Mathematical Analysis of Zienkiewicz–Zhu's Derivative Patch Recovery Technique

Zhimin Zhang* **and Harold Dean Victory, Jr.**

Department of Mathematics, Texas Tech University, Lubbock, Texas 79409

Received 11 January 1994; revised manuscript received 1 June 1995

Zienkiewicz–Zhu's derivative patch recovery technique is analyzed for general quadrilateral finite elements. Under certain regular conditions on the meshes, the arithmetic mean of the absolute error of the recovered gradient at the nodal points is superconvergent for the second-order elliptic operators. For rectangular meshes and the Laplacian, the recovered gradient is superconvergent in the maximum norm at the nodal points. Furthermore, it is proved for a model two-point boundary-value problem that the recovery technique results in an "ultra-convergent" derivative recovery at the nodal points for quadratic finite elements when uniform meshes are used. © 1996 John Wiley & Sons, Inc.

I. INTRODUCTION

Recently, *a posteriori* error estimates in the finite element method attract more and more attention (cf., e.g., [1]–[11] for a survey of various results), although the first article on this topic can be traced back to 1978 [12]. A closely related area is the analysis of superconvergence. This relationship is quite natural. Indeed, if we are able to recover a superconvergent solution from the post-processing, then we can use the recovered solution in place of the exact solution in computing an asymptotically exact error estimator. For the literature regarding superconvergence, the reader is referred to [13]–[17], [30], and references therein.

The derivative recovery technique has been explored since the early 1970s (cf., e.g., [18] and [19]), and many methods have been suggested, for example, averaging techniques (cf., [20]–[23]), projection techniques (cf., [18], [19]), extrapolation techniques [24], and interpolated finite elements [25], etc. Recently, O. C. Zienkiewicz and J. Z. Zhu proposed a new procedure [26] (Z–Z patch recovery) in which a continuous polynomial expansion of the function describing the derivatives is used on an element patch surrounding the nodes at which recovery is desired. This expansion is made to fit, in a least-squares sense, the derivatives of the finite element approximation at the Gaussian points. The

*To whom all correspondence should be addressed.

Numerical Methods for Partial Differential Equations 12, 507–524 (1996)

CCC 0749-159X/96/040507-18

nodal superconvergent recovery is observed from the numerical tests, although there is no theoretical justification for this behavior.

In a more recent work (see [4]), Babuška, Strouboulis, et al. developed a completely numerical methodology for checking the quality of *a posteriori* error estimators. Results from the application of the methodology to the study of several popular error estimators show that the error estimator based on the Z–Z patch recovery is the most robust one.

In this work, the Z–Z patch recovery is investigated theoretically for a class of curved isoparametric quadrilateral finite elements of an arbitrary degree r in each variable. It is shown that for a general second-order elliptic operator, the arithmetic mean of the absolute error of the recovered gradient at the nodal points is superconvergent; and for the Laplace operator with a rectangular mesh, the recovered gradient at the nodal points is superconvergent in the maximum norm. A better result is obtained for the quadratic finite element approximation to a model two-point boundary value problem. It is shown that the recovered derivative at the nodal points converges at a rate $O(h^4)$ when uniform meshes are used. Our analysis confirms the computational results in [26].

In our analysis, we assume that the exact solution of our model problem is sufficiently smooth and the meshes of the finite element approximation are regular. In many practical problems, because of the lack of smoothness of the data and the boundary, the solution has singularity at certain points. Usually, adaptive local mesh refinement is applied to deal with solution singularities, and very likely the mesh is unstructured and nonuniform. Since the interpolation property we have obtained for the recovery technique is completely local, we believe that the results in the present work, combined with some local analysis based on interior estimates (cf., [15]), should be a reasonable approach when the solution has singular behavior.

II. GRADIENT SUPERCONVERGENCE OF THE FINITE ELEMENT SOLUTION FOR THE SECOND-ORDER ELLIPTIC PROBLEM

In this article, the Dirichlet problem is chosen as the model problem. In order to concentrate on local recovery, some regularity assumptions are made on the exact solution in order to avoid the discussion of boundary behavior, corner singularities, etc. Throughout this article, C shall stand for generic constants independent of "essential" quantities that are not necessarily the same at each occurrence.

Let $\Omega \subset \mathbb{R}^2$ be a bounded domain with a sufficiently smooth boundary Γ. Consider:

$$-\sum_{i,j=1}^{2} \frac{\partial}{\partial x_i}\left(a_{ij}(\mathbf{x})\frac{\partial u}{\partial x_j}\right) + b(\mathbf{x})u = f(\mathbf{x}), \qquad \mathbf{x} := (x_1, x_2) \in \Omega; \tag{2.1}$$

$$u|_\Gamma = 0. \tag{2.2}$$

The weak solution of (2.1)–(2.2) is $u \in H_0^1(\Omega)$, which satisfies

$$a(u,v) = (f,v), \qquad \forall v \in H_0^1(\Omega), \tag{2.3}$$

where

$$a(u,v) = \int_\Omega \left(\sum_{i,j=1}^{2} a_{ij}(\mathbf{x})\frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_j} + b(\mathbf{x})uv\right) dx, \qquad (f,v) = \int_\Omega fv d\mathbf{x}.$$

Assume that $a_{ij}(\mathbf{x})$'s and $b(\mathbf{x})$ are sufficiently smooth with $b \geq 0, a_{ij}(\mathbf{x}) = a_{ji}(\mathbf{x})$, and that $a(u,v)$ is $H_0^1(\Omega)$-elliptic, i.e., there exists $\gamma > 0$ independent of $x \in \bar{\Omega}$ such that for all $\boldsymbol{\xi} = (\xi_1, \xi_2) \in \mathbb{R}^2$,

$$\sum_{i,j=1}^{2} a_{ij}(\mathbf{x})\xi_i\xi_j \geq \gamma(\xi_1^2 + \xi_2^2). \tag{2.4}$$

The aim of this article is to analyze the recovery technique. We shall not discuss the regularity of the model problem, but assume that the solution u of the model problem is sufficiently smooth to allow our analysis to be carried out.

The standard notation for Sobolev spaces is employed:

$$H^m(\Omega) := \{u \in L^2(\Omega), D^\alpha u \in L^2(\Omega), |\alpha| \leq m\}, \qquad m = 0, 1, 2, \cdots$$

$$H_0^1(\Omega) := \{u \in H^1(\Omega), u|_\Gamma = 0\},$$

$$W^{s,\infty}(\Omega) = \{u \in L^\infty(\Omega), D^\alpha u \in L^\infty(\Omega), |\alpha| \leq s\}.$$

Denote by $|\cdot|_{m,\Omega}$ and $\|\cdot\|_{m,\Omega}$, the semi-norm and the norm in $H^m(\Omega)$, which are defined, respectively, by

$$|u|_{m,\Omega}^2 = \sum_{|\alpha|=m} \|D^\alpha u\|_{L^2(\Omega)}^2, \qquad \|u\|_{m,\Omega}^2 = \sum_{|\alpha|\leq m} \|D^\alpha u\|_{L^2(\Omega)}^2,$$

and denote by $|\cdot|_{s,\infty,\Omega}$ and $\|\cdot\|_{s,\infty,\Omega}$, the semi-norm and the norm in $W^{s,\infty}(\Omega)$, which are defined, respectively, by

$$|u|_{s,\infty,\Omega} = \max_{|\alpha|=s} \|D^\alpha u\|_{\infty,\Omega}, \qquad \|u\|_{s,\infty,\Omega} = \max_{|\alpha|\leq s} \|D^\alpha u\|_{\infty,\Omega}.$$

To construct the finite element space in which the approximate solution will lie, let us cover Ω by curved isoparametric quadrilateral elements defined in the following manner. Let $L_r(\xi)$ be the Legendre polynomial of degree r on $[-1,1]$. Then $L_r(\xi)$ has r zeroes $g_k^{(r)} \in (-1,1), k = 1,\ldots,r$, and $L_r'(\xi)$ has $r-1$ zeros $l_k^{(r)} \in (-1,1), k = 1,\ldots,r-1$. Denote $l_0^{(r)} = -1, l_r^{(r)} = 1$; then

$$l_0^{(r)} < g_1^{(r)} < l_1^{(r)} < \cdots < g_k^{(r)} < l_k^{(r)} < \cdots < g_r^{(r)} < l_r^{(r)}.$$

Label $g_k^{(r)}, k = 1,\ldots,r$, as the Gaussian points on $(-1,1)$, and $l_k^{(r)}, k = 0,1,\ldots,r$, as the Lobatto points on $[-1,1]$. In two dimensions, denote $\hat{K} = [-1,1]^2$, and define

$$\{\hat{G}_i\}_{i=1}^{r^2} = \{(g_m^{(r)}, g_n^{(r)})\}_{m,n=1}^{r}, \qquad \{\hat{L}_j\}_{j=1}^{(r+1)^2} = \{(l_m^{(r)}, l_n^{(r)})\}_{m,n=0}^{r},$$

as the Gaussian and Lobatto points on $\hat{K}$, respectively.

Let $\hat{P}(r)$ denote the class of polynomials in $(\xi_1,\xi_2) \in \hat{K}$, which are of degree r, and $\hat{Q}(r)$, the class of polynomials of degree r in each variable separately. In other words, any element $p \in \hat{P}(r)$ has the form

$$p(\xi_1,\xi_2) = \sum_{i+j=0}^{r} c_{ij}\xi_1^i\xi_2^j,$$

whereas any $q \in \hat{Q}(r)$ will be of the form,

$$q(\xi_1, \xi_2) = \sum_{i=0}^{r}\sum_{j=0}^{r} c_{ij}\xi_1^i\xi_2^j.$$

Now any polynomial $\hat{v} \in \hat{Q}(r)$ is uniquely determined by its values at the Lobatto points $\hat{L}_j, j = 1, 2, \ldots, (r+1)^2, v_j = \hat{v}(\hat{L}_j), j = 1, \ldots, (r+1)^2$, or $v_{ij} = \hat{v}(l_i^{(r)}, l_j^{(r)}), i = 0, 1, \ldots, r, j = 0, 1, \ldots, r$. Let $N_i(\xi_1, \xi_2) \in \hat{Q}(r)$ be the basis functions, i.e., $N_i(\hat{L}_j) = \delta_{ij}$.

To define the isoparametric mapping, let $L_j := (x_1^{(j)}, x_2^{(j)}), j = 1, \ldots, (r+1)^2$, be $(r+1)^2$ points lying in Ω or on Γ and consider the mapping F_K: $\hat{K} \to R^2$ defined by

$$x_l = x_l(\xi_1, \xi_2) = \sum_{i=1}^{(r+1)^2} x_l^{(i)} N_i(\xi_1, \xi_2), \qquad l = 1, 2. \tag{2.5}$$

If the transformation (2.5) is one-to-one, mapping $\hat{K}$ onto a closed domain $K \subset \Omega, K$ is called a *curved quadrilateral element.* The points $L_j^K, j = 1, \ldots, (r+1)^2$, are the Lobatto points of K, and $G_j^K = (x_1(\hat{G}_j), x_2(\hat{G}_j)), j = 1, \ldots, r^2$, are the Gaussian points of K. As usual, h_K is the diameter of K, and $h = \max_K h_K$.

Suppose Ω can be decomposed by such elements K, and we denote this partition by $\mathcal{T}_h$. Let Ω_h be the interior of the union of all elements of the given partition, with Γ_h its boundary. For the sake of simplicity, assume that $\Omega = \Omega_h, \Gamma = \Gamma_h$. The same results in this work can be obtained for the general case by using the methods of P. Lesaint and M. Zlámal [27], [28].

The finite element space is defined as following:

$$S_0^{r,h} := \{v \in H_0^1(\Omega), v(x_1, x_2)|_K = \hat{v}(\xi_1^K(x_1, x_2), \xi_2^K(x_1, x_2)), K \in \mathcal{T}_h\},$$

with

$$\hat{v}(\xi_1, \xi_2) = \sum_{i=1}^{(r+1)^2} v_i N_i(\xi_1, \xi_2) \in \hat{Q}(r).$$

Here $\xi_l = \xi_l^K(x_1, x_2), l = 1, 2$, is the inverse mapping of F_K. Since the values v_j of v at nodes lying on Γ are zero, it is easy to see that $v|_\Gamma = 0$. Similarly, $S^{r,h}$ is defined without imposing zero trace on the boundary.

Definition 1. *$\mathcal{T}_h$ is r-strongly regular if F_K is a C^{r+1}-diffeomorphism for every $K \in \mathcal{T}_h$ and*

$$|D^\alpha x_l^K(\xi_1, \xi_2)| \le C_1 h_K^{|\alpha|}, \qquad \forall |\alpha| \le r+1, \qquad l = 1, 2, \tag{2.6}$$

$$C_2^{-1} h_K^2 \le J_K(\xi_1, \xi_2) \le C_2 h_K^2. \tag{2.7}$$

Here $J_K(\xi_1, \xi_2)$ is the Jacobian of the transformation described by (2.5), and the constants C_1 and C_2 depend at most on r.

Remark 2.1. *The sign of $J_K(\xi_1, \xi_2)$ changes if the local ordering of the nodes is taken in the opposite direction. Therefore, it is assumed without loss of generality that $J_K(\xi_1, \xi_2) > 0$ for any $K \in \mathcal{T}_h$.*

Condition I. For any two adjacent element K, K',

$$\left| \frac{1}{J_K} \frac{\partial x_l^K}{\partial \xi_1} \frac{\partial x_k^K}{\partial \xi_2} - \frac{1}{J_{K'}} \frac{\partial x_l^{K'}}{\partial \xi_1} \frac{\partial x_k^{K'}}{\partial \xi_2} \right| \leq Ch,$$

$l, k = 1, 2$, for some constant C dependent at most on r.

Remark 2.2. *Condition I is of a different nature than (2.7) in Definition 1 in that it involves two adjacent elements. This is satisfied, for example, if the meshes consist of elements that differ little from parallelograms having sides nearly parallel to the sides of its neighbors (cf. [28] Remark 6 (p. 678)). For a more thorough explanation of r-strongly regular, the reader is referred to Remarks 1, 2, and 3 of [28].*

For any function f defined on K, we define $\hat{f}$ on $\hat{K}$ via

$$\hat{f}(\xi_1, \xi_2) = f(x_1^K(\xi_1, \xi_2), x_2^K(\xi_1, \xi_2)).$$

We also define for $v \in H^1(\Omega)$,

$$|v|_h^2 = \sum_{K \in \mathcal{T}_h} \sum_{j=1}^{r^2} \hat{A}_j J_K(\hat{G}_j) \left(\left[\widehat{\frac{\partial v}{\partial x_1}}(\hat{G}_j) \right]^2 + \left[\widehat{\frac{\partial v}{\partial x_2}}(\hat{G}_j) \right]^2 \right), \tag{1.1}$$

where $\hat{A}_j > 0$ are the integration Gauss–Legendre weights. From [28], $|\cdot|_h$ is a norm on $S_0^{r,h}$. The finite element solution $u_h \in S_0^{r,h}$ satisfies

$$a(u_h, v) = (f, v), \qquad \forall v \in S_0^{r,h}. \tag{1.2}$$

The optimal error for $u - u_h$ in the H^1-norm is (cf. [29])

$$\|u - u_h\|_{1,\Omega} \leq Ch^r \|u\|_{r+1,\Omega}.$$

However, P. Lesaint and M. Zlámal (cf., [27] Theorem 4.1) have shown that the arithmetic mean of the gradient error at the Gaussian points is superconvergent and, more precisely, they proved the following theorem.

Theorem 1.1. *Let $\mathcal{T}_h$ be r-strongly regular and satisfy Condition I, and let u and u_h denote the solutions of (2.3) and (2.9), respectively. In addition, suppose $u \in H^{r+3}(\Omega)$. Then*

$$|u - u_h|_h \leq Ch^{r+1} \|u\|_{r+3,\Omega},$$

where C is a constant independent of h and u.

Remark 2.3. *In (2.9), an approximate bilinear form $a_h(u_h, v)$ could be used by employing approximate quadrature formulae to effect the integration. The resulting analysis would follow that of P. Lesaint and M. Zlámal in [27], and would be more complicated, but the conclusions would be precisely those in the present work.*

For the special case of the Laplacian with rectangular meshes, the following holds (cf. [21], Lemma 3.1).

Theorem 1.2. *Let u and u_h denote the solutions of (2.3) and (2.9), respectively. In addition, suppose $u \in W^{r+3,\infty}(\Omega)$. Then for the special case $a_{11}(\mathbf{x}) = a_{22}(\mathbf{x}) = 1, a_{12}(\mathbf{x}) = a_{21}(\mathbf{x}) = 0$ with Ω partitioned into quasi-uniform rectangles, there is a constant C independent of h and u such that,*

$$|\nabla(u - u_h)(G_i^K)| \leq Ch_K^{r+1}\|u\|_{r+3,\infty}, \qquad i = 1, 2, \ldots, r^2,$$

for any $K \in \mathcal{T}_h$.

An essential condition of Theorem 2.1 or Theorem 2.2 is the regularity assumption on u which is true whenever a_{ij}, b, f, and Γ are sufficiently smooth. But in many practical problems, Γ is not smooth. A typical situation is the polygonal boundary. In this case, u has the corner singularity in general. For example, the solution of the Poisson equation with a smooth f and Dirichlet boundary conditions resides in $H^{1+\pi/\omega-\epsilon}(\Omega)$ for any $\epsilon > 0$, where ω is the largest inner angle on Γ. Hence, for the square domain $(\omega = \pi/2), u$ is an element of $H^{3-\epsilon}(\Omega)$, and, for the L-shaped domain, $u \in H^{5/3-\epsilon}(\Omega)$ in general. We see that for polygon domain, the regularity assumption $u \in H^{r+3}(\Omega)(u \in W^{r+3,\infty}(\Omega))$ limits the application of Theorem 2.1 (Theorem 2.2) to some special situations, as in the example of Lesaint and Zlámal where Ω is the unit square and $u(x, y) = x(1-x)y(1-y)(1+2x+2y)$. See [27], Section 6; also see [26], Section 3.

For problems with singularities, more sophisticated analysis is required. This article considers only the situation when the global regularity assumption in Theorem 2.1 or in Theorem 2.2 is satisfied.

III. Z–Z PATCH RECOVERY

Denote by $\mathcal{N}_h$ the set of element nodes, and define $\overset{\circ}{\mathcal{N}}_h = \mathcal{N}_h \setminus \Gamma$. Clearly $\overset{\circ}{\mathcal{N}}_h$ is the set of element nodes that is not on the boundary. Denote either $\frac{\partial u}{\partial x}$ or $\frac{\partial u}{\partial y}$ as σ, and $\frac{\partial u_h}{\partial x}$ or $\frac{\partial u_h}{\partial y}$ as σ_h. It is well known that σ_h does not possess interelement continuity and has lower accuracy at element boundaries. The Z–Z patch recovery seeks to find, via a least-squares argument, a function $\sigma_r(x, y)$ defined on an element patch $S(O)$ surrounding the particular assembly node $O \in \overset{\circ}{\mathcal{N}}_h$, where (x, y) is the local coordinate relative to O. We represent $\sigma_r(x, y) = P(x, y)^T\vec{a}$ with either

$$P(x, y)^T = (1, x, y, \ldots, x^r, x^{r-1}y, \ldots, y^r),$$

$$\vec{a} = (a_1, \ldots, a_m)^T, \qquad m = \sum_{i+j=0}^{r} 1 = \frac{1}{2}(r+1)(r+2),$$

$$P(x, y)^T = (1, x, y, \ldots, x^r y^r),$$

$$\vec{a} = (a_1, \ldots, a_m)^T, \qquad m = \sum_{i=0}^{r}\sum_{j=0}^{r} 1 = (r+1)^2.$$

In other words, σ_r is made to fit σ_h at the Gaussian points in $S(O)$ in a least-squares sense.

In order to express matters more precisely, let us denote the Gaussian points as $\{(x_i, y_i)$: $i = 1, \ldots, n\}$; then the aforementioned least-squares fitting is to minimize

$$\begin{aligned} F(\hat{a}) &= \sum_{i=1}^{n} (\sigma_h(x_i, y_i) - \sigma_r(x_i, y_i))^2 \\ &= \sum_{i=1}^{n} (\sigma_h(x_i, y_i) - P(x_i, y_i)^T \vec{a})^2 \\ &= \sum_{i=1}^{n} (\sigma_h(x_i, y_i) - \hat{P}(\xi_i, \eta_i)^T \hat{a})^2, \end{aligned}$$

where

$$\hat{P}(\xi, \eta)^T = (1, \xi, \eta, \ldots, \xi^r, \xi^{r-1}\eta, \ldots, \eta^r), \qquad h\xi = x, h\eta = y,$$

$$\hat{a} = (a_1, a_2 h, a_3 h, a_4 h^2, \ldots, a_m h^r)^T.$$

We shall consider only the case where $P(x, y)$ is a linear combination of basis elements for polynomials of degree no more than r. We omit the discussion for the other case, whereby $P(x, y)$ is as a linear combination of basis elements for polynomials of degree r in each variable separately. The treatment for the latter case is similar and yields the same order of approximation.

The minimum $\hat{a}$ of F must of necessity satisfy

$$\sum_{i=1}^{n} \hat{P}(\xi_i, \eta_i) \hat{P}(\xi_i, \eta_i)^T \hat{a} = \sum_{i=1}^{n} \hat{P}(\xi_i, \eta_i) \sigma_h(x_i, y_i).$$

In matrix form, this condition becomes

$$A^T A \hat{a} = A^T \vec{b}_h, \tag{3.1}$$

where

$$A = \begin{pmatrix} 1 & \xi_1 & \eta_1 & \cdots & \xi_1^r & \xi_1^{r-1}\eta_1 & \cdots & \eta_1^r \\ 1 & \xi_2 & \eta_2 & \cdots & \xi_2^r & \xi_2^{r-1}\eta_2 & \cdots & \eta_2^r \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & \xi_n & \eta_n & \cdots & \xi_n^r & \xi_n^{r-1}\eta_n & \cdots & \eta_n^r \end{pmatrix},$$

$$\vec{b}_h = (\sigma_h(x_1, y_1), \ldots, \sigma_h(x_n, y_n))^T.$$

(The purpose of solving for $\hat{a}$ instead of $\vec{a}$ is to eliminate the numerical ill-conditioning of the system (3.1) for higher r. Obviously, $\hat{a}$ is independent of h.)

Assume that

$(A1)$ $$\text{Rank } (A) = \tfrac{1}{2}(r+1)(r+2) = m.$$

This is a reasonable assumption and is satisfied if enough Gaussian points are selected, or, equivalently, if the element patch is enlarged. In particular, for rectangular meshes, a natural element patch is a set of four elements that share a common vertex. It is easy to see that in such an element patch, there are $n = 4r^2$ Gaussian points.

Remark 3.1. *Since the set of* $4r^2$ *Gaussian point is* $\hat{Q}(2r-1)$*-unisolvent and, for* $r \geq 1, \hat{P}(r) \subset \hat{Q}(2r-1)$, *the element patch that contains 4 elements with a common vertex satisfies (A1) automatically. Also, we see that in order to satisfy the assumption (A1), we do not have to use all* $4r^2$ *Gaussian points. The minimum requirement is* $(r+1) \times (r+1)$ *Gaussian points.*

By assumption $(A1)$, $A^T A$ is invertible. Thus, (3.1) can be uniquely solved, and the value of σ_r at the assembly node O is

$$\sigma_r(0,0) = P(0,0)^T \vec{a} = \hat{P}(0,0)^T \hat{a} = a_1 = \vec{\alpha}^T \vec{b}_h, \tag{3.2}$$

where $\vec{\alpha} = (\alpha_1, \ldots, \alpha_n)^T$ is the first row of the matrix $(A^T A)^{-1} A^T$, which depends only on r and the relative positions of (ξ_i, η_i)'s. By (3.2), the least-squares fitting process defines a nodal recovery operator R such that

$$(R\sigma_h)(O) = \vec{\alpha}(O)^T \vec{b}_h(O),$$

for all $O \in \overset{\circ}{\mathcal{N}}_h$. Clearly, this operator can be applied to any function that is well defined at the Gaussian points. In particular, when applied to $\sigma = [\frac{\partial u}{\partial x}]$ or to $\sigma = [\frac{\partial u}{\partial y}]$, we get

$$(R\sigma)(O) = \vec{\alpha}(O)^T \vec{b}(O), \qquad \vec{b}(O) = (\sigma(x_1^{(O)}, y_1^{(O)}), \ldots, \sigma(x_n^{(O)}, y_n^{(O)}))^T.$$

By Hölder's inequality, for any $\sigma \in H^{r+1}(\Omega) + S^{r,h}, r \geq 1$,

$$|(R\sigma)(O)| \leq \|\vec{\alpha}(O)\|_1 \|\vec{b}(O)\|_\infty, \tag{3.3}$$

$$|(R\sigma)(O)| \leq \|\vec{\alpha}(O)\|_2 \|\vec{b}(O)\|_2. \tag{3.4}$$

Here $\|\cdot\|_1, \|\cdot\|_2$, and $\|\cdot\|_\infty$ are the l^1, Euclidean, and l^∞-norms for $\mathbb{R}^n$.

The following theorem states the interpolation property of the nodal recovery operator R.

Theorem 3.1. *Let (A1) be satisfied; then for any* $\sigma \in W^{r+1,\infty}(\Omega)$ *and* $O \in \overset{\circ}{\mathcal{N}}_h$,

$$|(\sigma - R\sigma)(O)| \leq C(r) h^{r+1} |\sigma|_{r+1,\infty,S(O)}, \tag{3.5}$$

with

$$C(r) = \frac{2^{r+1}}{(r+1)!} \sum_{i=1}^{n} |\alpha_i|.$$

Proof. To simplify the notation, we suppress the dependent index "O". Since $\vec{\alpha}^T$ is the first row of $(A^T A)^{-1} A^T$, then

$$\vec{\alpha}^T A = (1, 0, \ldots, 0), \tag{3.6}$$

and hence

$$\sum_{i=1}^{n} \alpha_i = 1, \tag{3.7}$$

$$\sum_{i=1}^{n} \alpha_i x_i^l y_i^{k-l} = h^k \sum_{i=1}^{n} \alpha_i \xi_i^l \eta_i^{k-l} = 0, \qquad 1 \leq k \leq r, l = 0, 1, \ldots, k. \tag{3.8}$$

By the Taylor expansion of σ in the local coordinates, we see

$$
\begin{aligned}
(R\sigma)(O) &= \sum_{i=1}^{n} \alpha_i \sigma(x_i, y_i) \\
&= \sum_{i=1}^{n} \alpha_i \left(\sum_{k=0}^{r} \frac{1}{k!} \left(x_i \frac{\partial}{\partial x} + y_i \frac{\partial}{\partial y} \right)^k \sigma(0,0) \right. \\
&\quad \left. + \frac{1}{r!} \int_0^1 (1-t)^r \left(x_i \frac{\partial}{\partial x} + y_i \frac{\partial}{\partial y} \right)^{r+1} \sigma(tx_i, ty_i) dt \right) \\
&= \sum_{k=0}^{r} \frac{1}{k!} \sum_{l=0}^{k} \binom{k}{l} \frac{\partial^k \sigma}{\partial x^l \partial y^{k-l}}(0,0) \sum_{i=1}^{n} \alpha_i x_i^l y_i^{k-l} \\
&\quad + \frac{1}{r!} \int_0^1 (1-t)^r \sum_{i=1}^{n} \alpha_i \left(x_i \frac{\partial}{\partial x} + y_i \frac{\partial}{\partial y} \right)^{r+1} \sigma(tx_i, ty_i) dt \\
&= \sigma(O) + \frac{1}{r!} \sum_{l=0}^{r+1} \binom{r+1}{l} \sum_{i=1}^{n} \alpha_i x_i^l y_i^{r+1-l} \\
&\quad \times \int_0^1 (1-t)^r \frac{\partial^{r+1} \sigma}{\partial x^l \partial y^{r+1-l}}(tx_i, ty_i) dt,
\end{aligned} \tag{3.9}
$$

when applying (3.7) and (3.8).

To estimate the residual term, note that since $|x_i| < h, |y_i| < h$, and

$$
\sum_{l=0}^{r+1} \binom{r+1}{l} = 2^{r+1}, \qquad \int_0^1 (1-t)^r dt = \frac{1}{r+1},
$$

we have

$$
\begin{aligned}
&\left| \frac{1}{r!} \sum_{l=0}^{r+1} \binom{r+1}{l} \sum_{i=1}^{n} \alpha_i x_i^l y_i^{r+1-l} \int_0^1 (1-t)^r \frac{\partial^{r+1} \sigma}{\partial x^l \partial y^{r+1-l}}(tx_i, ty_i) dt \right| \\
&\quad \le \frac{2^{r+1}}{(r+1)!} \sum_{i=1}^{n} |\alpha_i| h^{r+1} |\sigma|_{r+1,\infty,S(O)}.
\end{aligned} \tag{3.10}
$$

The conclusion follows by combining (3.9) and (3.10). ∎

Corollary. *If $S(O)$ contains 4 uniform rectangular elements that have O as a common vertex, then for any $\sigma \in W^{r+2,\infty}(\Omega)$ with r even,*

$$
|(\sigma - R\sigma)(O)| \le C(r+1) h^{r+2} |\sigma|_{r+2,\infty,S(O)}, \tag{3.11}
$$

where $C(r)$ is defined as same as in Theorem 3.1.

Proof. By hypothesis, $S(O)$ contains 4 uniform rectangular elements, and assumption $(A1)$ is automatically satisfied (see Remark 4). Furthermore, we know that $(-x_j, -y_j) \in S(O)$ whenever $(x_j, y_j) \in S(O)$. Hence, if $x_i^l y_i^{k-l}$ appears in (3.9), there exists a j such that $1 \le j \le n$ and $x_j^l y_j^{k-l} = (-x_i)^l (-y_i)^{k-l}$ also appears in (3.9), $1 \le k \le r+1, l = 0, 1, \dots, k$. Moreover, they have the same coefficient by symmetry, i.e., $\alpha_j = \alpha_i$. Consequently,

$$
\sum_{i=1}^{n} \alpha_i x_i^l y_i^{r+1-l} = \sum_{i=1}^{n} \alpha_i (-x_i)^l (-y_i)^{r+1-l} = (-1)^{r+1} \sum_{i=1}^{n} \alpha_i x_i^l y_i^{r+1-l}.
$$

When r is an even number,

$$\sum_{i=1}^{n} \alpha_i x_i^l y_i^{r+1-l} = -\sum_{i=1}^{n} \alpha_i x_i^l y_i^{r+1-l} = 0. \tag{3.12}$$

Following the same arguments as used in Theorem 3.1, we will have a residual term

$$\frac{1}{(r+1)!}\int_0^1 (1-t)^{r+1}\sum_{i=1}^{n}\alpha_i\left(x_i\frac{\partial}{\partial x}+y_i\frac{\partial}{\partial y}\right)^{r+2}\sigma(tx_i,ty_i)dt.$$

The estimate results in exactly the same manner as in Theorem 3.1 by changing r to $r+1$. ∎

Next a stability result for the recovery operator R will be shown. Let us denote $\mathbf{R} = (R, R)$, $\overset{\circ}{\mathcal{N}}_h = \{O_j\}_{j=1}^{N_h}$, where N_h is the number of internal element nodes, and assume that

$(A2)$ $\qquad \max_{1\le j\le N_h}\|\vec{\alpha}(O)\|_1 \le C_{1,BOUND}, \qquad \max_{1\le j\le N_h}\|\vec{\alpha}(O)\|_2 \le C_{2,BOUND},$

where $C_{1,BOUND}$ and $C_{2,BOUND}$ are constants independent of N_h or h. The following assumption is the standard quasi-uniformity of the meshes:

$(A3)$ $\qquad h \le \tau \min_K h_k, \qquad$ for some $\tau > 0$.

Remark 3.2. *For regular partitions, (A2) is a reasonable assumption. For example, when the partition is uniform, all* $\|\vec{\alpha}(O_j)\|_1$ *(or* $\|\vec{\alpha}(O_j)\|_2$*),* $j = 1, 2, \ldots, N_h$*, are the same and will remain so if the uniform mesh refinement is applied. If the partition is not uniform, but refinement follows in a uniform way,* $\|\vec{\alpha}(O_j)\|_1$ *(or* $\|\vec{\alpha}(O_j)\|_2$*) will be uniformly bounded independently of* N_h *(or* h*).*

As a consequence of $(A3)$ and (2.7), there exists $C_3 > 0$ independent of h such that

$$C_3^{-1}\hat{A}_j J_K(\hat{G}_j) \le \frac{1}{N_h} \le C_3\hat{A}_j J_K(\hat{G}_j), \qquad j = 1, 2, \ldots, r^2. \tag{3.13}$$

Theorem 3.2. *Under the assumptions of (A1), (A2), (A3), and (2.7),*

$$\frac{1}{N_h}\sum_{j=1}^{N_h}|\mathbf{R}\nabla(u-u_h)(O_j)| \le C_{STAB}|u-u_h|_h,$$

with $C_{STAB} = \sqrt{C_3 M}C_{2,BOUND}$*, where* M *is the maximum number of times a Gaussian point is repeatedly used for the recovery.*

Proof. Set $O = O_j$, and $\sigma = \frac{\partial}{\partial x}(u-u_h), \sigma = \frac{\partial}{\partial y}(u-u_h)$ in (3.4), respectively, and we have

$$\begin{aligned}\left|R\frac{\partial}{\partial x}(u-u_h)(O_j)\right| &\le \|\vec{\alpha}(O_j)\|_2\|\vec{b}_x(O_j)\|_2,\\ \left|R\frac{\partial}{\partial y}(u-u_h)(O_j)\right| &\le \|\vec{\alpha}(O_j)\|_2\|\vec{b}_y(O_j)\|_2,\end{aligned}$$

where

$$\vec{b}_x(O_j) = \left(\frac{\partial}{\partial x}(u-u_h)(x_1^{(j)},y_1^{(j)}),\ldots,\frac{\partial}{\partial x}(u-u_h)(x_n^{(j)},y_n^{(j)})\right)^T,$$

$$\vec{b}_y(O_j) = \left(\frac{\partial}{\partial y}(u-u_h)(x_1^{(j)},y_1^{(j)}),\ldots,\frac{\partial}{\partial y}(u-u_h)(x_n^{(j)},y_n^{(j)})\right)^T.$$

From $(A2)$,

$$|\mathbf{R}\nabla(u-u_h)(O_j)|^2 = \left[R\frac{\partial}{\partial x}(u-u_h)(O_j)\right]^2 + \left[R\frac{\partial}{\partial y}(u-u_h)(O_j)\right]^2 \leq C_{2,BOUND}^2 \sum_{i=1}^{n} |\nabla(u-u_h)(x_i^{(j)}, y_i^{(j)})|^2. \tag{3.14}$$

Consequently,

$$\begin{aligned} &\frac{1}{N_h}\sum_{j=1}^{N_h} |\mathbf{R}\nabla(u-u_h)(O_j)|^2 \\ &\leq C_{2,BOUND}^2 \frac{M}{N_h} \sum_{K\in\mathcal{T}_h}\sum_{j=1}^{r^2}\left(\left[\frac{\partial}{\partial x}(u-u_h)(G_j^K)\right]^2 + \left[\frac{\partial}{\partial y}(u-u_h)(G_j^K)\right]^2\right) \\ &\leq C_{2,BOUND}^2 C_3 M \sum_{K\in\mathcal{T}_h}\sum_{j=1}^{r^2} \hat{A}_j J_K(\hat{G}_j)\left(\left[\widehat{\frac{\partial u}{\partial x}}(\hat{G}_j) - \widehat{\frac{\partial u_h}{\partial x}}(\hat{G}_j)\right]^2\right. \\ &\quad + \left.\left[\widehat{\frac{\partial u}{\partial y}}(\hat{G}_j) - \widehat{\frac{\partial u_h}{\partial y}}(\hat{G}_j)\right]^2\right) \\ &= C_{2,BOUND}^2 C_3 M |u-u_h|_h^2, \end{aligned} \tag{3.15}$$

by virtue of (3.14), (3.13), and the definition of $|\cdot|_h$. From Hölder's inequality,

$$\frac{1}{N_h}\sum_{j=1}^{N_h}|\mathbf{R}\nabla(u-u_h)(O_j)| \leq \left(\frac{1}{N_h}\sum_{j=1}^{N_h}|\mathbf{R}\nabla(u-u_h)(O_j)|^2\right)^{1/2} \leq C_{STAB}|u-u_h|_h.$$

∎

We now show the two main results of this section.

Theorem 3.3. *Let $\mathcal{T}_h$ be r-strongly regular and satisfy Condition I, assumptions (A2), and (A3). Let the element patch be chosen so that (A1) is satisfied. Then*

$$\frac{1}{N_h}\sum_{j=1}^{N_h}|(\nabla u - \mathbf{R}\nabla u_h)(O_j)| \leq \sqrt{2}C(r)h^{r+1}|u|_{r+2,\infty} + C_{STAB}Ch^{r+1}\|u\|_{r+3,\Omega},$$

where C is the same constant as in Theorem 2.1 and $\{O_j\}_{j=1}^{N_h}$ is the set of all interior element nodes.

Proof. By the triangle inequality,

$$\|(\nabla u - \mathbf{R}\nabla u_h)(O_j)| \leq |(\nabla u - \mathbf{R}\nabla u)(O_j)| + |(\mathbf{R}\nabla(u-u_h)(O_j)|. \tag{3.16}$$

Using Theorem 3.1 twice by setting $O = O_j, \sigma = \frac{\partial u}{\partial x}$ and $\sigma = \frac{\partial u}{\partial y}$, respectively, we have

$$\left|\left(\frac{\partial u}{\partial x} - R\frac{\partial u}{\partial x}\right)(O_j)\right| \leq C(r)h^{r+1}\left|\frac{\partial u}{\partial x}\right|_{r+1,\infty,S(O_j)},$$

$$\left|\left(\frac{\partial u}{\partial y} - R\frac{\partial u}{\partial y}\right)(O_j)\right| \leq C(r)h^{r+1}\left|\frac{\partial u}{\partial y}\right|_{r+1,\infty,S(O_j)}.$$

Hence,

$$|(\nabla u - \mathbf{R}\nabla u)(O_j)| \leq \sqrt{2}C(r)h^{r+1}|u|_{r+2,\infty}. \tag{3.17}$$

Summing up (3.16) over $j = 1, 2 \cdots, N_h$, and dividing by N_h, we have

$$\begin{aligned}
&\frac{1}{N_h}\sum_{j=1}^{N_h}|(\nabla u - \mathbf{R}\nabla u_h)(O_j)| \\
&\leq \sqrt{2}C(r)h^{r+1}|u|_{r+2,\infty} + C_{STAB}|u - u_h|_h \\
&\leq \sqrt{2}C(r)h^{r+1}|u|_{r+2,\infty} + C_{STAB}Ch^{r+1}\|u\|_{r+3,\Omega}.
\end{aligned}$$

Here we have used (3.17), Theorem 3.2, and Theorem 2.1. ∎

Theorem 3.4. *Under the same assumptions as in Theorem 2.2, we have for any interior element node $O_j, j = 1, 2, \ldots, N_h$,*

$$|(\nabla u - \mathbf{R}\nabla u_h)(O_j)| \leq \sqrt{2}(C(r)|u|_{r+2,\infty} + C_{1,BOUND}C\|u\|_{r+3,\infty})h^{r+1},$$

where C is the same constant appearing in Theorem 2.2.

Proof. With uniform rectangular meshes, (A1) is automatically satisfied (see Remark 3.1). By using (3.3), we have as in (3.14),

$$\begin{aligned}
|\mathbf{R}\nabla(u - u_h)(O_j)| &\leq C_{1,BOUND}\sqrt{2}\max_{1\leq i\leq n}|\nabla(u - u_h)(x_i^{(j)}, y_i^{(j)})| \\
&\leq \sqrt{2}C_{1,BOUND}Ch^{r+1}\|u\|_{r+3,\infty},
\end{aligned} \tag{3.18}$$

where C is the same constant as in Theorem 2.2. Inserting (3.17) and (3.18) into (3.16), we obtain the conclusion. ∎

Conclusion and Remarks

1. Theorem 3.3 states that the arithmetic mean of the absolute error of the recovered gradient at the nodal points inside Ω is superconvergent for a class of regular curved quadrilateral meshes. Theorem 3.4 states that, for uniform rectangular meshes, the recovered gradient is superconvergent in the maximum norm at every nodal point inside Ω. This has been observed by practical computation [26].
2. Theorem 3.1 describes the interpolation properties of the Z-Z patch recovery. It clearly shows that the interpolation feature of the Z-Z patch recovery is insensitive to the type of meshes. A parallel result of Theorem 3.1 holds for triangular meshes. These results partially explain the robustness of the method under extreme distortions of the mesh.
3. Another feature of the analysis in this article is the concrete knowledge of the constants multiplying the rate of convergence. This is very important in the practical design of the finite element method.
4. The generalization of the analysis to three dimensions is straightforward. Here,

$$A = \begin{pmatrix}
1 & \hat{x}_1 & \hat{y}_1 & \hat{z}_1 & \cdots & \hat{x}_1^r & \hat{x}_1^{r-1}\hat{y}_1\hat{x}_1^{r-1}\hat{z}_1 & \cdots & \hat{z}_1^r \\
1 & \hat{x}_2 & \hat{y}_2 & \hat{z}_2 & \cdots & \hat{x}_2^r & \hat{x}_2^{r-1}\hat{y}_2\hat{x}_2^{r-1}\hat{z}_2 & \cdots & \hat{z}_2^r \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & & \vdots \\
1 & \hat{x}_n & \hat{y}_n & \hat{z}_n & \cdots & \hat{x}_n^r & \hat{x}_n^{r-1}\hat{y}_n\hat{x}_n^{r-1}\hat{z}_n & \cdots & \hat{z}_n^r
\end{pmatrix},$$

$$m = \sum_{i+j+k=0}^{r} 1 = \sum_{k=0}^{r} \sum_{i+kj=0}^{k} 1 = \sum_{k=0}^{r} \frac{1}{2}(k+1)(k+2) = \frac{1}{6}r^3 + r^2 + \frac{11}{6}r + 1,$$

and the coefficient vector $\vec{\alpha}^T$ is still the first row of the pseudo-inverse of A.

5. We would like to point out that when piecewise bilinear elements are used and quadrilaterals are produced by a bisection scheme of mesh subdivisions, the finite element triangulation is 1-strongly regular (see Definition 1) and satisfies Condition I, and Assumptions (A1), (A2), and (A3). Consequently, the Z–Z recovery technique will result in a superconvergent gradient field.

IV. ULTRA-CONVERGENCE ANALYSIS IN 1-D

The authors of [26] announced an $O(h^4)$ convergent rate for the nodal values of the recovered derivative with quadratic elements when uniform meshes are used. Since this rate is two orders higher than the optimal rate, it is then termed "ultra-convergence." In this section, we provide a theoretical justification of this celebrated phenomenon for the following two-point boundary-value problem:

$$-u'' + b(x)u = f \qquad \text{in } I = (0,1), \tag{4.1}$$

$$u(0) = u(1) = 0. \tag{4.2}$$

We assume that b and f are sufficiently smooth for our analysis. The weak formulation of (4.1), (4.2) is to find $u \in H_0^1(I)$ such that

$$(u', v') + (bu, v) = (f, v) \qquad \forall v \in H_0^1(I). \tag{4.3}$$

Let $\mathcal{T}_h, 0 < h < 1/2$, be a sequence of uniform subdivisions of $\bar{I}$,

$$\mathcal{T}_h = \{x_i\}_{i=0}^N, x_i = ih, h = 1/N;$$

let $I_i = (x_{i-1}, x_i)$, and set

$$S_h^0 = \{v \in H_0^1(I), \qquad v|_{I_i} \in P_2(I_i)\}.$$

We see that S_h^0 is the space of piecewise quadratic polynomials on I under the uniform subdivision $\mathcal{T}_h$.

The finite element solution of (4.3) is to find $u_h \in S_h^0$ such that

$$(u_h', v') + (bu_h, v) = (f, v) \qquad \forall v \in S_h^0. \tag{4.4}$$

The first step of ultra-convergence analysis is to reduce (4.4) to a simpler problem. Subtracting (4.4) from (4.3) yields

$$(u' - u_h', v') + (b(u - u_h), v) = 0 \qquad \forall v \in S_h^0. \tag{4.5}$$

Let $\tilde{u}_h \in S_h^0$ be given by

$$(u' - \tilde{u}_h', v') = 0 \qquad \forall v \in S_h^0. \tag{4.6}$$

Then we have the following "ultra-approximation" result between u_h and $\tilde{u}_h$ (see [30] Remark 1.3.1).

Lemma 4.1. *Let $u_h, \tilde{u}_h$ satisfy (4.5), (4.6), respectively. Then*

$$\|u_h' - \tilde{u}_h'\|_{L_\infty(I)} \le Ch^4\|u\|_{W^5_\infty(I)}, \tag{4.7}$$

where C is a constant independent of h and u.

By virtue of Lemma 4.1, we can reduce our discussion to a simple case:

$$-u'' = f \qquad \text{in } I = (0,1),$$
$$u(0) = u(1) = 0.$$

or

$$(u', v') = (f, v) \qquad \forall v \in H_0^1(I), \tag{4.8}$$

since the finite element solution of (4.8) satisfies (4.6).

In the following, we will construct the finite element solution $u_h \in S_h^0$ for (4.8) and prove the ultra-convergence property of Ru_h' at the nodal points. Here Ru_h' is the recovered derivative by the Z Z patch recovery technique.

We characterize S_h^0 by the following basis functions:

$$S_h^0 = \text{Span}\,\{N_i(x), i = 1, \dots, N-1; \qquad \phi_j(x), j = 1, \dots, N\},$$

where

$$N_i(x) = \begin{cases} 1 + (x - x_i)/h & x \in I_i, \\ 1 + (x_i - x)/h & x \in I_{i+1}, \\ 0 & \text{otherwise}; \end{cases} \quad \phi_j(x) = \begin{cases} (x - x_{j-1})(x_j - x)/h^2 & x \in I_j, \\ 0 & \text{otherwise}. \end{cases}$$

Observe that $N_i(x_j) = \delta_{ij}, \phi_j(x_i) = 0$, and

$$\int_0^1 N_i'(x)\phi_j'(x)dx = 0, \tag{4.9}$$

$$\int_0^1 \phi_i'(x)\phi_j'(x)dx = 0, \qquad i \ne j. \tag{4.10}$$

It is well known that the finite element solution of the variational problem (4.8) is exact at the nodal points, i.e., $u_h(x_i) = u(x_i), i = 0, 1, \dots, N$. We then have on I_i,

$$u_h(x) = u(x_{i-1})N_{i-1}(x) + u(x_i)N_i(x) + c_i\phi_i(x), \tag{4.11}$$

where

$$c_i = (f, \phi_i)/(\phi_i', \phi_i').$$

Associated with any interior node $x_i, i = 1, \dots, N-1$, there is an element patch $J_i = (x_i - h, x_i + h)$, and a linear mapping F_i from $\hat{I} = (-1, 1)$ onto J_i defined by $x = x_i + h\xi$. Given any function v on J_i, we define

$$\hat{v} = v \circ F_i, \qquad \text{or } \hat{v}(\xi) = v(F_i(\xi)) = v(x_i + h\xi).$$

Now, consider

$$u'(x_i) - Ru_h'(x_i) = \langle u' - Ru_h', \delta_i\rangle = h\langle \widehat{u' - Ru_h'}, \hat{\delta}_i\rangle = hE(\hat{u}'). \tag{4.12}$$

Here $\hat{\delta}_i = \delta \circ F_i$ and δ is the Delta function. Obviously, $E(\hat{u}')$ is a linear functional that is bounded in $W^4_\infty(\hat{I})$. We shall show that $E(\hat{u}')$ vanishes when $\hat{u}'$ is a polynomial of degree no more than 3. To this end, we first introduce the following result from [31].

Lemma 4.2. *Let u be the solution of (4.8), and let u_h be its finite element approximation on S_h^0. Assume that u is a polynomial of degree no more than 3 on J_i, then $Ru_h' = u'$ on J_i.*

Note that u_h' is a piecewise linear function on J_i with a possible jump at x_i, but Ru_h' is a quadratic polynomial on J_i. We see that $E(\hat{u}')$ vanishes when $\hat{u}'$ is a quadratic polynomial.

Next, we examine the case when

$$u(x) = a\left(\frac{x - x_{i-1}}{h}\right)^2 \left(\frac{x - x_{i+1}}{h}\right)^2, \qquad a \neq 0,$$

on J_i. Note that $u'(x_i) = 0$. Since $u(x)$ is symmetric with respect to x_i, we have

$$c_i = (-u'', \phi_i)/(\phi_i', \phi_i') = (-u'', \phi_{i+1})/(\phi_{i+1}', \phi_{i+1}') = c_{i+1}.$$

Recall (4.11), and we have on J_i,

$$u_h(x) = aN_i(x) + \begin{cases} c_i\phi_i(x) & x \in I_i \\ c_i\phi_{i+1}(x) & x \in I_{i+1}. \end{cases}$$

$$u_h'(x) = \begin{cases} a/h + c_i\phi_i'(x) & x \in I_i \\ -a/h + c_i\phi_{i+1}'(x) & x \in I_{i+1}. \end{cases}$$

Observe that

$$\phi_i'(x_i - \tau) = -\phi_{i+1}'(x_i + \tau), \qquad 0 \le \tau \le h,$$

we then have

$$u_h'(x_i - \tau) = -u_h'(x_i + \tau), \qquad 0 \le \tau \le h.$$

Hence, by the Z-Z patch recovery procedure

$$\begin{aligned} Ru_h'(x_i) &= c_i\alpha\left[u_h'\left(x_i - \frac{h}{2}\left(1 - \frac{1}{\sqrt{3}}\right)\right) + u_h'\left(x_i + \frac{h}{2}\left(1 - \frac{1}{\sqrt{3}}\right)\right)\right] \\ &\quad + c_i\beta\left[u_h'\left(x_i - \frac{h}{2}\left(1 + \frac{1}{\sqrt{3}}\right)\right) + u_h'\left(x_i + \frac{h}{2}\left(1 + \frac{1}{\sqrt{3}}\right)\right)\right] \\ &= 0 = u'(x_i). \end{aligned} \tag{4.13}$$

Here

$$x_i \pm \frac{h}{2}\left(1 - \frac{1}{\sqrt{3}}\right), \qquad x_i \pm \frac{h}{2}\left(1 + \frac{1}{\sqrt{3}}\right)$$

are the four Gauss points on the element patch J_i and α, β are weights of the least-squares fitting.

Since any $u \in P_4(J_i)$ can be decomposed into

$$u(x) = a\left(\frac{x - x_{i-1}}{h}\right)^2 \left(\frac{x - x_{i+1}}{h}\right)^2 + w(x)$$

for some $a \in R^1$ and $w \in P_3(J_i)$, we see from Lemma 4.2 and (4.13) that

$$Ru_h'(x_i) = u'(x_i) \qquad \forall u \in P_4(J_i), \tag{4.14}$$

i.e., the linear functional $E(\hat{u}')$ vanishes for all $\hat{u}' \in P_3(\hat{I})$. Therefore, by the Bramble–Hilbert Lemma, we have

$$\begin{aligned}|E(\hat{u}')| &\le C\|\hat{\delta}_i\|_{W_1^0(\hat{I})}|\hat{u}'|_{W_\infty^4(\hat{I})}\\ &\le Ch^{-1}\|\delta_i\|_{W_1^0(J_i)}h^4|u'|_{W_\infty^4(J_i)}\\ &= Ch^3|u|_{W_\infty^5(J_i)}.\end{aligned} \tag{4.15}$$

Note that $\|\delta_i\|_{W_1^0(J_i)} = 1$.

Combining (4.12), (4.15) with Lemma 4.1, we have proved the following main result of this section.

Theorem 4.1. *Let u be the solution of problem (4.1), (4.2), let $u_h \in S_h^0$ be its finite element approximation, and let Ru_h' be the recovered derivative from the Z–Z patch recovery technique. Then*

$$|u'(x_i) - Ru_h'(x_i)| \le Ch^4(|u|_{W_\infty^5(J_i)} + \|u\|_{W_\infty^3(I)}), \tag{4.16}$$

where C is a constant independent of h and u.

Remark 4.1. Since the above argument is completely local, the result of Theorem 4.1 holds for globally nonuniform but locally uniform meshes. The only requirement is on the element patch, namely that two elements have the same length.

Remark 4.2. The conclusion for general two-point boundary-value problems

$$-(a_2(x)u')' - (a_1(x)u)' + a_0(x)u = f \qquad \text{in } I, \tag{4.17}$$

with the boundary condition (4.2) is unknown at present. The reason is that we have only the "super-approximation" between u_h' and $\tilde{u}_h'$ in this case, i.e.,

$$\|u_h' - \tilde{u}_h'\|_{L_\infty(I)} \le Ch^3\|u\|_{W_\infty^3(I)},$$

instead of (4.7). See [30], Theorem 1.3.1, for the details.

Remark 4.3. The generalization to the higher-order polynomials is possible. If the finite element subspace S_h^0 contains piecewise polynomials of degree $2r, (r \ge 1)$, then we conjecture that under local uniform meshes,

$$|u'(x_i) - Ru_h'(x_i)| \le Ch^{2r+2}(|u|_{W_\infty^{2r+3}(J_i)} + \|u\|_{W_\infty^{2r+1}(I)}),$$

where u is the solution of the two-point boundary problem (4.1), (4.2).

The authors are grateful to Professor Ivo Babuška and Dr. J. Z. Zhu for many helpful discussions. The work of these authors was partially supported by National Science Foundation Grant DMS-90-23063.

REFERENCES

1. M. Ainsworth and A. Craig, "*A posteriori* error estimators in the finite element method," *Numer. Math.* **60,** 429 (1992).

2. M. Ainsworth and J. T. Oden, "A unified approach to *a posteriori* error estimation using element residual methods," *Numer. Math.* **65,** 23 (1993).

3. I. Babuška and R. Rodríguez, "The problem of the selection of an *a posteriori* error indicator based on smoothing techniques," *Int. J. Numer. Methods Eng.* **36,** 539 (1993).

4. I. Babuška, T. Strouboulis, C. S. Upadhyay, S. K. Gangaraj, and K. Copps, "Validation of *a posteriori* error estimators by numerical approach," *Int. J. Numer. Methods Eng.* **37,** 1073 (1994).

5. I. Babuška, O. C. Zienkiewicz, J. Grago, and E. R. de A. Oliveira, Eds., *Accuracy Estimates and Adaptive Refinements in Finite Element Computations,* Wiley and Sons, New York, 1986.

6. R. E. Bank and A. Weiser, "Some *a posteriori* error estimators for elliptic partial differential equations," *Math. Comp.* **44,** 283 (1985).

7. R. E. Ewing, "*A posteriori* error estimation," in *Reliability in Computational Mechanics,* J. T. Oden, Ed., Elsevier Science Publishers B.V., North Holland, 1990, p. 59.

8. J. T. Oden, L. Demkowicz, W. Rachowicz, and T. A. Westermann, "Toward a universal h–p adaptive finite element strategy, Part 2. *A posteriori* error estimation," *Comp. Methods Appl. Mech. Engrg.* **77,** 113 (1989).

9. J. T. Oden, "Error estimation and control in computational fluid dynamics," in *The Mathematics of Finite Elements and Applications VIII,* J. R. Whiteman, Ed., MAFELAP 1993, Academic Press, London.

10. R. Verfürth, "A review of *a posteriori* error estimation and adaptive mesh-refinement techniques," *Teubner skripten zur Numerik,* preprint.

11. O. C. Zienkiewicz and J. Z. Zhu, "A simple error estimator and adaptive procedure for practical engineering analysis," *Internat. J. Numer. Methods Eng.* **24,** 337 (1987).

12. I. Babuška and W. C. Rheinboldt, "Error estimates for adaptive finite element computations," *SIAM J. Numer. Anal.* **15,** 736 (1978).

13. R. E. Ewing, R. D. Lazarov, and J. Wang, "Superconvergence of the velocity along the Gauss lines in mixed finite element methods," *SIAM Numer. Anal.* **28,** 1015 (1991).

14. M. Křížek and P. Neittaanmäki, "On superconvergence techniques," *Acta Appl. Math.* **9,** 175 (1987).

15. L. B. Wahlbin, "Local behavior in finite element methods," in *Handbook of Numerical Analysis, Vol. II, Finite Element Methods* (Part 1), P. G. Ciarlet and J. L. Lions, Eds., Elsevier Science Publishers B.V., North-Holland, 1991, p. 353.

16. J. R. Whiteman and G. Goodsell, "A survey of gradient superconvergence for finite element approximations to second order elliptic problems on triangular and tetrahedral meshes," in *The Mathematics of Finite Elements and Applications VII,* J. R. Whiteman, Ed., MAFELAP 1990, Academic Press, London, p. 55.

17. Q. D. Zhu and Q. Lin, *Superconvergence Theory of the Finite Element Method,* Hunan Science Press, China, 1989.

18. E. Hinton and J. S. Campbell, "Local and global smoothing of discontinuous finite element functions using a least square method," *Internat. J. Numer. Meth. Eng.* **8,** 461 (1974).

19. J. T. Oden and H. J. Brauchli, "On the calculation of consistent stress distributions in finite element applications," *Internat. J. Numer. Meth. Eng.* **3,** 317 (1971).

20. Q. Lin and J. R. Whiteman, "Superconvergence of recovered gradients of finite element approximations on nonuniform rectangular and quadrilateral meshes," in *The Mathematics of Finite Elements and Applications VII,* J. R. Whiteman, Ed., MAFELAP 1990, Academic Press, London, p. 563.

21. M. T. Nakao, "Superconvergence of the gradient of Galerkin approximations for elliptic problems," *RAIRO Anal. Numér.* **21,** 679 (1987).

22. M. F. Wheeler and J. R. Whiteman, "Superconvergent recovery of gradients on subdomains for piecewise linear finite element approximations," *Numer. Meth. Partial Different. Eq.* **3,** 65 (1987).

23. J. R. Whiteman and G. Goodsell, "Some gradient superconvergence results in the finite element method," in *Proc. Numerical Analysis Summer School, Lancaster 1987,* Lecture Notes in Mathematics, 1397, P. R. Turner, Ed., Springer–Verlag, Berlin, 1987, p. 182.

24. R. Rannacher, "Extrapolation techniques in the finite element method (A survey)," Proc. of Nevanlinna, Helsinki Univ. of Tech. Inst. Math. Rep., 1988.

25. Q. Lin, N. Yan, and A. Zhou, "A rectangle test for interpolated finite elements," *Proc. Sys. Sci. Sys. Eng.,* 217 (1991).

26. O. C. Zienkiewicz and J. Z. Zhu, "The superconvergence patch recovery and *a posteriori* error estimates. Part 1: The recovery technique," *Internat. J. Numer. Meth. Eng.* **33,** 1331 (1992).

27. P. Lesaint and M. Zlámal, "Superconvergence of the gradient of finite element solutions," *RAIRO Anal. Numér.* **13,** 139 (1979).

28. M. Zlámal, "Superconvergence and reduced integration in the finite element method," *Math. Comp.* **32,** 663 (1978).

29. P. G. Ciarlet, *The Finite Element Method for Elliptic Problems,* North–Holland, Amsterdam, 1978.

30. L. B. Wahlbin, *Superconvergence in Galerkin Finite Element Methods,* Lecture Notes in Mathematics, **1605,** Springer, Berlin, 1995.

31. Z. Zhang and J. Z. Zhu, "Superconvergence of the derivative patch recovery technique and *a posteriori* error estimation," in the *IMA Volumes in Mathematics and its Applications,* "Modeling, mesh generation, and adaptive numerical methods for partial differential equations," I. Babuška, J. E. Flaherty, J. E. Hopcroft, W. D. Henshaw, J. E. Oliger, and T. Tezduyar, Eds., Springer, New York, 1995, p. 431.