SUPERCONVERGENCE ANALYSIS OF THE LINEAR FINITE ELEMENT METHOD AND A GRADIENT RECOVERY POSTPROCESSING ON ANISOTROPIC MESHES

WEIMING CAO

ABSTRACT. For the linear finite element method based on general unstructured anisotropic meshes in two dimensions, we establish the superconvergence in energy norm of the finite element solution to the interpolation of the exact solution for elliptic problems. We also prove the superconvergence of the postprocessing process based on the global L^2 -projection of the gradient of the finite element solution. Our basic assumptions are: (i) the mesh is quasi-uniform under a Riemannian metric and (ii) each adjacent element pair forms an approximate (anisotropic) parallelogram. The analysis follows the same methodology developed by Bank and Xu in 2003 for the case of quasi-uniform meshes, and the results can be considered as an extension of their conclusion to the adaptive anisotropic meshes. Numerical examples involving both internal and boundary layers are presented in support of the theoretical analysis.

1. INTRODUCTION

Recovery type error estimators are widely used in computations based on the finite element method, particularly in various engineering applications, due to their simplicity and robustness. They are often asymptotically exact, i.e., the ratio of the estimated error over the true error tends to 1 as the mesh gets finer and finer. These desirable properties are closely associated with the superconvergence of the finite element solution and its postprocessing. There have been extensive studies on the superconvergence and the recovery type error estimators; see, e.g., monographs by Wahlbin [28], Ainsworth and Oden [1], Babuška and Strouboulis [2], and Zhu [35], for an overview on these topics.

Since the superconvergence is generally resulted from the cancellation of certain lower order terms in the finite element approximation due to local mesh symmetry, its mathematical analysis typically begins with the FE approximations based on uniform meshes, then extends to discretizations on mildly structured meshes [30, 31], unstructured quasi-uniform meshes [3, 4, 15], and more recently on adaptively refined meshes [29]. For problems exhibiting anisotropic features, e.g., internal and boundary layers, FE solutions based on anisotropic meshes can be much more effective than those based on isotropic ones. Recovery type error estimators have also been applied to this type of computation [12, 14, 19, 25]. Numerical examples demonstrate that these estimators are still robust and reliable at the presence of

Received by the editor June 30, 2012 and, in revised form, April 29, 2013.

2010 *Mathematics Subject Classification.* Primary 65N30, 65N15, 65N50.

Key words and phrases. Superconvergence, anisotropic mesh, recovery type error estimates, post processing, linear finite element.

This work was supported in part by NSF grant DMS-0811232.

©2014 American Mathematical Society
Reverts to public domain 28 years from publication

highly anisotropic elements [20, 24]. There have been some theoretical studies on the superconvergence and recovery techniques on structured anisotropic meshes. For instance, Li and Wheeler [16] proved the superconvergence of the FE solution to singular perturbed equations on anisotropic tensor product meshes. Shi, Mao, and Chen [18, 27] published a number of papers on the superconvergence analysis on uniform meshes that stretch anisotropically along coordinate axes. Lin and Lin [17] showed the superconvergence of bilinear elements on anisotropic rectangular meshes. Zhang [32], Xu and Zhang [30] demonstrated that the superconvergence of the FE solution and the asymptotic exactness of their polynomial preserving recovery estimators hold on a class of structured anisotropic meshes. However, we have seen little theoretical work on the case of general adaptively refined unstructured anisotropic meshes.

The purpose of this paper is to provide an analysis of the superconvergence property for the linear finite element approximation of second order problems on adaptively refined anisotropic meshes and an analysis of the recovery type error estimator based on the global L^2 -projection. More specifically, we consider a class of triangulations that are quasi-uniform under a problem dependent Riemannian metric. By assuming each pair of adjacent elements in the partition forms an approximate parallelogram, we demonstrate that $\|\nabla u_N - \nabla u_I\|$ is superconvergent, where u_N and u_I are the finite element solution and the linear interpolation of the exact solution, respectively, and N is the total number of elements. Furthermore, we prove that $\|\nabla u - \mathcal{P}(\nabla u_N)\|$ is superconvergent, where $\mathcal{P}$ is the global L^2 -projection onto the space of continuous piecewise linear functions. These conclusions imply immediately the asymptotic exactness of the recovery based error estimator $\|\nabla u_N - \mathcal{P}(\nabla u_N)\|$. This work is an extension of Bank and Xu's results [3] on the recovery technique based on L^2 -projection on quasi-uniform meshes.

Needless to say, the convergence and superconvergence of the finite element solution and the error estimator based on its recovery relies intimately on the interplay between the anisotropic behaviors of the PDE solution and the anisotropic features of the meshes. Here we measure the anisotropic behavior of the PDE solution by using some non-negative matrices determined by its higher order directional derivatives introduced in [6, 7, 21, 22]. Then the error bounds can be expressed as a multiple of certain powers of N with the multiplicative constants being some integrals involving the mesh metrics and the anisotropic measures.

A sketch of the paper is as follows: We first extend the concept of approximate parallelograms to the anisotropic meshes. Then we describe the model problem, the anisotropic meshes, and the assumption about the anisotropic behavior of the PDE solutions. We prove the superconvergence of $\|u_N - u_I\|_1$ in Section 4 and the superconvergence of $\|u_N - \mathcal{P}(\nabla u_N)\|$ in Section 5. Numerical examples are presented in Section 6. We conclude the paper with some discussions in Section 7.

2. APPROXIMATE PARALLELOGRAMS ON ANISOTROPIC MESHES

Local mesh symmetry plays an essential role in the superconvergence analysis of the finite element approximations. For linear elements, this symmetry typically requires each pair of elements sharing a common edge to form a parallelogram or an approximate parallelogram. We generalize here the concept of $O(h^{1+\alpha})$ -approximate parallelogram for quasi-uniform meshes introduced in [3, 15, 26].

First, recall that a parallelogram can be characterized in several ways, e.g., the ratio of opposite side lengths is 1, or the distance between the midpoints of two diagonals is 0, or each pair of opposite sides are in parallel. For a family $\{\mathcal{T}_h\}$ of quasi-uniform triangulations, with h being the characteristic diameter of elements in $\mathcal{T}_h$, $\tau_1 \cup \tau_2$ forms an $O(h^{1+\alpha})$ approximate parallelogram, if the ratio of its opposite side lengths is of order $O(h^\alpha)$, or if the distance between the midpoints of two diagonals is $O(h^{1+\alpha})$, or the angle between the outward normals on each pair of opposite sides are within $O(h^\alpha)$ range of π ; see [3, 15, 26].

Instead of using the above geometric characters to define an approximate parallelogram, we use here the Jacobians of the affine mapping from an equilateral standard element $\hat{\tau}$ to τ_1 and τ_2 . Suppose τ_1 is with vertices i, j , and k , and τ_2 with vertices j, i , and k' . Let J_1 and J_2 be the Jacobians of the affine mappings from $\hat{\tau}$ to τ_1 and τ_2 with vertices k and k' being the images of the same vertex on $\hat{\tau}$, respectively. Then $\tau_1 \cup \tau_2$ forms a parallelogram iff $J_1 = -J_2$, and $\tau_1 \cup \tau_2$ forms an $O(h^{1+\alpha})$ approximate parallelogram, iff

$$\|I + J_1^{-1}J_2\| \simeq \|I + J_1J_2^{-1}\| = O(h^\alpha).$$

In addition, $\{\mathcal{T}_h\}$ is a regular family of partition iff $\text{cond}(J_\tau) = O(1)$ for all $\tau \in \mathcal{T}_h$.

For discretization involving anisotropic meshes, the element diameter is not a proper parameter for describing the asymptotic behaviors, since an element can have much different length scales in different directions. Instead, the total number N of elements is essential. Thus we introduce the following concept of the $O(N^{-(1+\alpha)/2})$ -approximate parallelogram.

Definition 2.1. Let τ_1 and τ_2 be a pair of elements sharing a common edge, and let J_1 and J_2 be the Jacobians of the affine mappings from equilateral standard element $\hat{\tau}$ to τ_1 and τ_2 . $\tau_1 \cup \tau_2$ is called forming an $O(N^{-(1+\alpha)/2})$ -approximate parallelogram, if

$$(1) \quad \|I + J_1^{-1}J_2\| = O(N^{-\alpha/2}).$$

When $\{\mathcal{T}_N\}$ is quasi-uniform, we have $h = \text{diam}(\tau) = O(N^{-1/2})$, and the above definition is reduced to the $O(h^{1+\alpha})$ -approximate parallelogram for shape regular elements described in [3, 15].

It should be noted that this definition is independent of the ordering of τ_1 and τ_2 . It is easy to see that condition (1) is equivalent to

$$(2) \quad \|I + J_2^{-1}J_1\| = O(N^{-\alpha/2}).$$

Indeed, we have from (1) that

$$\|J_2^{-1}J_1\| = \|(I - (I + J_1^{-1}J_2))^{-1}\| \leq \frac{1}{1 - \|I + J_1^{-1}J_2\|} = O(1)$$

Hence

$$\|I + J_2^{-1}J_1\| = \|J_2^{-1}J_1(I + J_1^{-1}J_2)\| \leq \|J_2^{-1}J_1\| \cdot \|I + J_1^{-1}J_2\| = O(N^{-\alpha/2}).$$

However, this definition is not equivalent to requiring that

$$(3) \quad \|I + J_1J_2^{-1}\| = O(N^{-\alpha/2}).$$

We give here two examples of the anisotropic element pairs, one forms an approximate parallelogram by satisfying (1), and another does not, even though it satisfies (3). We choose condition (1) over (3) since the former measures the deviation

of $\tau_1 \cup \tau_2$ from an exact parallelogram in the coordinate system for the standard element, which is fixed for all elements and all mesh sizes.

FIGURE 1. Left: a pair of elements that forms an approximate parallelogram; Right: a pair of elements that does not form an approximate parallelogram.

Example 1. Let $\epsilon = N^{-\alpha/2} \rightarrow 0$, and let the standard element be the equilateral triangle with vertices $(\cos \theta_m, \sin \theta_m)$, $\theta_m = \frac{\pi}{2} + \frac{2m}{3}\pi$, $m = 0, 1, 2$. Choose $\mathbf{x}_1 = (0, 0)$, $\mathbf{x}_2 = (1 - \epsilon, 0)$, $\mathbf{x}_3 = (1, \epsilon)$, and $\mathbf{x}_4 = (0, \epsilon)$. Elements $\tau_1 = \Delta \mathbf{x}_4 \mathbf{x}_1 \mathbf{x}_3$, and $\tau_2 = \Delta \mathbf{x}_2 \mathbf{x}_3 \mathbf{x}_1$; see Figure 1(a). The center of τ_1 is $\mathbf{x}_c = (\frac{1}{3}, \frac{2\epsilon}{3})$. Then the affine mapping from $\hat{\tau}$ to τ_1 (with $(0, 1)$ to $\mathbf{x}_4$) is

$$\mathbf{x} = J_1 \boldsymbol{\xi} + \mathbf{x}_c$$

with the Jacobian J_1 for the mapping being

$$J_1 = \begin{pmatrix} \frac{1}{\sqrt{3}}(\mathbf{x}_3 - \mathbf{x}_1), \mathbf{x}_4 - \mathbf{x}_c \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{3} \\ \frac{\epsilon}{\sqrt{3}} & \frac{\epsilon}{3} \end{pmatrix}.$$

Similarly, the Jacobian for the affine mapping from $\hat{\tau}$ to τ_2 (with $(0, 1)$ to $\mathbf{x}_2$) is

$$J_2 = \begin{pmatrix} -\frac{1}{\sqrt{3}} & \frac{1-2\epsilon}{3} \\ -\frac{\epsilon}{\sqrt{3}} & -\frac{\epsilon}{3} \end{pmatrix} = -J_1 + \begin{pmatrix} 0 & -\frac{2\epsilon}{3} \\ 0 & 0 \end{pmatrix}.$$

It is elementary to verify that

$$I + J_1^{-1} J_2 = J_1^{-1} \cdot \begin{pmatrix} 0 & -\frac{2\epsilon}{3} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\frac{\epsilon}{\sqrt{3}} \\ 0 & \epsilon \end{pmatrix},$$

and that

$$I + J_2 J_1^{-1} = \begin{pmatrix} 0 & -\frac{2\epsilon}{3} \\ 0 & 0 \end{pmatrix} \cdot J_1^{-1} = \begin{pmatrix} \epsilon & -1 \\ 0 & 0 \end{pmatrix}.$$

Thus (1) holds for $\tau_1 \cup \tau_2$, and it forms an $O(N^{-(1+\alpha)/2})$ -approximate parallelogram, even though the left and right sides always form a 45° angle, and their lengths differ by a constant multiple $\sqrt{2}$. Note that (3) is not satisfied in this case.

Example 2. Let $\mathbf{x}_1 = (-1, 0)$, $\mathbf{x}_2 = (0, -\epsilon)$, $\mathbf{x}_3 = (1, 0)$, and $\mathbf{x}_4 = (0, 2\epsilon)$. Elements $\tau_1 = \Delta \mathbf{x}_1 \mathbf{x}_2 \mathbf{x}_4$, and $\tau_2 = \Delta \mathbf{x}_3 \mathbf{x}_4 \mathbf{x}_2$; see Figure 1(b). The Jacobians for the mapping from $\hat{\tau}$ to τ_1 and τ_2 with $(1, 0)$ to $\mathbf{x}_1$ and $\mathbf{x}_3$, respectively, are

$$J_1 = \begin{pmatrix} 0 & -\frac{2}{3} \\ \sqrt{3}\epsilon & -\frac{\epsilon}{3} \end{pmatrix}, \quad J_2 = \begin{pmatrix} 0 & \frac{2}{3} \\ -\sqrt{3}\epsilon & -\frac{\epsilon}{3} \end{pmatrix} = -J_1 + \begin{pmatrix} 0 & 0 \\ 0 & -\frac{2\epsilon}{3} \end{pmatrix}.$$

Clearly,

$$I + J_1^{-1} J_2 = \begin{pmatrix} 0 & -\frac{2}{3\sqrt{3}} \\ 0 & 0 \end{pmatrix},$$

and

$$I + J_2 J_1^{-1} = \begin{pmatrix} 0 & 0 \\ \epsilon & 0 \end{pmatrix}.$$

Thus (1) does not hold for $\tau_1 \cup \tau_2$, and it does not form an $O(N^{-(1+\alpha)/2})$ -approximate parallelogram. Nevertheless, (3) is satisfied.

To further elaborate the geometric meaning of the approximate parallelograms, we list in the following several lemmas.

Lemma 2.1. *If $\tau_1 \cup \tau_2$ forms an $O(N^{-(1+\alpha)/2})$ -approximate parallelogram, then*

$$\frac{|\tau_2|}{|\tau_1|} = 1 + O(N^{-\alpha/2}).$$

This conclusion comes from the fact that (1) implies

$$\frac{|\tau_2|}{|\tau_1|} = \det(J_1^{-1} J_2) = \det(-I + (I + J_1^{-1} J_2)) = 1 + O(N^{-\alpha/2}).$$

Lemma 2.2. *A pair of adjacent elements τ_1 and τ_2 forms an $O(N^{-(1+\alpha)/2})$ -approximate parallelogram, if and only if there exists an affine mapping F such that $F(\tau_1) \cup F(\tau_2)$ forms a shape regular $O(N^{-(1+\alpha)/2})$ -approximate parallelogram.*

Proof. Note that the Jacobians for the affine mapping from $\hat{\tau}$ to $F(\tau_1)$ and $F(\tau_2)$ are $\tilde{J}_1 = FJ_1$ and $\tilde{J}_2 = FJ_2$, resp. If $F(\tau_1) \cup F(\tau_2)$ forms a shape regular $O(N^{-(1+\alpha)/2})$ -approximate parallelogram, then

$$\|I + \tilde{J}_1^{-1} \tilde{J}_2\| = O(N^{-\alpha/2}),$$

which implies (1), and thus $\tau_1 \cup \tau_2$ forms an $O(N^{-(1+\alpha)/2})$ -approximate parallelogram.

On the other hand, if $\tau_1 \cup \tau_2$ forms an $O(N^{-(1+\alpha)/2})$ -approximate parallelogram, choose $F = J_1^{-1}$. Then $F(\tau_1)$ is equilateral. $F(\tau_2)$ is shape regular since (1) implies $\text{cond}(FJ_2) \leq C$. Furthermore,

$$\|I + \tilde{J}_1^{-1} \tilde{J}_2\| = \|I + J_1^{-1} J_2\| = O(N^{-\alpha/2}).$$

Thus $F(\tau_1) \cup F(\tau_2)$ forms a shape regular $O(N^{-(1+\alpha)/2})$ -approximate parallelogram. $\square$

The above lemma indicates that an anisotropic $O(N^{-(1+\alpha)/2})$ -approximate parallelogram has to be the affine image of a shape regular $O(N^{-(1+\alpha)/2})$ -approximate parallelogram. Note that for a pair of shape regular elements, if they form an $O(h^{1+\alpha})$ -approximate parallelogram, then another element pair obtained by swapping the diagonals of this approximate parallelogram also forms an $O(h^{1+\alpha})$ -approximate parallelogram. Namely, edge swapping in a mesh does not affect whether two adjacent elements form an approximate parallelogram or not. This prompts us to introduce yet another equivalent condition for approximate parallelogram by using only the location of four vertices of the adjacent elements.

Lemma 2.3. *Let $Q = \tau_1 \cup \tau_2$, and let $\hat{Q}$ be the standard quadrilateral element, e.g., $[0, 1]^2$. J_Q is the Jacobian of the bilinear mapping from $\hat{Q}$ to Q , and $\mathbf{d}$ is the vector connecting two midpoints of the diagonals of Q . Then $\tau_1 \cup \tau_2$ forms an $O(N^{-(1+\alpha)/2})$ -approximate parallelogram, iff*

$$(4) \quad |J_Q^{-1}(\boldsymbol{\xi})\mathbf{d}| = O(N^{-\alpha/2}), \quad \forall \boldsymbol{\xi} \in \hat{Q}.$$

Proof. First we note that both conditions (1) and (4) are invariant under any affine transform of $\tau_1 \cup \tau_2$. Thus we only have to show the above lemma in the cases τ_1 and τ_2 are shape regular. In particular, we may assume, without loss of generality, that Q is with vertices $\mathbf{x}_1 = (0, 0)$, $\mathbf{x}_2 = (1, 0)$, $\mathbf{x}_3 = (1 + \delta_1, 1 + \delta_2)$, and $\mathbf{x}_4 = (0, 1)$. It is ready to verify that $\mathbf{d} = \frac{1}{2}(\mathbf{x}_1 + \mathbf{x}_3 - \mathbf{x}_2 - \mathbf{x}_4) = \frac{1}{2}[\delta_1, \delta_2]^T$ and that the bilinear mapping $\mathbf{x} = \mathbf{B}(\boldsymbol{\xi})$ from $\hat{Q}$ to Q is

$$(5) \quad \mathbf{B}(\boldsymbol{\xi}) = \mathbf{x}_1 + \xi_1(\mathbf{x}_2 - \mathbf{x}_1) + \xi_2(\mathbf{x}_4 - \mathbf{x}_1) + 2\xi_1\xi_2\mathbf{d}$$

and

$$J_Q(\boldsymbol{\xi}) = \frac{\partial \mathbf{x}}{\partial \boldsymbol{\xi}} = \begin{bmatrix} 1 + \delta_1\xi_2 & \delta_1\xi_1 \\ \delta_2\xi_2 & 1 + \delta_2\xi_1 \end{bmatrix}.$$

An elementary calculation leads to

$$J_Q^{-1}(\boldsymbol{\xi})\mathbf{d} = \frac{1}{2(1 + \delta_1\xi_2 + \delta_2\xi_1)} \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix}.$$

Since for all $\boldsymbol{\xi} \in [0, 1]^2$, we have

$$1 - \sqrt{2}|\mathbf{d}| \leq 1 + \delta_1\xi_2 + \delta_2\xi_1 \leq 1 + \sqrt{2}|\mathbf{d}|.$$

Thus (4) is satisfied iff $|\mathbf{d}| = O(N^{-\alpha/2})$, i.e., $\tau_1 \cup \tau_2$ forms an $O(N^{-(1+\alpha)/2})$ -approximate parallelogram. $\square$

We point out that to study the superconvergence of the finite element approximation on quadrilateral meshes, Zhang [33] showed that (4) is a necessary condition for a shape regular approximate parallelogram. The above lemma indicates that (4) is also a sufficient condition and can be used even in the case of anisotropic meshes.

3. MODEL PROBLEMS AND THEIR FE APPROXIMATION

We consider the following homogeneous Dirichlet problem of a second order elliptic equation:

$$(6) \quad \begin{cases} -\nabla \cdot (A \nabla u + \mathbf{b} u) + d u = f, & \text{in } \Omega, \\ u|_{\partial\Omega} = 0. \end{cases}$$

We assume in this paper A is a positive definite constant matrix, $\mathbf{b}, d$ are suitably smooth functions and equation (6) is strongly elliptic.

FE approximation: Let $\{\mathcal{T}_N\}$ be a family of regular partitions of Ω [9]. Here, N represents the total number of elements in $\mathcal{T}_N$. Define S_N be the space of continuous piecewise linear functions based on partition $\mathcal{T}_N$. $V_N = S_N \cap H_0^1(\Omega)$. Then the finite element method for solving (6) is to find the approximate solution $u_N \in V_N$ such that

$$(7) \quad a(u_N, v_N) = (f, v_N), \quad \forall v_N \in V_N,$$

where bilinear form $a(\cdot, \cdot)$ is defined as

$$a(u, v) = (A \nabla u, \nabla v) + (u, \mathbf{b} \cdot \nabla v) + (du, v).$$

Anisotropic meshes: Since the convergence and superconvergence of the finite element solution depend closely on the properties of meshes, we restrict our analysis on a class of anisotropic meshes that are quasi-uniform under a given metric. More

precisely, Let M be a continuous Riemannian metric on $\bar{\Omega}$. For each element $\tau \in \mathcal{T}_N$, let M_τ be the average of M over τ , and its eigen-decomposition is of the form

$$(8) \quad M_\tau = T_\tau \cdot \Lambda_\tau \cdot T_\tau',$$

where Λ_τ is diagonal and T_τ is orthonormal. Define

$$(9) \quad F_\tau = T_\tau \Lambda_\tau^{-\frac{1}{2}}.$$

A family of triangulations $\{\mathcal{T}_N\}$ is called quasi-uniform under metric M , if for all $\tau \in \mathcal{T}_N$, $\hat{\tau} = F_\tau^{-1}\tau$ are shape regular and of about the same size; see [6–8]. Let J_τ be the Jacobian of the affine mapping from standard element $\hat{\tau}$ to element τ . Then $\{\mathcal{T}_N\}$ is quasi-uniform under metric M if and only if

$$(10) \quad \|F_\tau^{-1} J_\tau\| \simeq (\|J_\tau^{-1} F_\tau\|)^{-1} \simeq (C_M/N)^{1/2}, \quad \forall \tau \in \mathcal{T}_N,$$

where

$$(11) \quad C_M = \int_{\Omega} |\det(M)|^{1/2}.$$

Clearly, the error for the finite element approximation depends on the interplay between the anisotropic features of the meshes and the anisotropic behavior of the solutions. In order to derive the convergence and superconvergence for the linear finite element approximation, we need the following assumption to characterize the anisotropic behavior of the second and third derivatives of solution u .

Assumption on D^2u and D^3u : Let $Q_2(\mathbf{x})$ and $Q_3(\mathbf{x})$ be 2×2 symmetric non-negative definite matrices. For $m = 2, 3$, we assume at each $\mathbf{x} \in \Omega$ that

$$(12) \quad |(\mathbf{s} \cdot \nabla)^m u(\mathbf{x})| \leq (\mathbf{s} \cdot Q_m(\mathbf{x}) \mathbf{s})^{m/2}, \quad \forall \mathbf{s} \in \mathcal{R}^2.$$

Such matrices always exist as long as D^2u and D^3u exist. For instance, $Q_2(\mathbf{x})$ can be chosen as $|\nabla^2 u| I$, where I is the identity matrix and $|\nabla^2 u| = \max_{\|\mathbf{s}\|=1} |(\mathbf{s} \cdot \nabla)^2 u|$. It can also be chosen as $\text{abs}(\nabla^2 u(\mathbf{x}))$, which is the matrix of the same eigenvectors as $\nabla^2 u(\mathbf{x})$ but with the eigenvalues equal to the absolute values of $\nabla^2 u(\mathbf{x})$'s. Note that the left-hand side of (12) is the directional derivative along $\mathbf{s}$, the “smaller” the matrix Q_m , the more precise it describes the anisotropic behaviors of $D^m u$. A geometric explanation of (12) and a numerical procedure to find these matrices can be found in [6, 7]. More recently, Mirebeau studied systematically how to find these matrices and applied them in anisotropic mesh generation. He also derived an explicit formula to define Q_3 ; see [21, 22] for details.

The effects of the anisotropic mesh for approximating a solution of anisotropic derivatives can be appreciated by the following inequalities. Let $\hat{u}$ be the pull-back of $u|_\tau$ under the affine mapping from the standard element $\hat{\tau}$ to element τ . Then we have $\hat{\nabla} = J_\tau^T \nabla$, and

$$(13) \quad |(\mathbf{s} \cdot \hat{\nabla})^m \hat{u}| \leq (\mathbf{s} \cdot J_\tau^T Q_m J_\tau \mathbf{s})^{m/2}, \quad \forall \mathbf{s} \in \mathcal{R}^2$$

which implies

$$(14) \quad |\hat{D}^m \hat{u}| \leq \|J_\tau^T Q_m J_\tau\|^{m/2}.$$

When $D^m u$ is highly anisotropic, the eigenvalues of Q_m are of very different magnitudes. A proper choice of J_τ may make $|\hat{D}^m \hat{u}|$ significantly smaller than $|D^m u|$.

a-priori error estimate: By the ellipticity of equation (6), it follows from the interpolation error estimates (see Theorem 2.1 in [7]) the following *a priori* error estimate for the linear finite element approximation of (6).

Theorem 3.1. *Suppose $\{\mathcal{T}_N\}$ is quasi-uniform under metric M , and u_N is the linear finite element approximation of the solution u of (6). Furthermore, assume the second derivative D^2u satisfies assumption (12) about its anisotropic behavior. Then there exists a constant c independent of u and N such that*

$$(15) \quad \|u - u_N\|_{1,\Omega} \leq cN^{-1/2} \cdot \left\{ C_M \int_{\Omega} \|F^{-1}\|^2 \|F^T Q_2 F\|^2 \right\}^{1/2},$$

where F and C_M are determined by M as in (9) and (11).

Remark 3.1. When $\nabla^2 u$ is positive definite on Ω , we may choose $Q_2 = \nabla^2 u$ in (12), and choose $M = c(\lambda_{max}/\lambda_{min})^{1/4} \cdot \nabla^2 u$, where λ_{max} and λ_{min} are the largest and smallest eigenvalues of $\nabla^2 u$, respectively; see [5, 7]. In this case, (15) is reduced to

$$\|u - u_N\|_{1,\Omega} \leq cN^{-1/2} \|\lambda_{min}^{1/4} \lambda_{max}^{3/4}\|_{L^1(\Omega)}.$$

In contrast, if we allow only shape regular elements, then the best isotropic mesh metric to minimize the error bound in (15) would be $M = \lambda_{max} I$, and the error bound for $\|u - u_N\|_{1,\Omega}$ would be $cN^{-1/2} \|\lambda_{max}\|_{L^1(\Omega)}$, which is clearly much larger than for the anisotropic case when λ_{max} and λ_{min} are of much different magnitudes.

4. SUPERCONVERGENCE OF FE SOLUTION

First we recall a basic lemma developed by Bank and Xu [3] for analyzing the “weak convergence” of linear interpolation on a triangular element. Let τ be an element with vertices $\mathbf{x}_k$, $k = 1, 2, 3$. Denote by θ_k the internal angle of τ at $\mathbf{x}_k$. Let $\ell_k = \mathbf{x}_{k-1} - \mathbf{x}_{k+1}$ be the edge (directed counterclockwise) opposite to $\mathbf{x}_k$, and $\ell_k = |\ell_k|$ is its length. $\mathbf{t}_k = \ell_k/\ell_k$ and $\mathbf{n}_k$ are the unit tangential and outward normal directions on ℓ_k , respectively, and

$$\xi_k = -\mathbf{n}_{k+1} \cdot A\mathbf{n}_{k-1}.$$

Furthermore, let ϕ_k be the nodal basis function associated with $\mathbf{x}_k$, $q_k = \phi_{k+1}\phi_{k-1}$, and $\psi_k = \phi_k(1 - \phi_k)$.

Lemma 4.1. *Let τ be an element in $\mathcal{T}_N$, and let u_I be the linear interpolation of u at the vertices of τ . Then*

$$(16) \quad \begin{aligned} \int_{\tau} \nabla(u - u_I) \cdot A \nabla v_N &= \sum_{k=1}^3 \int_{e_k} \frac{\xi_k q_k}{2 \sin \theta_k} \cdot \left\{ (\ell_{k+1}^2 - \ell_{k-1}^2) \frac{\partial^2 u}{\partial \mathbf{t}_k^2} + 4|\tau| \frac{\partial^2 u}{\partial \mathbf{t}_k \partial \mathbf{n}_k} \right\} \\ &\cdot \frac{\partial v_N}{\partial \mathbf{t}_k} - \int_{\tau} \sum_{k=1}^3 \frac{\xi_k \ell_k}{2 \sin^2 \theta_k} \left\{ \ell_{k+1} \psi_{k-1} \frac{\partial^3 u}{\partial^2 \mathbf{t}_{k+1} \partial \mathbf{t}_{k-1}} + \ell_{k-1} \psi_{k+1} \frac{\partial^3 u}{\partial^2 \mathbf{t}_{k-1} \partial \mathbf{t}_{k+1}} \right\}. \end{aligned}$$

This lemma plays a fundamental role in the superconvergence analysis [3, 30]. It separates not only the error contribution from each edge of the elements, but more importantly, the contributions from normal and tangential derivatives. It makes the cancellation of lower order error terms much easier. Moreover, (16) is an identity valid on any triangle. We shall use it on the standard element $\hat{\tau}$.

Theorem 4.1. *Let u_I and u_N be the linear interpolation and the finite element approximation of u , respectively. Suppose D^2u and D^3u satisfy assumption (12),*

and each pair of adjacent elements in $\mathcal{T}_N$ forms an $O(N^{-(1+\alpha)/2})$ -approximate parallelogram. Then

$$(17) \quad \|u_N - u_I\|_{1,\Omega} \leq c \left\{ \sum_{\tau} \int_{\tau} (1 + N^{-\alpha} \|J_{\tau}^{-1}\|^4 \|J_{\tau}\|^2) \cdot \|J_{\tau}^T Q_2 J_{\tau}\|^2 + \|J_{\tau}^{-1}\|^4 \|J_{\tau}\|^2 \cdot \|J_{\tau}^T Q_3 J_{\tau}\|^3 \right\}^{1/2}.$$

In addition, if the partition $\mathcal{T}_N$ is quasi-uniform under metric M , then we have

$$(18) \quad \|u_N - u_I\|_{1,\Omega} \leq c N^{-1/2} \left\{ \int_{\Omega} ((C_M)^2 N^{-1} + C_M N^{-\alpha} \|F^{-1}\|^4 \|F\|^2) \cdot \|F^T Q_2 F\|^2 + (C_M)^2 N^{-1} \|F^{-1}\|^4 \|F\|^2 \cdot \|F^T Q_3 F\|^3 \right\}^{1/2},$$

where F and C_M are determined by M as in (9) and (11).

Proof. By the ellipticity of $a(\cdot, \cdot)$ and the orthogonality of the finite element approximation,

$$\|u_N - u_I\|_{1,\Omega} \leq c \sup_{v_N} \frac{|a(u_N - u_I, v_N)|}{\|v_N\|_{1,\Omega}} = c \sup_{v_N} \frac{|a(u - u_I, v_N)|}{\|v_N\|_{1,\Omega}}.$$

To prove the theorem, we only have to bound $a(u - u_I, v_N)$ by a certain multiple of $\|v_N\|_1$ for arbitrary $v_N \in S_N$. Note that

$$(19) \quad a(u - u_I, v_N) = \sum_{\tau} \int_{\tau} \nabla(u - u_I) \cdot A \nabla v_N + \int_{\Omega} (u - u_I) (\mathbf{b} \cdot \nabla v_N + dv_N).$$

First we estimate the second term on the right-hand side of the above equation. On each element $\tau \in \mathcal{T}_N$, let $\mathcal{F}_{\tau}$ be the affine mapping from the equilateral standard element $\hat{\tau}$ to τ , $\hat{u}_{\tau} = u \circ \mathcal{F}_{\tau}$, and let $\hat{u}_I = \hat{\Pi}_1 \hat{u}_{\tau}$ be the linear interpolation of $\hat{u}_{\tau}$ on $\hat{\tau}$. It follows from assumption (12) on the anisotropic behavior of $D^2 u$ that

$$\begin{aligned} \left| \int_{\Omega} (u - u_I) \cdot (\mathbf{b} \cdot \nabla v_N + dv_N) \right| &\leq c \|u - u_I\| \|v_N\|_1 \\ &\leq c \|v_N\|_1 \left\{ \sum_{\tau} \int_{\tau} \|u - u_I\|^2 \right\}^{1/2} \\ &\leq c \|v_N\|_1 \left\{ \sum_{\tau} |\tau| \cdot \int_{\hat{\tau}} \|\hat{u} - \hat{u}_I\|^2 \right\}^{1/2} \\ &\leq c \|v_N\|_1 \left\{ \sum_{\tau} |\tau| \cdot \int_{\hat{\tau}} |\hat{D}^2 \hat{u}|^2 \right\}^{1/2} \\ &\leq c \|v_N\|_1 \left\{ \sum_{\tau} \int_{\tau} \|J_{\tau}^T Q_2 J_{\tau}\|^2 \right\}^{1/2}. \end{aligned}$$

Now we deal with the first term on the right-hand side of (19). On each element τ , let J_{τ} be the Jacobian of the mapping $\mathcal{F}_{\tau}$. Let $\hat{A}_{\tau} = J_{\tau}^{-1} A J_{\tau}^{-T}$ and $\hat{\xi}_k = -\hat{\mathbf{n}}_{k+1} \cdot A \hat{\mathbf{n}}_{k-1}$. Then by a change of variables, we have $\hat{\nabla} = J_{\tau}^T \nabla$ and

$$(20) \quad \sum_{\tau \in \mathcal{T}_N} \int_{\tau} \nabla(u - u_I) \cdot A \nabla v_N = \sum_{\tau \in \mathcal{T}_N} |\tau| \cdot \int_{\hat{\tau}} \hat{\nabla}(\hat{u}_{\tau} - \hat{u}_I) \cdot \hat{A}_{\tau} \hat{\nabla} \hat{v}_N.$$

Applying Lemma 4.1 to the integrals on $\hat{\tau}$ for each τ , and noting that $|\hat{\tau}| = \frac{3\sqrt{3}}{4}$, $|\hat{\ell}_k| = \sqrt{3}$ and $\theta_k = \frac{\pi}{3}$ for all $k = 1, 2, 3$, we have

$$(21) \quad \int_{\hat{\tau}} \hat{\nabla}(\hat{u}_\tau - \hat{u}_I) \cdot \hat{A}_\tau \hat{\nabla} \hat{v}_N = \sum_{k=1}^3 \int_{\hat{e}_k} \hat{\xi}_k q_k \cdot \frac{\partial^2 \hat{u}_\tau}{\partial \hat{\mathbf{t}}_k \partial \hat{\mathbf{n}}_k} \cdot \frac{\partial \hat{v}_N}{\partial \hat{\mathbf{t}}_k} \\ - 2 \sum_{k=1}^3 \int_{\hat{\tau}} \hat{\xi}_k (\psi_{k+1} \frac{\partial^3 \hat{u}_\tau}{\partial^2 \hat{\mathbf{t}}_{k+1} \partial \hat{\mathbf{t}}_{k-1}} + \psi_{k-1} \frac{\partial^3 \hat{u}_\tau}{\partial^2 \hat{\mathbf{t}}_{k-1} \partial \hat{\mathbf{t}}_{k+1}}) \cdot \frac{\partial \hat{v}_N}{\partial \hat{\mathbf{t}}_k}.$$

Summing up the above integrals for all $\tau \in \mathcal{T}_N$ and grouping together the two line integrals associated with a common edge $e = \tau \cap \tau'$, we have

$$\sum_{\tau \in \mathcal{T}_N} \int_{\tau} \nabla(u - u_I) \cdot A \nabla v_N = I_1 + I_2,$$

where

$$I_1 = \sum_{\tau \in \mathcal{T}_N} |\tau| \cdot \sum_{k=1}^3 \int_{\hat{e}_k} \hat{\xi}_k q_k \cdot \frac{\partial^2 \hat{u}_\tau}{\partial \hat{\mathbf{t}}_k \partial \hat{\mathbf{n}}_k} \cdot \frac{\partial \hat{v}_N}{\partial \hat{\mathbf{t}}_k} \\ = \sum_{e=\tau \cap \tau'} \left(\int_{\hat{e}} |\tau| q \hat{\xi}_\tau \frac{\partial^2 \hat{u}_\tau}{\partial \hat{\mathbf{t}} \partial \hat{\mathbf{n}}} \cdot \frac{\partial \hat{v}_N|_\tau}{\partial \hat{\mathbf{t}}} + \int_{\hat{e}'} |\tau'| q \hat{\xi}_{\tau'} \frac{\partial^2 \hat{u}_{\tau'}}{\partial \hat{\mathbf{t}} \partial \hat{\mathbf{n}}} \cdot \frac{\partial \hat{v}_N|_{\tau'}}{\partial \hat{\mathbf{t}}} \right), \\ I_2 = -2 \sum_{\tau \in \mathcal{T}_N} \left(|\tau| \sum_{k=1}^3 \int_{\hat{\tau}} \hat{\xi}_k \left(\psi_{k+1} \frac{\partial^3 \hat{u}_\tau}{\partial^2 \hat{\mathbf{t}}_{k+1} \partial \hat{\mathbf{t}}_{k-1}} + \psi_{k-1} \frac{\partial^3 \hat{u}_\tau}{\partial^2 \hat{\mathbf{t}}_{k-1} \partial \hat{\mathbf{t}}_{k+1}} \right) \cdot \frac{\partial \hat{v}_N}{\partial \hat{\mathbf{t}}_k} \right).$$

We first bound the integrals in I_2 . Note that $|\psi_k| \leq \frac{1}{4}$, $|q_k| \leq 1$,

$$|\hat{\xi}_k| \leq \|\hat{A}_\tau\| \leq \|J_\tau^{-1}\|^2 \|A\|$$

and

$$\left| \frac{\partial \hat{v}_N}{\partial \hat{\mathbf{t}}_k} \right| \leq |\hat{\nabla} \hat{v}_N| \leq \|J_\tau\| |\nabla v_N|.$$

Furthermore, by assumption (12), we have

$$\left| \frac{\partial^3 \hat{u}_\tau}{\partial^2 \hat{\mathbf{t}}_{k+1} \partial \hat{\mathbf{t}}_{k-1}} \right| \leq \|\hat{D}^3 \hat{u}_\tau\| \leq \|J_\tau^T Q_3 J_\tau\|^{3/2}.$$

Thus

$$|I_2| \leq c \sum_{\tau \in \mathcal{T}_N} \int_{\tau} \|J_\tau^{-1}\|^2 \|J_\tau\| \cdot \|J_\tau^T Q_3 J_\tau\|^{3/2} |\nabla v_N| \\ \leq c \left(\sum_{\tau \in \mathcal{T}_N} \int_{\tau} \|J_\tau^{-1}\|^4 \|J_\tau\|^2 \|J_\tau^T Q_3 J_\tau\|^3 \right)^{1/2} \cdot \|\nabla v_N\|_{L^2(\Omega)}.$$

Next, we deal with I_1 . Note that for each term in the sum up to I_1 we may assume $\hat{e} = \hat{e}'$, i.e., the common edge e of τ and τ' is the images of the same edge in the equilateral standard element under $\mathcal{F}_\tau$ and $\mathcal{F}_{\tau'}$. Otherwise, a rotation of the standard element by angle $\pm 120^\circ$ for one of the integrals will convert it to such a case. With this in mind, we have

$$\frac{\partial \hat{v}_N|_\tau}{\partial \hat{\mathbf{t}}} = - \frac{\partial \hat{v}_N|_{\tau'}}{\partial \hat{\mathbf{t}}}$$

and

$$I_1 = \sum_{e=\tau \cap \tau'} \int_{\hat{e}} \left(|\tau| \hat{\xi}_\tau \frac{\partial^2 \hat{u}_\tau}{\partial \hat{\mathbf{t}} \partial \hat{\mathbf{n}}} - |\tau'| \hat{\xi}_{\tau'} \frac{\partial^2 \hat{u}_{\tau'}}{\partial \hat{\mathbf{t}} \partial \hat{\mathbf{n}}} \right) \cdot \mathbf{q} \cdot \frac{\partial \hat{v}_N|_\tau}{\partial \hat{\mathbf{t}}}.$$

We split the integrand into

$$(22) \quad \begin{aligned} |\tau| \hat{\xi}_\tau \frac{\partial^2 \hat{u}_\tau}{\partial \hat{\mathbf{t}} \partial \hat{\mathbf{n}}} - |\tau'| \hat{\xi}_{\tau'} \frac{\partial^2 \hat{u}_{\tau'}}{\partial \hat{\mathbf{t}} \partial \hat{\mathbf{n}}} &= |\tau| (\hat{\xi}_\tau - \hat{\xi}_{\tau'}) \frac{\partial^2 \hat{u}_\tau}{\partial \hat{\mathbf{t}} \partial \hat{\mathbf{n}}} \\ &\quad + |\tau| \hat{\xi}_\tau \left(\frac{\partial^2 \hat{u}_\tau}{\partial \hat{\mathbf{t}} \partial \hat{\mathbf{n}}} - \frac{\partial^2 \hat{u}_{\tau'}}{\partial \hat{\mathbf{t}} \partial \hat{\mathbf{n}}} \right) + (|\tau| - |\tau'|) \hat{\xi}_{\tau'} \frac{\partial^2 \hat{u}_{\tau'}}{\partial \hat{\mathbf{t}} \partial \hat{\mathbf{n}}}. \end{aligned}$$

It follows from the assumption that $\tau \cup \tau'$ forms an $O(N^{-(1+\alpha)/2})$ -approximate parallelogram such that

$$\begin{aligned} |\hat{\xi}_\tau - \hat{\xi}_{\tau'}| &= |\hat{\mathbf{n}}_{k+1} \cdot (\hat{A}_\tau - \hat{A}_{\tau'}) \hat{\mathbf{n}}_{k-1}| \\ &\leq \|\hat{A}_\tau - \hat{A}_{\tau'}\| \\ &= \|J_\tau^{-1} A J_\tau^{-T} - J_{\tau'}^{-1} A J_{\tau'}^{-T}\| \\ &\leq \|(J_\tau^{-1} + J_{\tau'}^{-1}) A J_\tau^{-T}\| + \|J_{\tau'}^{-1} A (J_\tau^{-T} + J_{\tau'}^{-T})\| \\ &\leq \|J_\tau^{-1} + J_{\tau'}^{-1}\| \cdot \|A\| \cdot (\|J_\tau^{-1}\| + \|J_{\tau'}^{-1}\|) \\ &\leq \|(I + J_{\tau'}^{-1} J_\tau) J_\tau^{-1}\| \cdot \|A\| \cdot \|J_\tau^{-1}\| (1 + \|J_{\tau'}^{-1} J_\tau\|) \\ &\leq C N^{-\alpha/2} \|J_\tau^{-1}\|^2. \end{aligned}$$

On the other hand,

$$\frac{\partial^2 \hat{u}_\tau}{\partial \hat{\mathbf{t}} \partial \hat{\mathbf{n}}} = (\hat{\mathbf{t}} \cdot \hat{\nabla})(\hat{\mathbf{n}} \cdot \hat{\nabla}) \hat{u}_\tau = (J_\tau \hat{\mathbf{t}} \cdot \nabla)(J_\tau \hat{\mathbf{n}} \cdot \nabla) u$$

and

$$\frac{\partial^2 \hat{u}_{\tau'}}{\partial \hat{\mathbf{t}} \partial \hat{\mathbf{n}}} = (\hat{\mathbf{t}} \cdot \hat{\nabla})(\hat{\mathbf{n}} \cdot \hat{\nabla}) \hat{u}_{\tau'} = (J_{\tau'} \hat{\mathbf{t}} \cdot \nabla)(J_{\tau'} \hat{\mathbf{n}} \cdot \nabla) u.$$

By the fact that $J_\tau \hat{\mathbf{t}} = -J_{\tau'} \hat{\mathbf{t}}$, we have

$$\begin{aligned} \left| \frac{\partial^2 \hat{u}_\tau}{\partial \hat{\mathbf{t}} \partial \hat{\mathbf{n}}} - \frac{\partial^2 \hat{u}_{\tau'}}{\partial \hat{\mathbf{t}} \partial \hat{\mathbf{n}}} \right| &= |(J_\tau \hat{\mathbf{t}} \cdot \nabla)((J_\tau + J_{\tau'}) \hat{\mathbf{n}} \cdot \nabla) u| \\ &= |(\hat{\mathbf{t}} \cdot \hat{\nabla})(J_\tau^{-1} (J_\tau + J_{\tau'}) \hat{\mathbf{n}} \cdot \hat{\nabla}) \hat{u}_\tau| \\ &\leq |(J_\tau^{-1} (J_\tau + J_{\tau'}) \hat{\mathbf{n}}| \cdot |\hat{D}^2 \hat{u}_\tau| \\ &\leq \|I + J_\tau^{-1} J_{\tau'}\| \cdot |\hat{D}^2 \hat{u}_\tau| \\ &\leq c N^{-\alpha/2} |\hat{D}^2 \hat{u}_\tau|. \end{aligned}$$

Put the above two bounds and the bound from Lemma 2.1 into (22), we have

$$\begin{aligned} |I_1| &\leq c N^{-\alpha/2} \sum_{e=\tau \cup \tau'} |\tau| \cdot \|J_\tau^{-1}\|^2 \cdot |\hat{\nabla} \hat{v}_N| \cdot \int_{\hat{e}} |\hat{D}^2 \hat{u}_\tau| \\ &\leq c N^{-\alpha/2} \sum_{\tau \in \mathcal{T}_N} |\tau| \cdot \|J_\tau^{-1}\|^2 \cdot \|J_\tau\| \cdot |\nabla v_N| \cdot \left\{ \int_{\hat{\tau}} (|\hat{D}^2 \hat{u}_\tau|^2 + |\hat{D}^3 \hat{u}_\tau|^2) \right\}^{1/2} \\ &\leq c N^{-\alpha/2} \sum_{\tau \in \mathcal{T}_N} |\tau|^{1/2} \cdot \|J_\tau^{-1}\|^2 \cdot \|J_\tau\| \cdot |\nabla v_N| \\ &\quad \cdot \left\{ \int_{\tau} (\|J_\tau^T Q_2 J_\tau\|^2 + \|J_\tau^T Q_3 J_\tau\|^3) \right\}^{1/2} \\ &\leq c N^{-\alpha/2} \left\{ \sum_{\tau \in \mathcal{T}_N} \int_{\tau} \|J_\tau^{-1}\|^4 \|J_\tau\|^2 (\|J_\tau^T Q_2 J_\tau\|^2 + \|J_\tau^T Q_3 J_\tau\|^3) \right\}^{1/2} \cdot \|\nabla v_N\|_{L^2(\Omega)}, \end{aligned}$$

where in the second step we used the trace theorem to convert the integral on edge $\hat{e}$ into the integral on element $\hat{\tau}$.

Combining the estimates for I_1 , I_2 , and (4), we have for any $v_N \in S_N$,

$$(23) \quad |a(u - u_I, v_N)| \leq c \|\nabla v_N\| \cdot \left\{ \sum_{\tau} \int_{\tau} (1 + N^{-\alpha} \|J_{\tau}^{-1}\|^4 \|J_{\tau}\|^2) \cdot \|J_{\tau}^T Q_2 J_{\tau}\|^2 + \|J_{\tau}^{-1}\|^4 \|J_{\tau}\|^2 \cdot \|J_{\tau}^T Q_3 J_{\tau}\|^3 \right\}^{1/2},$$

which completes the proof of (17).

Finally, by the assumption that $\mathcal{T}_N$ is quasi-uniform under metric M , we have from (10) that

$$J_{\tau} \simeq (C_M N^{-1})^{1/2} F_{\tau}.$$

By putting it into (17) and use the continuity of M , we have the error bound (18) in integral form. $\square$

Remark 4.1. Similar to Remark 3.1, if $\nabla^2 u$ is positive definite on Ω , and the triangulation is quasi-uniform under metric $M = c(\lambda_{max}/\lambda_{min})^{1/4} \cdot \nabla^2 u$, we may also derive an error bound for $\|u_N - u_I\|_{1,\Omega}$ in terms of the eigenvalues of $\nabla^2 u$ and Q_3 . In particular, if $\alpha < 1$, then the leading term on the right-hand side in (18) is of magnitude $cN^{-(1+\alpha)/2} (\|\lambda_{min}^{1/4} \lambda_{max}^{3/4}\|_{L^1(\Omega)} \cdot \|\lambda_{min}^{-3/4} \lambda_{max}^{7/4}\|_{L^1(\Omega)})^{1/2}$.

5. SUPERCONVERGENCE OF POSTPROCESSING

Let $\mathcal{P}$ be the global L^2 -projection onto S_N . We discuss in this section the superconvergence of $\mathcal{P}(\nabla u_N)$ to ∇u . A key step leading to this result is the superconvergent error estimate for $\|\nabla u - \mathcal{P}(\nabla u_I)\|$. To obtain this estimate, we first derive an identity for the difference $\nabla u_q - \nabla u_I$, where u_q is the piecewise quadratic “projection based interpolation” of u defined as

$$u_q = u_I + \sum_{k=1}^3 \alpha_k q_k \quad \text{with} \quad \alpha_k = \frac{6}{\ell_k} \int_{e_k} (u - u_I),$$

where $q_k = \phi_{k+1} \phi_{k-1}$ is the side basis associated with edge e_k . Clearly, $\int_{e_k} (u - u_q) = 0$ on each edge e_k . In addition, u_q is continuous on Ω , and vanishes on $\partial\Omega$ if u does.

Lemma 5.1. *For any u with $u|_{\partial\Omega} = 0$ and any $\mathbf{w} \in (S_N)^2$, we have*

$$(24) \quad (\nabla u_q - \nabla u_I, \mathbf{w})_{\Omega} = \sum_{e_k \in \Omega} (\mathbf{w}(m_k) \cdot \mathbf{n}_k) \ell_k \left\{ \frac{1}{\ell_{k+1}} \int_{e_{k+1}} + \frac{1}{\ell_{k-1}} \int_{e_{k-1}} - \frac{1}{\ell_{k'+1}} \int_{e_{k'+1}} - \frac{1}{\ell_{k'-1}} \int_{e_{k'-1}} \right\} (u - u_I) + \sum_{e_k \in \partial\Omega} (\mathbf{w}(m_k) \cdot \mathbf{n}_k) \ell_k \left\{ \frac{1}{\ell_{k+1}} \int_{e_{k+1}} + \frac{1}{\ell_{k-1}} \int_{e_{k-1}} \right\} (u - u_I),$$

where the summations are over all interior edges $e_k \in \Omega$ and boundary edges $e_k \in \partial\Omega$ in the partition, respectively. $\mathbf{w}(m_k)$ is the value of $\mathbf{w}$ at the midpoint m_k of each e_k , refer to Figure 2.

FIGURE 2. Element pair around an edge $e_k = \tau \cup \tau'$.

Proof. By the facts that $u|_{\partial\Omega} = 0$ and that $\mathbf{w}$ is linear on each element, we have

$$(25) \quad (\nabla u_q - \nabla u_I, \mathbf{w})_{\Omega} = -(u_q - u_I, \nabla \cdot \mathbf{w})_{\Omega} = - \sum_{\tau} (\nabla \cdot \mathbf{w}|_{\tau}) (u_q - u_I, 1)_{\tau}.$$

Note that $(q_k, 1)_{\tau} = \frac{1}{12} |\tau|$, we have

$$(26) \quad (u_q - u_I, 1)_{\tau} = \left(\sum_{k=1}^3 \alpha_k q_k, 1 \right)_{\tau} = \frac{|\tau|}{12} \sum_{k=1}^3 \alpha_k = \frac{|\tau|}{2} \sum_{k=1}^3 \frac{1}{\ell_k} \int_{e_k} (u - u_I).$$

On the other hand, let $\mathbf{w}|_{\tau} = \sum_{k=1}^3 \mathbf{w}_k \phi_k$. By $\nabla \phi_k = -(\nabla \phi_{k-1} + \nabla \phi_{k+1})$, we have

$$\nabla \cdot \mathbf{w}|_{\tau} = \sum_{k=1}^3 \mathbf{w}_k \cdot \nabla \phi_k = - \sum_{k=1}^3 \mathbf{w}_k \cdot (\nabla \phi_{k-1} + \nabla \phi_{k+1}) = - \sum_{k=1}^3 (\mathbf{w}_{k-1} + \mathbf{w}_{k+1}) \cdot \nabla \phi_k.$$

Furthermore, by $\nabla \phi_k = -\frac{\ell_k}{2|\tau|} \mathbf{n}_k$ and $\mathbf{w}(m_k) = \frac{1}{2}(\mathbf{w}_{k+1} + \mathbf{w}_{k-1})$,

$$(27) \quad \nabla \cdot \mathbf{w}|_{\tau} = -\frac{1}{|\tau|} \sum_{k=1}^3 \ell_k \mathbf{w}(m_k) \cdot \mathbf{n}_k.$$

By putting the above equation into (25), we have

$$\begin{aligned} (\nabla u_q - \nabla u_I, \mathbf{w})_{\Omega} &= \sum_{\tau} \frac{1}{|\tau|} (u_q - u_I, 1)_{\tau} \cdot \sum_{k=1}^3 \ell_k (\mathbf{w}(m_k) \cdot \mathbf{n}_k)|_{\tau} \\ &= \sum_{e_k \in \Omega} (\mathbf{w}(m_k) \cdot \mathbf{n}_k|_{\tau}) \ell_k \left(\frac{1}{|\tau|} (u_q - u_I, 1)_{\tau} - \frac{1}{|\tau'|} (u_q - u_I, 1)_{\tau'} \right) \\ &\quad + \sum_{e_k \in \partial\Omega} \frac{\ell_k}{|\tau|} (\mathbf{w}(m_k) \cdot \mathbf{n}_k|_{\tau}) (u_q - u_I, 1)_{\tau}, \end{aligned}$$

where in the second step we used the fact that $\mathbf{n}_k|_{\tau} = -\mathbf{n}_{k'}|_{\tau'}$, and $\mathbf{w}$ is continuous across element edges. Finally, by putting (26) into the right-hand side of the above equation, we have formula (24). $\square$

Theorem 5.1. *Suppose that D^2u and D^3u satisfy assumption (12), and that each pair of adjacent elements in $\{\mathcal{T}_N\}$ forms an $O(N^{-(1+\alpha)/2})$ -approximate parallelogram. Then*

$$(28) \quad \begin{aligned} \|\nabla u - \mathcal{P}\nabla u_I\| &\leq c \left\{ \sum_{\tau} \int_{\tau} \|J_{\tau}^{-1}\|^2 \cdot (N^{-\alpha} \|J_{\tau}^T Q_2 J_{\tau}\|^2 + \|J_{\tau}^T Q_3 J_{\tau}\|^3) \right\}^{1/2} \\ &\quad + c \left\{ \sum_{\tau \in \Omega_h^b} \int_{\tau} \|J_{\tau}^{-1}\|^2 \cdot \|J_{\tau}^T Q_2 J_{\tau}\|^2 \right\}^{1/2}, \end{aligned}$$

where $\Omega_h^b = \bigcup_{\tau \cap \partial\Omega \neq \emptyset} \tau$ is the union of elements next to the boundary $\partial\Omega$. In addition, if partition $\mathcal{T}_N$ is quasi-uniform under metric M , then we have

$$(29) \quad \begin{aligned} \|\nabla u - \mathcal{P}\nabla u_I\| &\leq cN^{-1/2} \left\{ \int_{\Omega} \|F^{-1}\|^2 \cdot (C_M N^{-\alpha} \|F^T Q_2 F\|^2 \right. \\ &\quad \left. + (C_M)^2 N^{-1} \|F^T Q_3 F\|^3) \right\}^{1/2} \\ &\quad + cN^{-1/2} \left\{ C_M \int_{\Omega_h^b} \|F^{-1}\|^2 \cdot \|F^T Q_2 F\|^2 \right\}^{1/2}, \end{aligned}$$

where F and C_M are determined by M as in (9) and (11).

Proof. By the triangle inequality, we have

$$(30) \quad \|\nabla u - \mathcal{P}\nabla u_I\| \leq \|\nabla u - \mathcal{P}\nabla u\| + \|\mathcal{P}(\nabla u - \nabla u_q)\| + \|\mathcal{P}(\nabla u_q - \nabla u_I)\|.$$

The first term on the right-hand side can be bounded by the interpolation error estimate. Note on each element τ ,

$$\Pi_1(\nabla u) = \hat{\Pi}_1(J_{\tau}^{-1} \hat{\nabla} \hat{u}) = J_{\tau}^{-1} (\hat{\Pi}_1 \hat{\nabla} \hat{u}).$$

Hence

$$\begin{aligned} \|\nabla u - \mathcal{P}\nabla u\|^2 &\leq \|\nabla u - \Pi_1(\nabla u)\|^2 = \sum_{\tau} \int_{\tau} |\nabla u - \Pi_1(\nabla u)|^2 \\ &= \sum_{\tau} |\tau| \int_{\hat{\tau}} |J_{\tau}^{-1} \hat{\nabla} \hat{u} - J_{\tau}^{-1} (\hat{\Pi}_1 \hat{\nabla} \hat{u})|^2 \\ &\leq \sum_{\tau} |\tau| \int_{\hat{\tau}} \|J_{\tau}^{-1}\|^2 \cdot |\hat{\nabla} \hat{u} - \hat{\Pi}_1(\hat{\nabla} \hat{u})|^2 \\ &\leq \sum_{\tau} |\tau| \int_{\hat{\tau}} \|J_{\tau}^{-1}\|^2 \cdot |\hat{D}^3 \hat{u}|^2 \\ &\leq \sum_{\tau} |\tau| \int_{\hat{\tau}} \|J_{\tau}^{-1}\|^2 \cdot \|J_{\tau}^T Q_3 J_{\tau}\|^3 \\ &= \sum_{\tau} \int_{\tau} \|J_{\tau}^{-1}\|^2 \cdot \|J_{\tau}^T Q_3 J_{\tau}\|^3, \end{aligned}$$

where in the second last step, we used the assumption (12) about the anisotropic behavior of D^3u .

The second term on the right-hand side of (30) can also be bounded by using the error estimates for quadratic interpolation,

$$\begin{aligned}
 \|\mathcal{P}(\nabla u - \nabla u_q)\|^2 &\leq \|\nabla u - \nabla u_q\|^2 = \sum_{\tau} \int_{\tau} |\nabla u - \nabla u_q|^2 \\
 &\leq \sum_{\tau} |\tau| \int_{\hat{\tau}} \|J_{\tau}^{-1}\|^2 |\hat{\nabla} \hat{u} - \hat{\nabla} \hat{u}_q|^2 \\
 &\leq \sum_{\tau} |\tau| \int_{\hat{\tau}} \|J_{\tau}^{-1}\|^2 |\hat{D}^3 \hat{u}|^2 \\
 &\leq \sum_{\tau} \int_{\tau} \|J_{\tau}^{-1}\|^2 \|J_{\tau}^T Q_3 J_{\tau}\|^3.
 \end{aligned}$$

Now we estimate the third term $\|\mathcal{P}(\nabla u_q - \nabla u_I)\|$. By the fact that

$$\begin{aligned}
 (31) \quad \|\mathcal{P}(\nabla u_q - \nabla u_I)\| &= \sup_{\mathbf{w} \in (S_N)^2} \frac{|(\mathcal{P}(\nabla u_q - \nabla u_I), \mathbf{w})|}{\|\mathbf{w}\|} \\
 &= \sup_{\mathbf{w} \in (S_N)^2} \frac{|(\nabla u_q - \nabla u_I, \mathbf{w})|}{\|\mathbf{w}\|},
 \end{aligned}$$

we only have to bound $|(\nabla u_q - \nabla u_I, \mathbf{w})|$ by a certain multiple of $\|\mathbf{w}\|$ for any $\mathbf{w} \in (S_N)^2$. It follows from Lemma 5.1 that

$$\begin{aligned}
 (32) \quad &|(\nabla u_q - \nabla u_I, \mathbf{w})| \\
 &\leq \sum_{e_k \in \Omega} |\mathbf{w}(m_k)| \ell_k \left\{ \left| \frac{1}{\ell_{k+1}} \int_{e_{k+1}} (u - u_I) - \frac{1}{\ell_{k'+1}} \int_{e_{k'+1}} (u - u_I) \right| \right. \\
 &\quad \left. + \left| \frac{1}{\ell_{k-1}} \int_{e_{k-1}} (u - u_I) - \frac{1}{\ell_{k'-1}} \int_{e_{k'-1}} (u - u_I) \right| \right\} \\
 &\quad + \sum_{e_k \in \partial\Omega} |\mathbf{w}(m_k)| \ell_k \left\{ \left| \frac{1}{\ell_{k+1}} \int_{e_{k+1}} (u - u_I) + \frac{1}{\ell_{k-1}} \int_{e_{k-1}} (u - u_I) \right| \right\}.
 \end{aligned}$$

In order to estimate the difference between the averages of $u - u_I$ on opposite edges e_{k+1} and $e_{k'+1}$, we consider $Q = \tau \cup \tau'$ as the image of $\hat{Q} = [0, 1]^2$ under a bilinear transform $\mathbf{x} = \mathbf{B}(\xi)$. Then e_{k+1} and $e_{k'+1}$ are the images of two opposite sides of $\hat{Q}$, say $\eta = 1$ and $\eta = 0$. Furthermore, let $\tilde{u} = u \circ \mathbf{B}$ and $J_Q = \frac{\partial \mathbf{x}}{\partial \xi}$. Thus

$$\begin{aligned}
 (33) \quad &\frac{1}{\ell_{k+1}} \int_{e_{k+1}} (u - u_I) - \frac{1}{\ell_{k'+1}} \int_{e_{k'+1}} (u - u_I) \\
 &= \int_0^1 (\tilde{u} - \Pi_1 \tilde{u})(\xi, 1) d\xi - \int_0^1 (\tilde{u} - \Pi_1 \tilde{u})(\xi, 0) d\xi \\
 &= \int_0^1 \xi(\xi - 1) \frac{\partial^2 \tilde{u}}{\partial \xi^2}(\xi, 1) d\xi - \int_0^1 \xi(\xi - 1) \frac{\partial^2 \tilde{u}}{\partial \xi^2}(\xi, 0) d\xi \\
 &= \int_{\hat{Q}} \xi(\xi - 1) \frac{\partial^3 \tilde{u}}{\partial \xi^2 \partial \eta}(\xi, \eta) d\eta d\xi.
 \end{aligned}$$

Now we bound the third order derivative in the above integral. Note that $|D^3 \tilde{u}| \leq \sup_{|\mathbf{s}|=1} \left| \frac{\partial^3 \tilde{u}}{\partial \mathbf{s}^3} \right|$. By using the tensor form of the chain rule for derivatives (see, e.g., Theorem 37.1 in [9]), we have for any unit vector $\mathbf{s} = [s_1, s_2]^T$,

$$\frac{\partial^3 \tilde{u}}{\partial \mathbf{s}^3} = D^3 u(\mathbf{x}_{\mathbf{s}}, \mathbf{x}_{\mathbf{s}}, \mathbf{x}_{\mathbf{s}}) + 3 D^2 u(\mathbf{x}_{\mathbf{s}}, \mathbf{x}_{\mathbf{s}\mathbf{s}}) + Du(\mathbf{x}_{\mathbf{s}\mathbf{s}\mathbf{s}}).$$

Here

$$\mathbf{x}_s = \frac{\partial \mathbf{x}}{\partial \mathbf{s}} = \frac{\partial \mathbf{x}}{\partial \boldsymbol{\xi}} \cdot \mathbf{s} = J_Q \mathbf{s}$$

and $\mathbf{x}_{ss\mathbf{s}} = \frac{\partial^3 \mathbf{x}}{\partial \mathbf{s}^3} = 0$, since $\mathbf{B}$ is bilinear. By formula (5) for $\mathbf{B}(\boldsymbol{\xi})$,

$$\mathbf{x}_{ss\mathbf{s}} = \frac{\partial^2 \mathbf{x}}{\partial \mathbf{s}^2} = s_1 s_2 \mathbf{d},$$

where $\mathbf{d}$ is the vector connecting the midpoints of two diagonals of Q . Therefore, we have

$$|D^3 u(\mathbf{x}_s, \mathbf{x}_s, \mathbf{x}_s)| \leq |D^3 u(J_Q \mathbf{s}, J_Q \mathbf{s}, J_Q \mathbf{s})| \leq [J_Q \mathbf{s} \cdot Q_3 J_Q \mathbf{s}]^{3/2} \leq \|J_Q^T Q_3 J_Q\|^{3/2}$$

and

$$|D^2 u(\mathbf{x}_s, \mathbf{x}_{ss\mathbf{s}})| \leq |J_Q \mathbf{s} \cdot D^2 u \mathbf{d}| \leq \|J_Q^T D^2 J_Q\| |J_Q^{-1} \mathbf{d}| \leq cN^{-\alpha/2} \|J_Q^T Q_2 J_Q\|,$$

where in the last step we used Lemma 2.3 under the assumption that $Q = \tau \cup \tau'$ forms an $O(N^{-(1+\alpha)/2})$ -approximate parallelogram. Combining them together, we have

$$|D^3 \tilde{u}| \leq cN^{-\alpha/2} \|J_Q^T Q_2 J_Q\| + \|J_Q^T Q_3 J_Q\|^{3/2}.$$

Therefore, we may derive from (33) that

$$\begin{aligned} & \left| \frac{1}{\ell_{k+1}} \int_{e_{k+1}} (u - u_I) - \frac{1}{\ell_{k'+1}} \int_{e_{k'+1}} (u - u_I) \right| \\ & \leq \int_{\hat{Q}} (cN^{-\alpha/2} \|J_Q^T Q_2 J_Q\| + \|J_Q^T Q_3 J_Q\|^{3/2}) \\ & \leq \int_Q (cN^{-\alpha/2} \|J_Q^T Q_2 J_Q\| + \|J_Q^T Q_3 J_Q\|^{3/2}) \cdot \det(J_Q^{-1}). \end{aligned}$$

Note that

$$\|J_\tau^{-1} J_Q(\boldsymbol{\xi})\| \leq \text{const}, \quad \forall \boldsymbol{\xi} \in \hat{Q},$$

which is obviously true in the case when τ and Q are shape regular. It is also true when they are anisotropic, because $J_\tau^{-1} J_Q$ is invariant under any affine transformation of Q . Therefore,

$$\|J_Q^T Q_2 J_Q\| \leq c \|J_\tau^T Q_2 J_\tau\|, \quad \text{on } \tau.$$

The term $\|J_Q^T Q_3 J_Q\|$ can be bounded similarly. In addition,

$$\det(J_Q^{-1}) \leq c \det(J_\tau^{-1}) \leq \frac{c}{|\tau|}.$$

Hence we have

$$\begin{aligned} & \left| \frac{1}{\ell_{k+1}} \int_{e_{k+1}} (u - u_I) - \frac{1}{\ell_{k'+1}} \int_{e_{k'+1}} (u - u_I) \right| \\ & \leq \frac{c}{|\tau|} \int_\tau (N^{-\alpha/2} \|J_\tau^T Q_2 J_\tau\| + \|J_\tau^T Q_3 J_\tau\|^{3/2}). \end{aligned}$$

Put it into (32) and note that

$$\frac{\ell_k}{|\tau|} = \frac{4}{3\sqrt{3}} \cdot \frac{|J_\tau \hat{e}_k|}{\det(J_\tau)} \leq \frac{4}{3} \cdot \frac{\|J_\tau\|}{\det(J_\tau)} \leq \frac{4}{3} \|J_\tau^{-1}\|,$$

then we have

$$\begin{aligned}
 |(\nabla u_q - \nabla u_I, \mathbf{w})| &\leq c \sum_{\tau} \left(\sum_{k=1}^3 \frac{\ell_k}{|\tau|} |\mathbf{w}(m_k)| \right) \\
 &\quad \cdot \int_{\tau} (N^{-\alpha/2} \|J_{\tau}^T Q_2 J_{\tau}\| + \|J_{\tau}^T Q_3 J_{\tau}\|^{3/2}) \\
 (34) \quad &\leq c \sum_{\tau} \int_{\tau} \left(\sum_{k=1}^3 |\mathbf{w}(m_k)| \right) \cdot \|J_{\tau}^{-1}\| \cdot (N^{-\alpha/2} \|J_{\tau}^T Q_2 J_{\tau}\| + \|J_{\tau}^T Q_3 J_{\tau}\|^{3/2}) \\
 &\leq c \|\mathbf{w}\| \cdot \left\{ \sum_{\tau} \int_{\tau} \|J_{\tau}^{-1}\|^2 \cdot (N^{-\alpha} \|J_{\tau}^T Q_2 J_{\tau}\|^2 + \|J_{\tau}^T Q_3 J_{\tau}\|^3) \right\}^{1/2},
 \end{aligned}$$

where we have used the fact that

$$\int_{\tau} \left(\sum_{k=1}^3 |\mathbf{w}(m_k)| \right)^2 \leq c \int_{\tau} |\mathbf{w}|^2.$$

This completes the estimate for the first sum on the right-hand side of (32).

For the second sum on the right-hand side of (32) involving edges on the domain boundary, note that $u - u_I = 0$ on $\partial\Omega$, thus we have for any linear function p_1 ,

$$\begin{aligned}
 \left| \frac{1}{\ell_{k+1}} \int_{e_{k+1}} (u - u_I) + \frac{1}{\ell_{k-1}} \int_{e_{k-1}} (u - u_I) \right| &= \left| \int_{\partial\hat{\tau}} (\hat{u} - \hat{u}_I) \right| \\
 &= \left| \int_{\partial\hat{\tau}} (\hat{u} + p_1 - \hat{\Pi}_1(\hat{u} + p_1)) \right| \\
 &\leq c \|\hat{u} + p_1\|_{C(\hat{\tau})} \leq c \|\hat{u} + p_1\|_{H^2(\hat{\tau})}.
 \end{aligned}$$

Since p_1 is arbitrary, it follows from Bramble-Hilbert Lemma [9] that

$$\begin{aligned}
 \left| \frac{1}{\ell_{k+1}} \int_{e_{k+1}} (u - u_I) + \frac{1}{\ell_{k-1}} \int_{e_{k-1}} (u - u_I) \right| &\leq c \inf_{\forall p_1} \|\hat{u} + p_1\|_{H^2(\hat{\tau})} \leq c |\hat{u}|_{H^2(\hat{\tau})} \\
 &= c \left\{ \int_{\hat{\tau}} |\hat{D}^2 \hat{u}|^2 \right\}^{1/2} \\
 &\leq c \left\{ \int_{\tau} \frac{1}{|\tau|} \|J_{\tau}^T Q_2 J_{\tau}\|^2 \right\}^{1/2}.
 \end{aligned}$$

Therefore, we have

$$\begin{aligned}
 \sum_{e_k \in \partial\Omega} |\mathbf{w}(m_k)| \ell_k \left\{ \left| \frac{1}{\ell_{k+1}} \int_{e_{k+1}} (u - u_I) + \frac{1}{\ell_{k-1}} \int_{e_{k-1}} (u - u_I) \right| \right\} \\
 (35) \quad &\leq c \sum_{e_k \in \partial\Omega} |\mathbf{w}(m_k)| |\tau|^{1/2} \cdot \left\{ \int_{\tau} \frac{\ell_k^2}{|\tau|^2} \|J_{\tau}^T Q_2 J_{\tau}\|^2 \right\}^{1/2} \\
 &\leq c \left\{ \sum_{e_k \in \partial\Omega} |\mathbf{w}(m_k)|^2 |\tau| \right\}^{1/2} \cdot \left\{ \sum_{\tau \in \Omega_h^b} \int_{\tau} \|J_{\tau}^{-1}\|^2 \|J_{\tau}^T Q_2 J_{\tau}\|^2 \right\}^{1/2} \\
 &\leq c \|\mathbf{w}\| \cdot \left\{ \sum_{\tau \in \Omega_h^b} \int_{\tau} \|J_{\tau}^{-1}\|^2 \|J_{\tau}^T Q_2 J_{\tau}\|^2 \right\}^{1/2}.
 \end{aligned}$$

By combining (34) and (35), we complete the estimate of the right hand side of (32), and hence the proof of (28).

The proof of (29) is similar to that of (18). $\square$

By using Theorems 4.1 and 5.1, and the triangle inequality

$$(36) \quad \begin{aligned} \|\nabla u - \mathcal{P}\nabla u_N\| &\leq \|\nabla u - \mathcal{P}\nabla u_I\| + \|\mathcal{P}(\nabla u_N - \nabla u_I)\| \\ &\leq \|\nabla u - \mathcal{P}\nabla u_I\| + \|\nabla u_N - \nabla u_I\|, \end{aligned}$$

we have immediately the following result on the superconvergence of the post-processed finite element solution.

Theorem 5.2. *Under the same assumption of Theorem 5.1, we have*

$$(37) \quad \begin{aligned} \|\nabla u - \mathcal{P}\nabla u_N\| &\leq c \left\{ \sum_{\tau} \int_{\tau} (1 + N^{-\alpha} \|J_{\tau}^{-1}\|^4 \|J_{\tau}\|^2) \cdot \|J_{\tau}^T Q_2 J_{\tau}\|^2 \right. \\ &\quad \left. + \|J_{\tau}^{-1}\|^4 \|J_{\tau}\|^2 \cdot \|J_{\tau}^T Q_3 J_{\tau}\|^3 \right\}^{1/2} \\ &\quad + c \left\{ \sum_{\tau \in \Omega_h^b} \int_{\tau} \|J_{\tau}^{-1}\|^2 \cdot \|J_{\tau}^T Q_2 J_{\tau}\|^2 \right\}^{1/2}. \end{aligned}$$

In addition, if the partition $\mathcal{T}_N$ is quasi-uniform under metric M , then

$$(38) \quad \begin{aligned} \|\nabla u - \mathcal{P}\nabla u_N\| &\leq c N^{-1/2} \left\{ \int_{\Omega} ((C_M)^2 N^{-1} + C_M N^{-\alpha/2} \|F^{-1}\|^4 \|F\|^2) \right. \\ &\quad \left. \cdot \|F^T Q_2 F\|^2 + (C_M)^2 N^{-1} \|F^{-1}\|^4 \|F\|^2 \cdot \|F^T Q_3 F\|^3 \right\}^{1/2} \\ &\quad + c N^{-1/2} \left\{ C_M \int_{\Omega_h^b} \|F^{-1}\|^2 \cdot \|F^T Q_2 F\|^2 \right\}^{1/2}. \end{aligned}$$

Remark 5.1. In practical computation, the mesh metric M is typically chosen to be a certain scalar multiple of Q_2 . In this case, the integrals in (37) involving elements next to the domain boundary are only an $O(N^{-1/2})$ portion of the same integral over the entire Ω , and the second integral in (38) is of order $O(N^{-1/2})$, which implies $\|\nabla u - \mathcal{P}\nabla u_N\|$ converges at the rate of $O(N^{-(1+\min(1,\alpha))/2})$. It also implies the asymptotic exactness (as $N \rightarrow \infty$) of the error estimator $\|\nabla u_N - \mathcal{P}\nabla u_N\|$ for solutions satisfying condition $\|\nabla u - \nabla u_N\| \geq cN^{-1/2}$.

Remark 5.2. Comparing the estimates for $\|\nabla u_N - \nabla u_I\|$ in Theorem 4.1 and $\|\nabla u - \mathcal{P}(\nabla u_I)\|$ in Theorem 5.1, we note that there is an extra factor $\text{cond}(J_{\tau}) = \|J_{\tau}\| \|J_{\tau}^{-1}\|$ for the bound of the former. This factor measures the aspect ratio of τ . Thus for highly anisotropic meshes, $\|\nabla u_N - \nabla u_I\|$ can be much larger than $\|\nabla u - \mathcal{P}(\nabla u_I)\|$, even though both of them are of the same convergence rate. This is a new feature not existing in the superconvergence estimates on isotropic meshes. We believe this extra factor is not due to the technical treatment in our analysis, but rather the fact that $\|\nabla u_N - \nabla u_I\|$ is basically the H^1 -norm of the H^1 -projection of $u - u_I$, while $\|\nabla u - \mathcal{P}(\nabla u_I)\|$ is mainly the L^2 -error of an L^2 -projection. While L^2 -projection is insensitive to the aspect ratio of elements, H^1 -projection is not. Numerical results in the next section seem to support this observation.

6. NUMERICAL RESULTS

We present in this section some numerical results for the linear finite element solution and its error estimates of the model problem. We choose in (6) the domain $\Omega = [0, 1]^2$, and the coefficients $A = I_2$, $\mathbf{b} = \mathbf{0}$, $d = 1$. The right-hand side function f and the Dirichlet boundary condition are selected so that (6) admits the following exact solutions:

Example (1). $u(x, y) = \tanh[-50(r - 0.3)]$, where r is the distance to the center of Ω .

Example (2). $u(x, y) = [\cosh(kx) + \cosh(ky)] / [2 \cosh(k)]$ with $k = 50$.

The first example involves an internal layer along a circle in the domain, while the second one a boundary layer along the top and right sides of the domain.

To obtain the adaptive meshes and the finite element solutions, we start with a Delaunay triangulation of Ω , and calculate the finite element solution based on it. Then we recover the second order derivatives of the solution, and define a mesh metric on Ω . We use a two-dimensional anisotropic generator, **bamg** [13], to create an adaptive mesh that is quasi-uniform under the mesh metric, and calculate again the finite element solution based on the adaptive mesh. This process is repeated 20 times, which is hand-picked but seems sufficient to obtain a convergent mesh and the finite element solution. In practical computation, more judicious criterions, e.g., a threshold on the differences between the meshes or solutions in consecutive iterates, should be applied to control this iterative process.

In order to defined the mesh metrics, we first approximate the second derivatives $\nabla^2 u$ by

$$(39) \quad \tilde{H} = \nabla(P(\nabla u_N)),$$

where P is the L^2 -projection onto S_N . We choose here the L^2 -projection of u_N because it is the quantity involved in our analysis; see Theorem 5.2. Other types of postprocessing, e.g., local averaging, are also possible. Since $\tilde{H}$ is piecewise constant, it is further approximated by a continuous piecewise linear function, still denoted by $\tilde{H}$, with the nodal values equal to the weighted average over the element patch around each node. The anisotropic behavior of $\nabla^2 u$ is then characterized by

$$Q_2 = \text{abs}(\tilde{H}) + \delta I,$$

where $\text{abs}(\tilde{H})$ is the matrix of the same eigenvectors as $\tilde{H}$, but with eigenvalues the absolute value of $\tilde{H}$'s. $\delta \geq 0$ is a user defined parameter to avoid the mesh metric based on it being singular in case $\nabla^2 u$ becomes singular. A smaller δ makes Q_2 a more precise characterization of the anisotropic behavior of $\nabla^2 u$ and the mesh metric more effective, while a larger δ results in a less aggressive anisotropic refinement and is easier for refinement. An optimal value of δ is problem dependent and should be selected dynamically. However, this is not our main objective in this paper. We choose for simplicity in all our numerical examples $\delta = 1$ (a small value compared to the magnitude of the second derivatives of u).

The mesh metric M is then defined as in [5, 7] for the minimization of the H^1 -error for the linear interpolation of u on anisotropic meshes,

$$(40) \quad M = c [\lambda_{\max}/\lambda_{\min}]^{1/4} Q_2,$$

where $\lambda_{\max}$ and $\lambda_{\min}$ are the largest and smallest eigenvalues of Q_2 , and c is a constant to control the total number of elements.

In order to check the $O(N^{-(1+\alpha)/2})$ -approximate parallelogram property for the adaptive meshes used in our computation, we introduce a quantity $\mathcal{M}_{ap}$ to measure how close a pair of adjacent elements is to a parallelogram. Suppose τ_1 and τ_2 share a common edge e . J_1 and J_2 are the Jacobian of the affine mapping from an equilateral standard element $\hat{\tau}$ to τ_1 and τ_2 , respectively, with the two vertices

opposite to e being the images of the same vertex of $\hat{\tau}$. Define

$$(41) \quad \mathcal{M}_{ap}(e) = \frac{1}{2} (\| I + J_1^{-1} J_2 \| + \| I + J_2^{-1} J_1 \|),$$

where the matrix norm is chosen to be Frobenius norm for easy calculation. Note that when τ_1 and τ_2 form an $O(N^{-(1+\alpha)/2})$ -approximate parallelogram, both terms on the right-hand side of (41) are of magnitude $O(N^{-\alpha/2})$ (see Section 2), thus $\mathcal{M}_{ap} = O(N^{-\alpha/2})$. We use their average in (41) to produce a measurement symmetric for both sides of an edge. As noted in Lemma 2.2, this measure is invariant under any affine mapping of the element pair. When τ_1 and τ_2 form an exact parallelogram, $\mathcal{M}_{ap}(e_k) = 0$. When they are congruent but form a triangle, $\mathcal{M}_{ap} = \frac{2\sqrt{3}}{3} \approx 1.155$. If τ_1 is shape regular but τ_2 is of a high aspect ratio, then $\mathcal{M}_{ap} \gg 1$.

We notice that Bank and Xu established in [4] the convergence of $\tilde{H}$ (with some additional smoothing steps to improve the convergence rate) to $\nabla^2 u$ for quasi-uniform meshes. In the case of adaptive anisotropic meshes, there is no such theory yet, even though the idea of using $\tilde{H}$ for mesh refinement has been in practice for a long time; see, e.g., [8, 10–12]. When $\tilde{H}$ is a good approximation of $\nabla^2 u$, the above defined mesh metric will be optimal for minimizing the H^1 semi-norm of the interpolation error, and thus for minimizing the a priori bound of the finite element solution error as stated in Theorem 3.1.

We examine the following quantities involved at various stages of our analysis:

$$\begin{aligned} E_{u_N} &= \|\nabla u - \nabla u_N\|, & E_{NpN} &= \|\nabla u_N - P(\nabla u_N)\|, \\ E_{pu_I} &= \|\nabla u - P(\nabla u_I)\|, & E_{pu_N} &= \|\nabla u - P(\nabla u_N)\|, \\ E_{NI} &= \|\nabla u_N - \nabla u_I\|, & E_{pNI} &= \|P(\nabla u_N - \nabla u_I)\|. \end{aligned}$$

In addition, we check the effectivity factor η of the error estimator E_{NpN} :

$$\eta = E_{NpN} / E_{u_N}.$$

For each of the above errors E and $|\eta - 1|$, we estimate its rate of convergence β by linear regression of $\ln E$ vs. $\ln(1/\sqrt{N})$, i.e.,

$$E \approx cN^{-\beta/2}.$$

For comparison, we also report the results for solutions based on uniform meshes obtained from bisecting uniform square meshes on Ω with the 135° diagonals. Note that these uniform meshes can be considered as quasi-uniform under mesh metric I and each pair of the adjacent elements forms an $O(N^{-(1+\alpha)/2})$ -approximate parallelogram with $\alpha = \infty$. In this case, E_{NI} will be of order N^{-1} according to Theorem 4.1, and E_{pu_I} and E_{pu_N} will also be of order N^{-1} in the absence of boundary errors according to Theorems 5.1 and 5.2.

We list in Tables 1 and 2 the results for Example (1) with adaptive and uniform meshes, respectively, and Tables 3 and 4 for Example (2) similarly. Clearly, in all the cases, E_{u_N} converges linearly, i.e., at the rate of $O(N^{-1/2})$, and the effectivity factor $\eta \rightarrow 1$, which indicates that the error estimator E_{NpN} is asymptotically exact. In addition, E_{NI} converges quadratically on uniform meshes as predicted in Theorem 4.1 with $\alpha = \infty$. On adaptive meshes, the superconvergence of E_{NI} is much weaker, with $\alpha \approx 0.12$ only. This result is similar to that reported in [3] for the case of adaptive meshes composed of shape regular elements. The slower convergence rate with adaptive meshes results mainly from the fact that there are elements in the triangulation that do not form approximate parallelograms due to

TABLE 1. Example (1) based on adaptive meshes generated from **bamg** with mesh metrics defined by (40).

N	E_{u_N}	E_{NpN}	E_{pu_I}	E_{pu_N}	E_{NI}	E_{pNI}	$ \eta - 1 $
4070	6.62e-1	6.57e-1	1.95e-1	1.90e-1	1.36e-1	8.06e-2	7.81e-3
8024	4.45e-1	4.41e-1	1.09e-1	1.10e-1	9.11e-2	5.36e-2	8.74e-3
16082	3.01e-1	2.99e-1	6.18e-2	6.32e-2	6.16e-2	3.63e-2	7.78e-3
32114	2.10e-1	2.09e-1	3.64e-2	3.82e-2	4.07e-2	2.39e-2	7.09e-3
64112	1.46e-1	1.45e-1	2.26e-2	2.37e-2	2.78e-2	1.62e-2	6.78e-3
128762	1.04e-1	1.04e-1	1.41e-2	1.53e-2	1.99e-2	1.14e-2	6.65e-3
257114	7.40e-2	7.36e-2	8.82e-3	9.81e-3	1.30e-2	7.55e-3	6.11e-3
Order β	1.053	1.051	1.485	1.425	1.126	1.137	0.147

TABLE 2. Example (1) based on uniform meshes.

N	E_{u_N}	E_{NpN}	E_{pu_I}	E_{pu_N}	E_{NI}	E_{pNI}	$ \eta - 1 $
4050	3.33e+0	3.01e+0	2.04e+0	2.22e+0	6.33e-1	3.74e-1	9.63e-2
7938	2.45e+0	2.35e+0	1.09e+0	1.26e+0	3.64e-1	2.63e-1	4.35e-2
16200	1.76e+0	1.72e+0	5.28e-1	6.44e-1	1.91e-1	1.61e-1	2.23e-2
32258	1.26e+0	1.25e+0	2.61e-1	3.29e-1	9.90e-2	9.11e-2	1.15e-2
64082	9.01e-1	8.95e-1	1.30e-1	1.67e-1	5.07e-2	4.87e-2	5.92e-3
128018	6.39e-1	6.37e-1	6.46e-2	8.38e-2	2.56e-2	2.51e-2	3.00e-3
256328	4.52e-1	4.52e-1	3.22e-2	4.20e-2	1.29e-2	1.27e-2	1.51e-3
Order β	0.965	0.924	2.013	1.927	1.891	1.655	1.976

TABLE 3. Example (2) based on adaptive meshes generated from **bamg** with mesh metrics defined by (40).

N	E_{u_N}	E_{NpN}	E_{pu_I}	E_{pu_N}	E_{NI}	E_{pNI}	$ \eta - 1 $
4030	8.06e-2	7.88e-2	1.72e-2	1.75e-2	1.21e-2	7.39e-3	2.28e-2
8086	5.45e-2	5.37e-2	9.25e-3	9.52e-3	7.16e-3	4.46e-3	1.48e-2
15926	3.79e-2	3.74e-2	5.51e-3	5.92e-3	5.38e-3	3.26e-3	1.20e-2
32064	2.67e-2	2.65e-2	3.21e-3	3.39e-3	3.49e-3	2.00e-3	7.93e-3
63867	1.89e-2	1.88e-2	2.01e-3	2.20e-3	2.27e-3	1.34e-3	6.74e-3
129809	1.32e-2	1.31e-2	1.21e-3	1.32e-3	1.56e-3	8.78e-4	5.01e-3
256928	9.35e-3	9.31e-3	7.67e-4	8.70e-4	1.16e-3	6.51e-4	4.32e-3
Order β	1.031	1.023	1.485	1.438	1.129	1.178	0.797

geometric constraints, or the approximate parallelograms are rather poorly shaped; see Figure 3 and Figure 6 for a close-up look of typical meshes. Strictly speaking, these meshes belong to a more general type satisfying the so-called condition (α, σ) introduced in [3, 30] for shape regular meshes; namely, most element pairs in the partition form $O(N^{-(1+\alpha)/2})$ -approximate parallelograms and the measure of those that do not is at most $O(N^{-\sigma/2})$. The extension of our analysis to this type of meshes is straightforward. We choose not to include it here for the purpose of

TABLE 4. Example (2) based on uniform meshes.

N	E_{u_N}	E_{NpN}	E_{puI}	E_{puN}	E_{NI}	E_{pNI}	$ \eta - 1 $
4050	1.07e+00	7.60e-01	6.91e-01	6.91e-01	1.37e-04	1.33e-04	2.90e-01
7938	7.86e-01	6.16e-01	4.54e-01	4.54e-01	6.99e-05	6.86e-05	2.16e-01
16200	5.58e-01	4.72e-01	2.83e-01	2.83e-01	3.42e-05	3.39e-05	1.55e-01
32258	3.99e-01	3.54e-01	1.76e-01	1.76e-01	1.72e-05	1.71e-05	1.11e-01
64082	2.84e-01	2.61e-01	1.08e-01	1.08e-01	8.66e-06	8.63e-06	7.96e-02
128018	2.01e-01	1.90e-01	6.53e-02	6.53e-02	4.33e-06	4.33e-06	5.66e-02
256328	1.42e-01	1.37e-01	3.93e-02	3.93e-02	2.16e-06	2.16e-06	4.01e-02
Order β	0.975	0.835	1.387	1.387	2.000	1.986	0.957

simplifying our presentation. Instead, we report some statistics about the measurement $\mathcal{M}_{ap}$ for how close each pair of elements is to a parallelogram. We display in Figure 4 and Figure 7 the distribution of $\mathcal{M}_{ap}$ for a typical adaptive mesh generated by **bamg** in Examples (1) and (2), respectively. To further quantify this distribution, we plot in Figure 5(a) and Figure 8(a) the percentile of the element pairs whose $\mathcal{M}_{ap}$ are below a given value for each of the adaptive meshes reported in Table 1 and Table 3. It is clear from these two figures that as the mesh gets refined by **bamg**, $\mathcal{M}_{ap}$ decreases in general, and more and more element pairs become closer and closer to exact parallelograms. We plot in Figure 5(b) and Figure 8(b) the average value, $\overline{\mathcal{M}}_{ap}$, of $\mathcal{M}_{ap}$ over the entire mesh for each of the adaptive meshes. Based on linear regression of $\ln \overline{\mathcal{M}}_{ap}$ vs. $\ln N$, we find that $\overline{\mathcal{M}}_{ap} \approx O(N^{-\bar{\alpha}/2})$ with $\bar{\alpha} = 0.208$ and 0.278 for the adaptive meshes generated by **bamg** in Example (1) and Example (2), respectively.

It is also noted that E_{puI} exhibits superconvergence for both examples and both types of meshes as described in Theorem 5.1. However, with uniform meshes, the convergence rate is about $O(N^{-1})$ for Example (1) and $O(N^{-1.387/2})$ for Example (2). We believe the loss of the convergence rate in Example (2) is due to the boundary term in (28). Note that the solution of Example (1) is nearly 0 around the boundary of Ω , while it is not the case for Example (2). With adaptive meshes, E_{puI} converges at a rate $O(N^{-1.485/2})$ for both examples, and the effect of boundary terms for Example (2) seems to be masked by that of non-ideal approximate parallelograms.

By comparing E_{puI} to E_{NI} , we note that on uniform meshes E_{NI} is always smaller than E_{puI} , while on adaptive meshes it is the other way around. This phenomenon reflects their dependence on the element aspect ratio stated in Remark 5.2; namely, higher element aspect ratios have a stronger negative impact on the accuracy of E_{NI} than on E_{puI} . On the other hand, for both examples on adaptive meshes, E_{puI} seems to behave very similarly to E_{puN} , even though the bound for the latter involves an extra factor of the element aspect ratio $\text{cond}(F_\tau)$; see Theorems 5.1 and 5.2. It raises the question of whether this extra factor in the bound for E_{puN} is indeed necessary. In our analysis this extra factor is introduced due to the use of triangle inequality (36), which relies on the estimate (17) of E_{NI} . Unfortunately, we still do not know if it is possible not to use E_{NI} to provide E_{puN} with a similar bound as for E_{puI} .

FIGURE 3. Example (1): Left: An adaptive mesh generated by **bamg** based on metric (40) with the total number of elements $N = 4070$. Right: Its close-up look.

FIGURE 4. Example (1): Left: The distribution of measure $\mathcal{M}_{ap}$ defined in (41) for the closeness of a pair of adjacent elements to a parallelogram, on the adaptive mesh with $N = 4070$. The color around an edge e represents the magnitude of $\mathcal{M}_{ap}(e)$ for the pair sharing e . Right: Its close-up look.

FIGURE 5. Example (1): Left: Percentile curves for the number of element pairs whose M_{ap} are below a given value specified by the horizontal axis. The 7 curves from the bottom up represent the 7 adaptive meshes reported in Table 1 in increasing order of N . Right: The average value of M_{ap} vs. the number N of elements for each of the adaptive meshes.

FIGURE 6. Example (2): Left: An adaptive mesh generated by **bamg** based on metric (40) with the total number of elements $N = 4030$. Right: Its close-up look.

FIGURE 7. Example (2): Left: The distribution of measure $\mathcal{M}_{ap}$ defined in (41) for the closeness of a pair of adjacent elements to a parallelogram, on the adaptive mesh with $N = 4050$. The color around an edge e represents the magnitude of $\mathcal{M}_{ap}(e)$ for the pair sharing e . Right: Its close-up look.

FIGURE 8. Example (2): Left: Percentile curves for the number of element pairs whose $\mathcal{M}_{ap}$ are below a given value specified by the horizontal axis. The 7 curves from the bottom up represent the 7 adaptive meshes reported in Table 3 in increasing order of N . Right: The average value of $\mathcal{M}_{ap}$ vs. the number N of elements for each of the adaptive meshes.

7. CONCLUSION AND DISCUSSIONS

For the linear finite element method based on adaptive anisotropic meshes in two dimensions, we established the superconvergence of the finite element solutions to the interpolation of exact solutions in energy norm. We also proved the superconvergence of the recovery process using the global L^2 -projection of the gradients of the finite element solutions. Our analysis is based on the notion that the mesh is quasi-uniform under a Riemannian metric and the notion that each adjacent element pair forms an approximate parallelogram. These notions are extensions of the corresponding concepts in classical superconvergence studies on meshes composed of shape regular elements. In particular, our analysis follows the same methodology developed in Bank and Xu [3] for the case of quasi-uniform meshes, and can be viewed as an extension of their conclusion to the case of unstructured anisotropic meshes. Numerical examples involving both internal and boundary layers are in support of these superconvergence properties.

It should be noted that our analysis is concerned with the asymptotic behavior of the solution and its post-processing. Therefore, the conclusion is meaningful only when the mesh is sufficiently fine for the underlying problem. In the initial stages of an adaptive finite element solution where mesh resolution is far from enough, the error from data approximation can be more pronounced and the error estimates based on post-processing may not be accurate. In this case, except the anisotropic behavior of the PDEs is known a-priori, one should be extremely careful about the use of anisotropic coarse meshes, since a highly anisotropic element may introduce a much larger error than an isotropic element does if used inappropriately. Therefore, it is generally advised to begin an adaptive process with isotropic coarse meshes and other types of error estimators should be considered, too.

Several issues remain unresolved in this paper. First, it is observed in numerical experiments that the convergence of $\|\nabla u - \mathcal{P}(\nabla u_N)\|$ is similar to $\|\nabla u - \mathcal{P}(\nabla u_I)\|$ and both are less sensitive to high element aspect ratio than $\|\nabla u_N - \nabla u_I\|$. However, the estimate for $\|\nabla u - \mathcal{P}(\nabla u_N)\|$ involves an extra factor of the element aspect ratio than the one for $\|\nabla u - \mathcal{P}(\nabla u_I)\|$. It seems a sharper analysis tool is needed to remove this extra factor if it is indeed non-essential.

Second, with unstructured adaptive meshes the superconvergence for both $\|\nabla u_N - \nabla u_I\|$ and $\|\nabla u - \mathcal{P}(\nabla u_N)\|$ is rather weak due to the non-uniformity of the elements. Bank and Xu proposed in [4] to reduce the dependence on mesh uniformity and strengthen the superconvergence of the gradient recovery by employing several cycles of multigrid type smoothing in addition to the L^2 -projection. Their numerical results showed such a treatment may improve the superconvergence substantially. This idea is yet to be tested for the case of anisotropically refined meshes. Furthermore, we hope that with the improved superconvergence property for the recovered gradients the convergence for the approximation of $\nabla^2 u$ by $\tilde{H}$ in (39) can be established in the case of anisotropic meshes.

There are several other types of gradient recovery techniques used in practice. The most popular one is probably Zienkiewicz-Zhu's superconvergence patch recovery (SPR) based on local averaging [36, 37], and a more recent variation, the polynomial preserving recovery (PPR), by Zhang [23, 32, 34] based on higher order local approximation of the solution. There has been extensive work on the analysis of these techniques on uniform meshes, structured and unstructured isotropic

meshes [15, 30], and recently on structured anisotropic meshes [32]. With the establishment of the superconvergence of $\|\nabla u_N - \nabla u_I\|$, it is now possible to establish the superconvergence of these local gradient recovery techniques on general adaptively refined anisotropic meshes. We will report a detailed analysis and numerical results for these techniques in a separate paper.

REFERENCES

- [1] Mark Ainsworth and J. Tinsley Oden, *A Posteriori Error Estimation in Finite Element Analysis*, Pure and Applied Mathematics (New York), Wiley-Interscience [John Wiley & Sons], New York, 2000. MR1885308 (2003b:65001)
- [2] Ivo Babuška and Theofanis Strouboulis, *The Finite Element Method and its Reliability*, Numerical Mathematics and Scientific Computation, The Clarendon Press Oxford University Press, New York, 2001. MR1857191 (2002k:65001)
- [3] Randolph E. Bank and Jinchao Xu, *Asymptotically exact a posteriori error estimators. I. Grids with superconvergence*, SIAM J. Numer. Anal. **41** (2003), no. 6, 2294–2312 (electronic), DOI 10.1137/S003614290139874X. MR2034616 (2004k:65194)
- [4] Randolph E. Bank and Jinchao Xu, *Asymptotically exact a posteriori error estimators. II. General unstructured grids*, SIAM J. Numer. Anal. **41** (2003), no. 6, 2313–2332 (electronic), DOI 10.1137/S0036142901398751. MR2034617 (2004m:65212)
- [5] Weiming Cao, *On the error of linear interpolation and the orientation, aspect ratio, and internal angles of a triangle*, SIAM J. Numer. Anal. **43** (2005), no. 1, 19–40 (electronic), DOI 10.1137/S0036142903433492. MR2177954 (2006k:65023)
- [6] Weiming Cao, *An interpolation error estimate in $\mathcal{R}^2$ based on the anisotropic measures of higher order derivatives*, Math. Comp. **77** (2008), no. 261, 265–286 (electronic), DOI 10.1090/S0025-5718-07-01981-3. MR2353953 (2008j:65218)
- [7] Weiming Cao, *An interpolation error estimate on anisotropic meshes in $\mathcal{R}^n$ and optimal metrics for mesh refinement*, SIAM J. Numer. Anal. **45** (2007), no. 6, 2368–2391 (electronic), DOI 10.1137/060667992. MR2361894 (2008k:65264)
- [8] Long Chen, Pengtao Sun, and Jinchao Xu, *Optimal anisotropic meshes for minimizing interpolation errors in L^p -norm*, Math. Comp. **76** (2007), no. 257, 179–204, DOI 10.1090/S0025-5718-06-01896-5. MR2261017 (2008e:65385)
- [9] P. G. Ciarlet, *The Finite Element Methods for Elliptic Problems*, SIAM Classics in Applied Mathematics, SIAM, Philadelphia, 2002.
- [10] L. Formaggia and S. Perotto, *New anisotropic a priori error estimates*, Numer. Math. **89** (2001), no. 4, 641–667, DOI 10.1007/s002110100273. MR1865506 (2002j:65110)
- [11] L. Formaggia and S. Perotto, *Anisotropic error estimates for elliptic problems*, Numer. Math. **94** (2003), no. 1, 67–92, DOI 10.1007/s00211-002-0415-z. MR1971213 (2004c:65127)
- [12] Wagdi G. Habashi, Julien Dompierre, Yves Bourgault, Djaffar Ait-Ali-Yahia, Michel Fortin, and Marie-Gabrielle Vallet, *Anisotropic mesh adaptation: towards user-independent, mesh-independent and solver-independent CFD. I. General principles*, Internat. J. Numer. Methods Fluids **32** (2000), no. 6, 725–744, DOI 10.1002/(SICI)1097-0363(20000330)32:6<725::AID-FLD935>3.0.CO;2-4. MR1911901 (2003d:76106)
- [13] F. Hecht, Bidimensional anisotropic mesh generator (*User's Manual*), INRIA, Rocquencourt, 1997. <http://www-rocq.inria.fr/gamma/cdrom/www/bamg/eng.htm>
- [14] Gerd Kunert and Serge Nicaise, *Zienkiewicz-Zhu error estimators on anisotropic tetrahedral and triangular finite element meshes*, M2AN Math. Model. Numer. Anal. **37** (2003), no. 6, 1013–1043, DOI 10.1051/m2an:2003065. MR2026406 (2004k:65196)
- [15] A. M. Lakhany, I. Marek, and J. R. Whiteman, *Superconvergence results on mildly structured triangulations*, Comput. Methods Appl. Mech. Engrg. **189** (2000), no. 1, 1–75, DOI 10.1016/S0045-7825(99)00281-9. MR1779678 (2001f:74041)
- [16] Jichun Li and Mary F. Wheeler, *Uniform convergence and superconvergence of mixed finite element methods on anisotropically refined grids*, SIAM J. Numer. Anal. **38** (2000), no. 3, 770–798, DOI 10.1137/S0036142999351212. MR1781203 (2001f:65137)

- [17] Qun Lin and Jia-Fu Lin, *Superconvergence on anisotropic meshes*, Int. J. Inf. Syst. Sci. **3** (2007), no. 2, 261–266. MR2297053 (2008c:65337)
- [18] Shipeng Mao, Shaochun Chen, and Dongyang Shi, *Convergence and superconvergence of a nonconforming finite element on anisotropic meshes*, Int. J. Numer. Anal. Model. **4** (2007), no. 1, 16–38. MR2289730 (2007k:65188)
- [19] F. Marcuzzi, M. Morandi Cecchi, and M. Venturin, *An anisotropic unstructured triangular adaptive mesh algorithm based on error and error gradient information*, Math. Comput. Simulation **78** (2008), no. 5-6, 645–652, DOI 10.1016/j.matcom.2008.04.006. MR2435592 (2009e:65197)
- [20] Stefano Micheletti and Simona Perotto, *Reliability and efficiency of an anisotropic Zienkiewicz-Zhu error estimator*, Comput. Methods Appl. Mech. Engrg. **195** (2006), no. 9-12, 799–835, DOI 10.1016/j.cma.2005.02.009. MR2195291 (2006i:65179)
- [21] Jean-Marie Mirebeau, *Optimal meshes for finite elements of arbitrary order*, Constr. Approx. **32** (2010), no. 2, 339–383, DOI 10.1007/s00365-010-9090-y. MR2677884 (2011g:65278)
- [22] J.-M. Mirebeau, *Approximation adaptative et anisotrope par éléments finis: Théorie et Algorithmes*, Ph.D. dissertation, Laboratoire Jacques-Louis Lions, Université Pierre et Marie Curie, 2010.
- [23] Ahmed Naga and Zhimin Zhang, *A posteriori error estimates based on the polynomial preserving recovery*, SIAM J. Numer. Anal. **42** (2004), no. 4, 1780–1800 (electronic), DOI 10.1137/S0036142903413002. MR2114301 (2005k:65231)
- [24] M. Picasso, *Numerical study of the effectivity index for an anisotropic error indicator based on Zienkiewicz-Zhu error estimator*, Comm. Numer. Methods Engrg. **19** (2003), no. 1, 13–23, DOI 10.1002/cnm.546. MR1952014 (2004c:65131)
- [25] M. Picasso, *An anisotropic error indicator based on Zienkiewicz-Zhu error estimator: application to elliptic and parabolic problems*, SIAM J. Sci. Comput. **24** (2003), no. 4, 1328–1355 (electronic), DOI 10.1137/S1064827501398578. MR1976219 (2004e:65124)
- [26] R. Rannacher and S. Turek, *Simple nonconforming quadrilateral Stokes element*, Numer. Methods Partial Differential Equations **8** (1992), no. 2, 97–111, DOI 10.1002/num.1690080202. MR1148797 (92i:65170)
- [27] Dong-yang Shi, Shi-peng Mao, and Shao-chun Chen, *An anisotropic nonconforming finite element with some superconvergence results*, J. Comput. Math. **23** (2005), no. 3, 261–274. MR2133919 (2006h:65194)
- [28] Lars B. Wahlbin, *Superconvergence in Galerkin Finite Element Methods*, Lecture Notes in Mathematics, vol. 1605, Springer-Verlag, Berlin, 1995. MR1439050 (98j:65083)
- [29] Haijun Wu and Zhimin Zhang, *Can we have superconvergent gradient recovery under adaptive meshes?*, SIAM J. Numer. Anal. **45** (2007), no. 4, 1701–1722, DOI 10.1137/060661430. MR2338406 (2008k:65252)
- [30] Jinchao Xu and Zhimin Zhang, *Analysis of recovery type a posteriori error estimators for mildly structured grids*, Math. Comp. **73** (2004), no. 247, 1139–1152 (electronic), DOI 10.1090/S0025-5718-03-01600-4. MR2047081 (2005f:65141)
- [31] Ningning Yan and Aihui Zhou, *Gradient recovery type a posteriori error estimates for finite element approximations on irregular meshes*, Comput. Methods Appl. Mech. Engrg. **190** (2001), no. 32-33, 4289–4299, DOI 10.1016/S0045-7825(00)00319-4. MR1832657 (2002c:65189)
- [32] Zhimin Zhang, *Polynomial preserving recovery for meshes from Delaunay triangulation or with high aspect ratio*, Numer. Methods Partial Differential Equations **24** (2008), no. 3, 960–971, DOI 10.1002/num.20300. MR2402584 (2009c:65327)
- [33] Zhimin Zhang, *Polynomial preserving gradient recovery and a posteriori estimate for bilinear element on irregular quadrilaterals*, Int. J. Numer. Anal. Model. **1** (2004), no. 1, 1–24. MR2052728 (2005b:65119)
- [34] Zhimin Zhang and Ahmed Naga, *A new finite element gradient recovery method: superconvergence property*, SIAM J. Sci. Comput. **26** (2005), no. 4, 1192–1213 (electronic), DOI 10.1137/S1064827503402837. MR2143481 (2006d:65137)
- [35] Q.D. Zhu, *Theory on High Accuracy Post-processing in the Finite Element Method*, Science Press, Beijing, 2008 (in Chinese).

- [36] O. C. Zienkiewicz and J. Z. Zhu, *The superconvergent patch recovery and a posteriori error estimates. I. The recovery technique*, Internat. J. Numer. Methods Engrg. **33** (1992), no. 7, 1331–1364, DOI 10.1002/nme.1620330702. MR1161557 (93c:73098)
- [37] O. C. Zienkiewicz and J. Z. Zhu, *The superconvergent patch recovery and a posteriori error estimates. II. Error estimates and adaptivity*, Internat. J. Numer. Methods Engrg. **33** (1992), no. 7, 1365–1382, DOI 10.1002/nme.1620330703. MR1161558 (93c:73099)

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF TEXAS AT SAN ANTONIO, SAN ANTONIO,
TEXAS 78249

E-mail address: `weiming.cao@utsa.edu`