \cdot \mathbf{n}, u_h - u \rangle \\
 &= a(u - u_h, w - v) + \left\langle \frac{\partial w}{\partial \mathbf{n}_a} + \boldsymbol{\phi} \cdot \mathbf{n} - \xi, u_h - u \right\rangle.
 \end{aligned}$$

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

$$\inf_{\xi \in S_h} \left\| \frac{\partial w}{\partial \mathbf{n}_a} + \boldsymbol{\phi} \cdot \mathbf{n} - \xi \right\|_{-\frac{1}{2}, \partial\Omega} \leq Ch^{s-1} \tau^{1-s} \|\boldsymbol{\phi}\|_0.$$

It is also not hard to see that

$$\inf_{v \in S_h^0} |a(u - u_h, w - v)| \leq Ch^{s-1} \tau^{1-s} \|u - u_h\|_1 \|\boldsymbol{\phi}\|_0.$$

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

$$|(\nabla(u_h - u), \boldsymbol{\phi})| \leq Ch^{s-1} \tau^{1-s} \|u - u_h\|_1 \|\boldsymbol{\phi}\|_0,$$

which implies that

$$(28) \quad \|\mathbf{Q}_\tau(\mathbf{q} - \mathbf{q}_h)\|_0 \leq Ch^{s-1} \tau^{1-s} \|u - u_h\|_1.$$

This completes the proof of the theorem. \square

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

$$\|\mathbf{q} - \mathcal{G}_\tau \mathbf{q}_h\|_0 \leq C(u, \mathbf{q}) (\tau^{r+1} + \tau^{1-s} h^{k+s-1}).$$

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

$$\|\mathbf{q} - \mathcal{G}_\tau \mathbf{q}_h\|_0 \leq C(u, \mathbf{q}) (\tau^{r+1} + \tau^{-k} h^{2k}).$$

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

$$\|\mathbf{q} - \mathcal{G}_\tau \mathbf{q}_h\|_0 \leq C(u, \mathbf{q}) \left(h^{\alpha(r+1)} + h^{(2-\alpha)k} \right),$$

which is optimized when

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

The corresponding error estimate is given by

$$\|\mathbf{q} - \mathcal{G}_\tau \mathbf{q}_h\|_0 \leq C(u, \mathbf{q}) h^{\frac{2k(r+1)}{r+k+1}}.$$

With $k = 3$ and $r = 3$, the above estimate implies an accuracy of order $h^{\frac{8}{3}}$ 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*

$$(29) \quad \|g - g_h\|_{0,\partial\Omega} \leq Ch^{k+1}\|g\|_{k+1,\partial\Omega}.$$

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

$$\|\mathbf{q} - \tilde{\mathbf{Q}}_\tau \mathbf{q}_h\|_0 \leq C\tau^{r+1}\|\mathbf{q}\|_{r+1,\tau} + Ch^{s-1}\tau^{1-s}\|u - u_h\|_1 + Ch^{k+1}\tau^{-\frac{1}{2}}\|g\|_{k+1,\partial\Omega}.$$

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

$$(u - u_h, \nabla \cdot \phi) = a(u - u_h, w - v) - \left\langle \frac{\partial w}{\partial \mathbf{n}_a}, g - g_h \right\rangle.$$

Thus,

$$\begin{aligned} |(u - u_h, \nabla \cdot \phi)| &\leq Ch^{s-1}\|u - u_h\|_1\|w\|_s + \left\| \frac{\partial w}{\partial \mathbf{n}_a} \right\|_{0,\partial\Omega} \|g - g_h\|_{0,\partial\Omega} \\ &\leq Ch^{s-1}\|u - u_h\|_1\|\nabla \cdot \phi\|_{s-2} + Ch^{k+1}\|w\|_{\frac{3}{2}}\|g\|_{k+1,\partial\Omega} \\ &\leq Ch^{s-1}\tau^{1-s}\|u_h - u\|_1\|\phi\|_0 + Ch^{k+1}\tau^{-1/2}\|g\|_{k+1,\partial\Omega}\|\phi\|_0. \end{aligned}$$

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}_τ .

4.2. Discontinuous projection space \mathcal{L}_τ . The flux approximation scheme (13) or (14) is well defined for discontinuous projection space \mathcal{L}_τ . 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}_τ contains discontinuous finite element functions.

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

$$(30) \quad \int_{\partial K} v^2 ds \leq C \left((\tau^{-1} + \epsilon^{-1}) \int_K v^2 dx + \epsilon \int_K |\nabla v|^2 dx \right),$$

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 $\mathbf{q} = \mathbf{a}\nabla u$ be the flux/gradient with the obvious approximation $\mathbf{q}_h = \mathbf{a}\nabla u_h$ and let \mathcal{G}_τ be a postprocessing operator given by \mathbf{Q}_τ . Then,*

for any projection space \mathcal{L}_τ which is a piecewise polynomial of order r , we have for any $\epsilon > 0$

$$(31) \quad \begin{aligned} \|\mathbf{q} - \mathcal{G}_\tau \mathbf{q}_h\|_0 &\leq C\tau^{r+1} \|\mathbf{q}\|_{r+1,\tau} \\ &\quad + C\tau^{-\frac{1}{2}} ((\tau^{-\frac{1}{2}} + \epsilon^{-\frac{1}{2}}) \|u - u_h\|_0 + \epsilon^{\frac{1}{2}} \|u - u_h\|_1). \end{aligned}$$

Proof. Since $\mathcal{G}_\tau = \mathbf{Q}_\tau$, then

$$\|\mathbf{q} - \mathcal{G}_\tau \mathbf{q}_h\|_0 \leq \|\mathbf{q} - \mathbf{Q}_\tau \mathbf{q}\|_0 + \|\mathbf{Q}_\tau \mathbf{q} - \mathbf{Q}_\tau \mathbf{q}_h\|_0.$$

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

$$(32) \quad (\nabla(u - u_h), \boldsymbol{\phi}) = \sum_{K \in \Omega_\tau} \int_K (u_h - u) \nabla \cdot \boldsymbol{\phi} dx + \sum_{K \in \Omega_\tau} \int_{\partial K} (u - u_h) \boldsymbol{\phi} \cdot \mathbf{n}_K ds.$$

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

$$(33) \quad \left| \sum_{K \in \Omega_\tau} \int_K (u_h - u) \nabla \cdot \boldsymbol{\phi} dx \right| \leq \|u - u_h\|_0 \|\nabla \cdot \boldsymbol{\phi}\|_0 \leq C\tau^{-1} \|u - u_h\|_0,$$

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

$$(34) \quad \begin{aligned} \left| \sum_{K \in \Omega_\tau} \int_{\partial K} (u - u_h) \boldsymbol{\phi} \cdot \mathbf{n}_K ds \right| &\leq \sum_{K \in \Omega_\tau} \int_{\partial K} |u - u_h| |\boldsymbol{\phi} \cdot \mathbf{n}_K| ds \\ &\leq \sum_{K \in \Omega_\tau} \|u - u_h\|_{0,\partial K} \|\boldsymbol{\phi}\|_{0,\partial K} \\ &\leq \left(\sum_{K \in \Omega_\tau} \|u - u_h\|_{0,\partial K}^2 \right)^{\frac{1}{2}} \left(\sum_{K \in \Omega_\tau} \|\boldsymbol{\phi}\|_{0,\partial K}^2 \right)^{\frac{1}{2}}. \end{aligned}$$

It follows from (30) that

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

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

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

Substituting the above two estimates into (34) yields

$$\begin{aligned} &\left| \sum_{K \in \Omega_\tau} \int_{\partial K} (u - u_h) \boldsymbol{\phi} \cdot \mathbf{n}_K ds \right| \\ &\leq C((\tau^{-1} + \epsilon^{-1}) \|u - u_h\|_0^2 + \epsilon \|\nabla(u - u_h)\|_0^2)^{\frac{1}{2}} (\tau^{-1} \|\boldsymbol{\phi}\|_0^2 + \tau \|\nabla \boldsymbol{\phi}\|_0^2)^{\frac{1}{2}}. \end{aligned}$$

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

$$\tau^{-1} \|\boldsymbol{\phi}\|_0^2 + \tau \|\nabla \boldsymbol{\phi}\|_0^2 \leq C\tau^{-1}.$$

Thus,

$$(35) \quad \left| \sum_{K \in \Omega_\tau} \int_{\partial K} (u - u_h) \phi \cdot \mathbf{n}_K ds \right| \leq C\tau^{-\frac{1}{2}}((\tau^{-\frac{1}{2}} + \epsilon^{-\frac{1}{2}})\|u - u_h\|_0 + \epsilon^{\frac{1}{2}}\|\nabla(u - u_h)\|_0).$$

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

$$|(\nabla(u - u_h), \phi)| \leq C\tau^{-\frac{1}{2}}((\tau^{-\frac{1}{2}} + \epsilon^{-\frac{1}{2}})\|u - u_h\|_0 + \epsilon^{\frac{1}{2}}\|\nabla(u - u_h)\|_0),$$

which completes the proof. \square

The discontinuous projection space \mathcal{L}_τ has many distinguished features in theory and application. In practical implementation, it allows a local and parallel computation of the projected flux $\mathbf{Q}_\tau \mathbf{q}_h$. 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 \mathbf{q} and the solution u satisfy

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

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

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

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:

$$(37) \quad \|\mathbf{q} - \mathcal{G}_\tau \mathbf{q}_h\|_0 \leq C(u, \mathbf{q})h^{\frac{(k+0.5)(r+1)}{r+1.5}}.$$

With $k = 2$ and $r = 3$ we obtain

$$\|\mathbf{q} - \mathcal{G}_\tau \mathbf{q}_h\|_0 \leq C(u, \mathbf{q})h^{\frac{20}{9}},$$

which is much better than the optimal order error estimate $\mathcal{O}(h^2)$ for the straightforward gradient approximation $\mathbf{q}_h = \mathbf{a}\nabla u_h$.

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 \mathbf{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 ϵ properly in Theorem 4.3, one is able to determine a value of α 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 ∇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 τ which is larger than the original mesh size h . By adjusting the size τ 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 Ω into many nonoverlapping and simply connected subdomains Ω_i . Each subdomain is a union of finite elements of Ω_h . Assume that the partition $\{\Omega_i\}$ is regular in the sense that each Ω_i is of diameter proportional to τ and contains a ball of diameter also proportional to τ . The finite element space \mathcal{L}_τ is defined as

$$\mathcal{L}_\tau = \{ \mathbf{v} = (v_1, \dots, v_d) : v_i|_{\Omega_i} \in P_r \ \forall i \}.$$

In other words, for $\mathbf{v} \in \mathcal{L}_\tau$, each component of \mathbf{v} in Ω_i is the restriction of a polynomial of order r . Notice that \mathbf{v} can be discontinuous on the interface between the subdomains. Here are some points of why discontinuous fitting functions are preferable:

- Each subdomain (element) Ω_i is constructed from the elements of Ω_h by regrouping. Thus, the implementation of the projection operator \mathbf{Q}_τ is computationally feasible, since the corresponding numerical integration for the matrix problem of \mathbf{Q}_τ is easy to compute.
- As the functions \mathbf{v} can be totally discontinuous on the interfaces, the boundary of the element Ω_i does not need to be straight lines.
- The projection operator \mathbf{Q}_τ 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

$$\| \mathbf{q} - \mathcal{G}_\tau \mathbf{q}_h \|_0 \leq C \tau^{r+1} \| \mathbf{q} \|_{r+1, \tau} + \tau^{-\frac{1}{2}} (h^{-\frac{1}{2}} \| u - u_h \|_0 + h^{\frac{1}{2}} \| u - u_h \|_1),$$

where \mathcal{G}_τ is given by \mathbf{Q}_τ . If the exact solution u is sufficiently smooth, then

$$(38) \quad \| \mathbf{q} - \mathcal{G}_\tau \mathbf{q}_h \|_0 \leq C (\tau^{r+1} \| \mathbf{q} \|_{r+1, \tau} + \tau^{-\frac{1}{2}} h^{k+0.5} \| u \|_{k+1, h}).$$

For a given τ , we can choose the order of polynomials of \mathcal{L}_τ properly such that $\tau^{r+1} \leq \sqrt{\frac{h}{\tau}} h^k$. 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 τ properly to obtain the best possible superconvergence. For example, if $\mathbf{q} \in [H^2(\Omega)]^d$ and $u \in H^2(\Omega)$ and S_h contains continuous piecewise linear functions, we take the projection space \mathcal{L}_τ to be the restriction of linear functions on each subdomain (or patch) Ω_i . For any $\mathbf{v} \in \mathcal{L}_\tau$, we have

$$\mathbf{v}|_{\Omega_i} = a_0^i + \sum_{j=1}^d a_j^i x_j.$$

In computing the projection of the flux over each Ω_i , we need only to compute the coefficients $a_0^i, a_1^i, \dots, a_d^i$ on each Ω_i by using the standard least-squares method. As $k = r = 1$, we have that

$$\|\mathbf{q} - \mathcal{G}_\tau \mathbf{q}_h\|_0 \leq C\tau^2 \|\mathbf{q}\|_{2,\tau} + C(h/\tau)^{0.5} h \|u\|_{2,h}.$$

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

$$\|\mathbf{q} - \mathcal{G}_\tau \mathbf{q}_h\|_0 \leq Ch^{\frac{6}{5}} (\|\mathbf{q}\|_{2,\tau} + \|u\|_{2,h}).$$

The gain for the convergence order for the flux or gradient is then $1/5$. If we use quadratic functions for \mathcal{L}_τ , then we need to take $\tau = h^{\frac{3}{7}}$, and the gain for the convergence order is $2/7$. 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

$$\|\mathbf{q} - \mathcal{G}_\tau \mathbf{q}_h\|_0 \leq CL^{k+1} h^{k+1} \|\mathbf{q}\|_{k+1,\tau} + C\sqrt{L^{-1}} h^k \|u\|_{k+1,h},$$

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 $h^{\frac{1}{2}}$ 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}_τ 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 Ω_h does not need to be a refinement of the elements of Ω_τ . It is well known that, for any $\epsilon \in (0, \frac{1}{2})$, $\mathcal{L}_\tau \subset H^{\frac{3}{2}-\epsilon}(\Omega)$. 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

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

where \mathcal{G}_τ is either $\tilde{\mathcal{Q}}_\tau$ or \mathcal{Q}_τ . If u is sufficiently smooth, we can choose the order of polynomials of \mathcal{L}_τ 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 τ 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}_τ is a finite element space of class C^ℓ for $0 \leq \ell \leq k - 1.5$. In case that u has only a limited regularity, we need to choose the mesh size τ properly to get the best possible superconvergence.

TABLE 1
The value of β in $\|\mathbf{q} - \mathcal{G}_\tau \mathbf{q}_h\|_0 \leq Ch^{k+\beta}$.

	$r = 1$	$r = 2$	$r = 3$	$r = 4$	$r = 5$
$k = 1$	0.4286	0.6667	0.8182	0.9231	1.0000
$k = 2$		0.3333	0.5455	0.6923	0.8000
$k = 3$			0.2727	0.4615	0.6000
$k = 4$				0.2308	0.4000

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

$$\|\mathbf{q} - \mathbf{q}_h\|_0 \leq \|\mathbf{q} - \mathcal{G}_\tau \mathbf{q}\|_0 + \|\mathcal{G}_\tau(\mathbf{q} - \mathbf{q}_h)\|_0 + \|\mathcal{G}_\tau \mathbf{q}_h - \mathbf{q}_h\|_0.$$

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

$$(39) \quad \|\mathbf{q} - \mathbf{q}_h\|_0 \cong h^k \|\mathbf{q}\|_{k,h}.$$

From (15), (22), and (28), and assuming that \mathbf{q} has the needed regularity, we obtain

$$\begin{aligned} & \|\mathbf{q} - \mathcal{G}_\tau \mathbf{q}\|_0 + \|\mathcal{G}_\tau(\mathbf{q} - \mathbf{q}_h)\|_0 \\ & \leq C\tau^{r+1} \|\mathbf{q}\|_{r+1,\tau} + \|\mathcal{G}_\tau(\mathbf{q} - \mathbf{q}_h)\|_0 \\ & \leq C \frac{\tau^{r+1}}{h^k} \frac{\|\mathbf{q}\|_{r+1,\tau}}{\|\mathbf{q}\|_{k,h}} \|\mathbf{q} - \mathbf{q}_h\|_0 + C \left(\frac{h}{\tau}\right)^{s-1} \|\mathbf{q} - \mathbf{q}_h\|_0 \\ & = \alpha \|\mathbf{q} - \mathbf{q}_h\|_0, \end{aligned}$$

where

$$\alpha = C \frac{\tau^{r+1}}{h^k} \frac{\|\mathbf{q}\|_{r+1,\tau}}{\|\mathbf{q}\|_{k,h}} + C \left(\frac{h}{\tau}\right)^{s-1}.$$

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

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

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

$$(1 - \alpha) \|\mathbf{q} - \mathbf{q}_h\|_0 \leq \|\mathcal{G}_\tau \mathbf{q}_h - \mathbf{q}_h\|_0.$$

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

$$(40) \quad \|\mathcal{G}_\tau \mathbf{q}_h - \mathbf{q}_h\|_0 \leq \varepsilon,$$

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

$$\eta_\tau = \max_{K \in \Omega_\tau} \|\mathcal{G}_\tau \mathbf{q}_h - \mathbf{q}_h\|_{0,K}.$$

We choose a parameter $\theta \in (0, 1)$. For a given coarse mesh element $K \in \Omega_\tau$, we refine K if

$$\|\mathcal{G}_\tau \mathbf{q}_h - \mathbf{q}_h\|_{0,K} \geq \theta \eta_\tau.$$

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 $\|\mathcal{G}_\tau \mathbf{q}_h - \mathbf{q}_h\|_0$ is in fact equivalent to the true error due to the fact that

$$\begin{aligned} \|\mathcal{G}_\tau \mathbf{q}_h - \mathbf{q}_h\|_0 &\leq \|\mathcal{G}_\tau(\mathbf{q}_h - \mathbf{q})\|_0 + \|\mathcal{G}_\tau \mathbf{q} - \mathbf{q}\|_0 + \|\mathbf{q} - \mathbf{q}_h\|_0 \\ &\leq (1 + \alpha)\|\mathbf{q} - \mathbf{q}_h\|_0. \end{aligned}$$

Note that we refine the coarse mesh Ω_τ instead of the fine mesh Ω_h . The fine mesh Ω_h is always produced from Ω_τ 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 Ω_τ and Ω_h are needed in the computation. The coarse mesh Ω_τ is produced by the adaptive strategy of section 7. The fine mesh Ω_h is always produced from Ω_τ . To produce Ω_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 Ω_τ and Ω_h are used for the projection \mathcal{G}_τ and for the solution of the finite element approximation u_h , respectively.

The proposed algorithms are tested for

$$(41) \quad -\nabla \cdot (a \nabla u) = f \quad \text{on } \Omega, \quad u = g \quad \text{on } \partial\Omega$$

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}_τ consisting of piecewise linear functions over Ω_τ . Equation (14) is thus solved for each element in Ω_τ 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 Ω_τ 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.

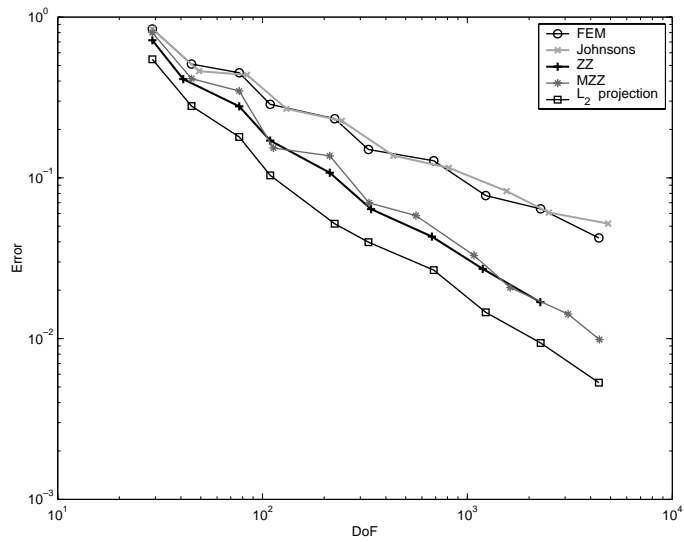


FIG. 1. Comparison of the mesh quality produced by different recovery methods compared to the error estimator of Johnson [16] and Eriksson and Johnson [12]. All the errors are measured in L^2 for the gradients. FEM refers to the error $\|\nabla u_h - \nabla u\|_0$, where u_h is the finite element solution over S_h .

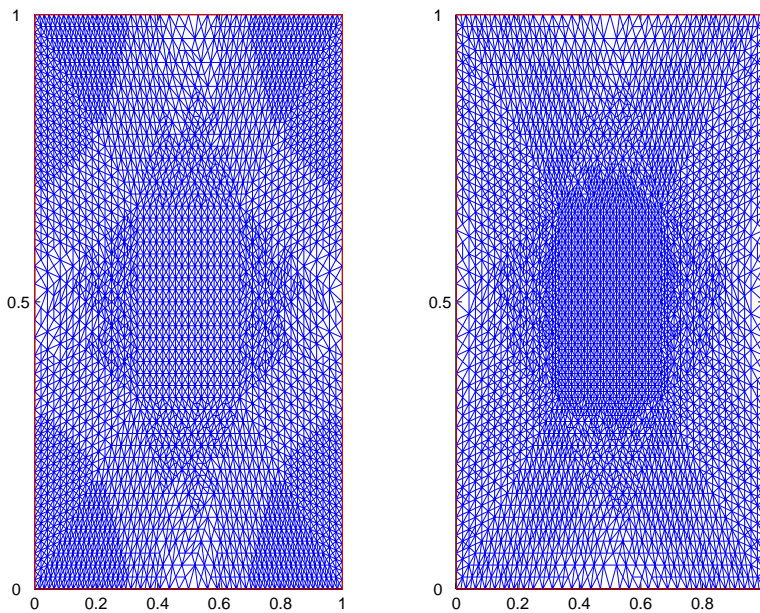


FIG. 2. The left is the mesh produced by the superconvergence estimator of this paper, and the right is the mesh of the residual error estimator of Johnson [16] and Eriksson and Johnson [12]. The L^2 error for the left mesh is 0.0094 with 2279 nodes, and the right mesh has an L^2 error of 0.061 with 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 $\log_{10} \|\nabla u_h - \nabla u\|_0 / \log_{10}(\sqrt{DoF})$ 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

- [1] J. H. BRAMBLE AND A. H. SCHATZ, *Higher order local accuracy by averaging in the finite element method*, Math. Comp., 31 (1977), pp. 94–111.
- [2] F. BREZZI, J. DOUGLAS, R. DURÁN, AND L. MARINI, *Mixed finite elements for second order elliptic problems in three variables*, Numer. Math., 52 (1987), pp. 237–250.
- [3] F. BREZZI, J. DOUGLAS, M. FORTIN, AND L. MARINI, *Efficient rectangular mixed finite elements in two and three spaces variables*, RAIRO Modél Math. Anal. Numér., 21 (1987), pp. 581–604.
- [4] F. BREZZI, J. DOUGLAS, AND L. MARINI, *Two families of mixed finite elements for second order elliptic problems*, Numer. Math., 47 (1985), pp. 217–235.
- [5] 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.
- [6] Z. CHEN AND S. DAI, *On the efficiency of adaptive finite element methods for elliptic problems with discontinuous coefficients*, SIAM J. Sci. Comput., to appear.
- [7] C. CHEN AND Y. HUANG, *High Accuracy Theory for Finite Element Methods*, Hunan Science and Technology Publishing House, Hunan, People's Republic of China, 1995 (in Chinese).
- [8] J. DOUGLAS AND T. DUPONT, *Superconvergence for Galerkin methods for the two-point boundary problem via local projections*, Numer. Math, 21 (1973), pp. 270–278.
- [9] J. DOUGLAS AND J. ROBERTS, *Global estimate for mixed finite elements methods for second order elliptic equations*, Math. Comp., 44 (1985), pp. 39–52.
- [10] J. DOUGLAS AND J. WANG, *Superconvergence of mixed finite element spaces on rectangular domains*, Calcolo, 26 (1989), pp. 121–134.
- [11] J. DOUGLAS AND J. WANG, *A new family of spaces in mixed finite element methods for rectangular elements*, Comput. Appl. Math., 12 (1993), pp. 183–197.
- [12] K. ERIKSSON AND C. JOHNSON, *Adaptive finite element methods for parabolic problems I: A linear model problem*, SIAM J. Numer. Anal., 28, (1991), pp. 43–77.
- [13] R. E. EWING, R. D. LAZAROV, AND J. WANG, *Superconvergence of the velocity along the Gauss lines in mixed finite element methods*, SIAM J. Numer. Anal., 28 (1991), pp. 1015–1029.
- [14] P. GRISVARD, *Elliptic Problems on Nonsmooth Domains*, Monogr. Stud. Math. 24, Pitman, Boston, MA, 1985.
- [15] 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. Part 1: A smooth problem and globally quasi-uniform meshes*, Math. Comp., 70 (2001), pp. 897–909.
- [16] C. JOHNSON, *Numerical Solution of Partial Differential Equations by the Finite Element Method*, Studentlitteratur, Lund, Sweden, Cambridge University Press, Cambridge, UK, 1987.
- [17] M. KRIZEK AND P. NEITTAANMAKI, *Superconvergence phenomenon in the finite element method arising from averaging gradients*, Numer. Math., 45 (1984), pp. 105–116.
- [18] R. LAZAROV, A. B. ANDREEV, AND M. HATRI, *Superconvergence of the gradients in the finite element method for some elliptic and parabolic problems*, in Variational-Difference Methods in Mathematical Physics, Part II, Moscow, 1984, pp. 13–25.
- [19] Q. LIN, *An integral identity and interpolated postprocess in superconvergence*, Research Report 90-7, Institute of Systems Science, Academia Sinica, Beijing, China, 1990, pp. 1–6.
- [20] Q. LIN, *Global error expansion and superconvergence for higher order interpolation of finite elements*, J. Comput. Math., supplementary issue, 1992, pp. 286–289.
- [21] Q. LIN AND Q. ZHU, *The Preprocessing and Postprocessing for the Finite Element Method*, Hunan Scientific & Technical Publishers, Hunan, People's Republic of China, 1994.
- [22] P. MORIN, R. H. NOCHETTO, AND K. SIEBERT, *Data oscillation and convergence of adaptive mesh*, SIAM J. Numer. Anal., 38 (2000), pp. 466–488.

- [23] L. A. OGANESYAN AND L. A. RUKHOVETZ, *Study of the rate of convergence of variational difference scheme for second order elliptic equations in two-dimensional field with a smooth boundary*, U.S.S.R. Comput. Math. and Math. Phys., 9 (1969), pp. 158–183.
- [24] P. RAVIART AND J. THOMAS, *A mixed finite element method for 2nd order elliptic problems*, in the Mathematics Aspects of Finite Element Methods, Lecture Notes in Math. 606, Springer-Verlag, Berlin, 1977, pp. 292–315.
- [25] A. H. SCHATZ, I. H. SLOAN, AND L. B. WAHLBIN, *Superconvergence in finite element methods and meshes that are locally symmetric with respect to a point*, SIAM J. Numer. Anal., 33 (1996), pp. 505–521.
- [26] L. B. WAHLBIN, *Superconvergence in Galerkin Finite Element Methods*, Lecture Notes in Math. 1605, Springer-Verlag, Berlin, 1995.
- [27] J. WANG, *A superconvergence analysis for finite element solutions by the least-squares surface fitting on irregular meshes for smooth problems*, J. Math. Study, 33 (2000), pp. 229–243.
- [28] J. WANG, *Superconvergence and extrapolation for mixed finite element methods on rectangular domains*, Math. Comp., 56 (1991), pp. 477–503.
- [29] J. WANG AND X. YE, *Superconvergence of finite element approximations for the Stokes problem by projection methods*, SIAM J. Numer. Anal., 39 (2001), pp. 1001–1013.
- [30] Z. M. ZHANG, JR., *Ultraconvergence of the patch recovery techniques II*, Math. Comp., 69 (1999), pp. 141–158.
- [31] Z. M. ZHANG AND H. D. VICTORY, JR., *Mathematical analysis of Zienkiewicz-Zhu's derivative patch recovery technique*, Numer. Methods Partial Differential Equations, 12 (1996), pp. 507–524.
- [32] O. C. ZIENKIEWICZ AND J. Z. ZHU, *The superconvergent patch recovery and a posteriori error estimates*, Part 1, Internat. J. Numer. Methods Engrg., 33 (1992), pp. 1331–1364.
- [33] O. C. ZIENKIEWICZ AND J. Z. ZHU, *The superconvergent patch recovery and a posteriori error estimates*, Part 2, Internat. J. Numer. Methods Engrg., 33 (1992), pp. 1365–1382.
- [34] Z. M. ZHANG AND J. Z. ZHU, *Superconvergence of the derivative path recovery technique and a posteriori error estimation*, in The Modeling, Mesh Generation, and Adaptive Numerical Methods for Partial Differential Equations, Springer, New York, 1995, pp. 431–450.
- [35] M. ZLÁMAL, *Superconvergence and reduced integration in the finite element method*, Math. Comp., 32 (1978), pp. 663–685.