MATHEMATICS OF COMPUTATION
Volume 77, Number 261, January 2008, Pages 265–286
S 0025-5718(07)01981-3
Article electronically published on April 17, 2007

AN INTERPOLATION ERROR ESTIMATE IN $\mathcal{R}^2$ BASED ON THE ANISOTROPIC MEASURES OF HIGHER ORDER DERIVATIVES

WEIMING CAO

ABSTRACT. In this paper, we introduce the magnitude, orientation, and anisotropic ratio for the higher order derivative $\nabla^{k+1}u$ (with $k \geq 1$) of a function u to characterize its anisotropic behavior. The magnitude is equivalent to its usual Euclidean norm. The orientation is the direction along which the absolute value of the $k+1$-th directional derivative is about the smallest, while along its perpendicular direction it is about the largest. The anisotropic ratio measures the strength of the anisotropic behavior of $\nabla^{k+1}u$. These quantities are invariant under translation and rotation of the independent variables. They correspond to the area, orientation, and aspect ratio for triangular elements. Based on these measures, we derive an anisotropic error estimate for the piecewise polynomial interpolation over a family of triangulations that are quasi-uniform under a given Riemannian metric. Among the meshes of a fixed number of elements it is identified that the interpolation error is nearly the minimum on the one in which all the elements are aligned with the orientation of $\nabla^{k+1}u$, their aspect ratios are about the anisotropic ratio of $\nabla^{k+1}u$, and their areas make the error evenly distributed over every element.

1. Introduction

Numerous examples have shown that long and thin elements are useful in computation of problems with boundary or internal layers [1, 2, 14, 15, 19, 22]. A practical question is in what direction an element should be long and how long and thin it should be. More generally, given a fixed number of degree of freedom, what are the characteristics of the optimal or nearly optimal mesh that produces the smallest approximation error? Here we confine ourselves to the approximation problem of interpolation by piecewise polynomials. For linear interpolation, the answer to the above question has been made clear through a number of works over the past 20 years, [18, 19, 13, 21, 8]. The main conclusion is: given the area of a general triangular element τ, the error (in various norms) for the linear interpolation of a function u at the vertices of τ is nearly the minimum when τ is aligned with the eigenvector (associated with the smaller eigenvalue) of the Hessian $\nabla^2 u$, and the aspect ratio (or stretch ratio) of τ is about the square root of the ratio of the larger eigenvalue of $\nabla^2 u$ to the smaller one. The globally optimal or nearly optimal mesh

Received by the editor June 24, 2005 and, in revised form, October 11, 2006.
2000 *Mathematics Subject Classification.* Primary 65D05, 65L50, 65N15, 65N50.
Key words and phrases. Interpolation error, anisotropic mesh, anisotropic measure, aspect ratio, mesh alignment, mesh metric.
This work was supported in part by NSF grant DMS-0209313.

can be further characterized by the equidistribution of the interpolation error over each element [17, 10, 11].

In the case of piecewise interpolation by polynomials of degree $k \geq 2$, the conclusion is far from clear. There are only a few papers considering the anisotropic error estimates and mesh refinement for higher order elements. For instance, denote by $\Pi_k u$ the interpolation of u by polynomials of degree k. Apel derived in [2] the following estimate for the interpolation error $u - \Pi_k u$ over an anisotropic element τ:

$$|u - \Pi_k u|_{W^{m,q}(\tau)} \leq c|\tau|^{1/q-1/p} \sum_{i+j=k-m+1} h_1^i h_2^j |\partial_{x^i y^j} u|_{W^{m,p}(\tau)},$$

where $W^{m,p}$ is the usual Sobolev space of functions whose up to m-th order derivatives are L^p-integrable. h_1 and h_2 are the lengths of τ along x and y directions, respectively. This estimate indicates qualitatively that when the partial derivatives of u are of different magnitudes in different directions, an element can be long and thin in the direction of smaller partial derivatives without compromising the overall accuracy of interpolation. The difficulty in using this estimate for adaptive mesh refinement is that it does not specify in what direction the partials are considered small and thus how the element should be aligned, nor does it specify how much the element aspect ratio should be. For example, if u is a function of $x + y$ only, there is no hint in this error estimate indicating that the element should be long and thin along the constant u direction $(1, -1)^T$. Analogously to the analysis of linear interpolation errors, Huang [16] provided an estimate for the $W^{m,p}$-norm of $u - \Pi_k u$ in terms of the eigenvalues and eigenvectors of the follow matrix:

$$H(D^{k-1}u) = \sum_{i+j=k-1} \mathrm{abs}(\nabla^2(\partial_{x^i y^j} u)),$$

where $\mathrm{abs}(A) = \sqrt{A^T A}$ for a real matrix A. It is seen from his estimate that the optimal triangle should be aligned with the eigenvector (associated with the smaller eigenvalue) of H, and its aspect ratio should equal to the square root of the ratio of the larger eigenvalue of H to the smaller one. These conclusions can be readily applied to anisotropic mesh generation and refinement. However, since condensing $\nabla^{k+1} u$ to H inevitably loses some information about its anisotropic behavior, this error estimate may not be accurate, and the direction and aspect ratio based on H may be far from optimal. For instance, let ϵ be a small positive number. Consider the interpolation of $u = (\epsilon x)^{k+1} + y^{k+1}$ on a triangle by polynomials of degree k. The best aspect ratio for the L^p-error is about ϵ^{-1}, which transforms the problem into interpolating $\hat{u} = \epsilon^{(k+1)/2}(\hat{x}^{k+1} + \hat{y}^{k+1})$ on a shape regular element of the same area. But the aspect ratio predicted by the eigenvalues of H would be $\epsilon^{-(k+1)/2}$, with which the L^p-norm of the interpolation error would be about the same magnitude as that of interpolating $\hat{u} = \epsilon^{(3-k)/4} \hat{x}^{k+1}$ on a shape regular element of the same area.

Needless to say the error for interpolation by polynomials of degree k depends on the $k + 1$-th derivatives of the interpolated function u. Accurately understanding the anisotropic behavior of $\nabla^{k+1} u$ is crucial for the anisotropic error estimates and mesh refinements. As pointed out in [2], page 7, a tight error estimate for $k \geq 2$ relies on a "sufficiently fine description of the properties of the solution u". A major motivation for our work in this paper is to find a way to measure quantitatively the anisotropic behavior of $\nabla^{k+1} u$. More precisely, we introduce for any $k \geq 1$

the magnitude, orientation, and anisotropic ratio of $\nabla^{k+1}u$. The magnitude is equivalent to the usual Euclidean norm of $\nabla^{k+1}u$. The orientation of $\nabla^{k+1}u$ is the direction along which the absolute value of the $k+1$-th directional derivative is about the smallest, while along its perpendicular direction is about the largest. The anisotropic ratio measures the strength of the anisotropic behavior of $\nabla^{k+1}u$. A critical feature for these definitions is that they are invariant under translation and rotation of the xy-coordinates. These quantities correspond to the three geometric features of an anisotropic triangular element: the size, the orientation, and the aspect ratio. One may determine the size, orientation and aspect ratio of the triangular elements according to the magnitude, orientation, and anisotropic ratio of $\nabla^{k+1}u$ in mesh generation and refinement.

Another motivation of this work is to find the connection among the interpolation error, the geometric features of the triangular elements, and the anisotropic features of $\nabla^{k+1}u$. We derive an error estimate for interpolation over a family of triangulations that are quasi-uniform under a given Riemannian metric M. The estimate is formulated in terms of the magnitude, orientation, and anisotropic ratio of $\nabla^{k+1}u$ and the metric M. Based on this estimate we identify an optimal metric which leads to the smallest error bound for the $W^{m,p}$-seminorm of the interpolation error. When a triangulation is quasi-uniform under the optimal metric, all elements are aligned with the orientation of $\nabla^{k+1}u$, their aspect ratios approximately equal to the anisotropic ratio of $\nabla^{k+1}u$, and the error over each element is about evenly distributed. In this case, the total interpolation error can be bounded by [1]

$$\Big(\sum_{\tau\in\mathcal{T}_N}|u-\Pi_k u|^p_{m,p,\tau}\Big)^{1/p}\le cN^{-(k+1-m)/2}\|(S_{k+1})^{-(k+1-m)/2}D_{k+1}\|_{L^{2/(k+1-m+2/p)}(\Omega)},$$

where N is the total number of elements, D_{k+1} and S_{k+1} are the magnitude and anisotropic ratio of $\nabla^{k+1}u$, respectively. This error bound extends the optimal interpolation error estimates for linear elements in [2, 10, 11, 16] to higher order elements in $\mathcal{R}^2$. The above conclusions also agree with those based on the exact error formulas in the model problems of linear interpolation of a quadratic function ($k=1$) and quadratic interpolation of a cubic function ($k=2$) presented in [8] and [9], respectively.

An outline of this paper is as follows: in Section 2 we introduce the magnitude, orientation, and anisotropic ratio for $\nabla^{k+1}u$. In Section 3 we derive the error estimate for interpolation by piecewise polynomials of degree k over triangulations that are quasi-uniform under a given Riemannian metric M. Optimal metrics which lead to the smallest error bound are identified. We give in Section 4 an example to support the optimality of the metrics predicted by our error estimate. Section 5 contains some discussions.

Throughout the paper, we use c as a generic constant which is independent of the mesh and the functions involved. It may take different values in different places.

2. Anisotropic measures of higher order derivatives

Second order derivative $\nabla^2 u$. It is well-known that the anisotropic features of the Hessian matrix $\nabla^2 u$ can be characterized by its eigenvalues and eigenvectors.

[1]Throughout this paper, we use for $\alpha>0$ the notation $\|v\|_{L^{\alpha)}}=[\int|v|^\alpha]^{1/\alpha}$ to denote a non-negative functional on L^α of functions whose α-th power is Lebesgue integrable. It is the usuall L^α norm when $\alpha\ge 1$.

For fixed $\boldsymbol{x}$, let

$$\nabla^2 u(\boldsymbol{x}) = R_{\phi_2} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} R_{\phi_2}^T, \tag{1}$$

where $|\lambda_1| \le |\lambda_2|$ are the eigenvalues of $\nabla^2 u(\boldsymbol{x})$, and R_{ϕ_2} is the matrix of rotation by angle ϕ_2 counter-clockwise. We define the orientation of $\nabla^2 u(\boldsymbol{x})$ to be the direction of the eigenvector associated with the smaller eigenvalue of $\nabla^2 u$, i.e., the direction of angle ϕ_2 from the x-axis. We also define the anisotropic ratio of $\nabla^2 u$ to be

$$\mathcal{S}_2 = \sqrt{|\frac{\lambda_2}{\lambda_1}|},$$

and the magnitude of $\nabla^2 u$ as

$$\mathcal{D}_2 = \frac{1}{2}(|\lambda_1| + |\lambda_2|).$$

Let $\boldsymbol{\xi} = (\xi, \eta)^T$; we introduce a homogenuous polynomial of $\boldsymbol{\xi}$ as

$$p_2(\boldsymbol{\xi}) = \frac{1}{2!}(\boldsymbol{\xi} \cdot \nabla)^2 u(\boldsymbol{x}).$$

When $\|\boldsymbol{\xi}\| = 1$, $p_2(\boldsymbol{\xi})$ is the (scaled) second order directional derivative of $u(\boldsymbol{x})$ along $\boldsymbol{\xi}$. The above defined anisotropic measures of $\nabla^2 u(\boldsymbol{x})$ can be determined equivalently in terms of the level curves of p_2. Indeed, when $\lambda_1\lambda_2 \le 0$, p_2 is the product of two linear functions of $\boldsymbol{\xi}$. It is not difficult to verify that

$$p_2(\boldsymbol{\xi}) = \mathcal{D}_2 \ \ell_1(\boldsymbol{\xi}) \ \ell_2(\boldsymbol{\xi}), \tag{2}$$

where

$$\ell_i(\boldsymbol{\xi}) = \xi \sin\beta_i - \eta\cos\beta_i, \qquad i = 1, 2, \tag{3}$$

with β_1 and β_2 being the angles of the two level-0 lines of p_2 from the ξ-axis. Moreover, these two lines divide the $\xi\eta$-plane into four sectors. The orientation ϕ_2 of $\nabla^2 u(\boldsymbol{x})$ is the bisector of the two smaller sectors. Also, the anisotropic ratio $\mathcal{S}_2 = \tan(\alpha)$, where 2α is just the opening the smaller sectors. See [9] for more details.

When $\lambda_1\lambda_2 > 0$, $p_2(\boldsymbol{\xi})$ is always positive or negative, and the level curves of p_2 are concentric ellipses. In this case, we define an associated polynomial $p_2^*(\boldsymbol{\xi}) = \frac{1}{2!}\boldsymbol{\xi} \cdot \hat{H}\boldsymbol{\xi}$ with

$$\hat{H} = R_{\phi_2} \begin{bmatrix} \lambda_1 & 0 \\ 0 & -\lambda_2 \end{bmatrix} R_{\phi_2}^T. \tag{4}$$

Then p_2^* can be factored into two linear functions as in (2), and the magnitude, orientation, and anisotropic ratio of $\nabla^2 u$ can be determined equivalently in terms of the level-0 lines of $p_2^*(\boldsymbol{\xi})$.

Higher order derivative $\nabla^m u$. For $m \ge 3$, $\nabla^m u(\boldsymbol{x})$ is no longer a matrix. To characterize its anisotropic behavior, we study the following homogeneous polynomial for the (scaled) m-th order directional derivatives at $\boldsymbol{x}$:

$$p_m(\boldsymbol{\xi}) = \frac{1}{m!}(\boldsymbol{\xi} \cdot \nabla)^m u(\boldsymbol{x}). \tag{5}$$

By the fundamental theorem of algebra, p_m can be factored as

$$p_m(\boldsymbol{\xi}) = \mathcal{D}_m \ \ell_1 \ \ell_2 \cdots \ell_{m_1} \cdot q_1 \ q_2 \cdots q_{m_2}, \tag{6}$$

where $\mathcal{D}_m$ is a non-negative number depending on $\boldsymbol{x}$, and $m_1 + 2m_2 = m$. Each ℓ_i is a linear function in the form of (3), and each q_i is a quadratic function in the form $q_i(\boldsymbol{x}) = \boldsymbol{\xi} \cdot H_i \boldsymbol{\xi}$, where the matrix

$$H_i = R_{\theta_i} \begin{bmatrix} \lambda_1^{(i)} & 0 \\ 0 & \lambda_2^{(i)} \end{bmatrix} R_{\theta_i}^T \tag{7}$$

with $\lambda_1^{(i)}\lambda_2^{(i)} > 0$, $|\lambda_1^{(i)}| \le |\lambda_2^{(i)}|$, and $|\lambda_1^{(i)} + \lambda_2^{(i)}| = 1$. This decomposition is unique up to the signs of ℓ_i and q_i. For $i = 1, 2, \cdots, m_2$, let $\hat{H}_i$ be defined by H_i in the same way as $\hat{H}$ is by $\nabla^2 u$ in (4), and let $q_i^*(\boldsymbol{\xi}) = \boldsymbol{\xi} \cdot \hat{H}_i \boldsymbol{\xi}$. Then we define a polynomial $p_m^*(\boldsymbol{\xi})$ as follows:

$$p_m^*(\boldsymbol{\xi}) = \mathcal{D}_m\ \ell_1\ \ell_2 \cdots \ell_{m_1} \cdot q_1^*\ q_2^* \cdots q_{m_2}^*. \tag{8}$$

Now each q_i^* can be factored as $q_i^*(\boldsymbol{\xi}) = \ell_{i'}(\boldsymbol{\xi})\ell_{i''}(\boldsymbol{\xi})$ with $\ell_{i'}$ and $\ell_{i''}$ in the form of (3). Thus p_m^* can be expressed as

$$p_m^*(\boldsymbol{\xi}) = \mathcal{D}_m \ell_1\ \ell_2 \cdots \ell_m. \tag{9}$$

We define the magnitude of $\nabla^m u$ as the coefficient $\mathcal{D}_m$ in (6). It is an upper bound for all the m-th order directional derivatives of u at $\boldsymbol{x}$. It is also equivalent to $\sum\limits_{i+j=m} |\partial_{x^i y^j} u(\boldsymbol{x})|$; see Lemma 2.2 below. We define the orientation and the anisotropic ratio of $\nabla^m u$ in terms of the directions of the level-0 lines of p_m^*. Specifically, assume without loss of generality that

$$-\frac{\pi}{2} < \beta_1 \le \beta_2 \le \cdots \le \beta_m \le \frac{\pi}{2}.$$

Consider the ranges for the central angle of the m sectors:

$$[\beta_1, \beta_m], \qquad \text{and} \qquad [\beta_j, \pi + \beta_{j-1}], \quad j = 2, 3, \cdots, m.$$

Let $[\underline{\beta}, \overline{\beta}]$ be the shortest interval among all of them; see Figure 1. Clearly $\overline{\beta} - \underline{\beta} \le (1 - 1/m)\pi$. If $\overline{\beta} - \underline{\beta} << 1$, then the absolute value of the m-th directional derivative along the line of angle

$$\phi_m = \frac{1}{2}(\underline{\beta} + \overline{\beta})$$

from the x-axis is about the smallest, while along its perpendicular direction it is about the largest. We define the direction of angle ϕ_m from the x-axis as the orientation of $\nabla^m u$. We may also use $\cot(\frac{1}{2}(\overline{\beta} - \underline{\beta}))$ to define the anisotropic ratio of $\nabla^m u$. When $\bar{\beta} - \underline{\beta} \ge \pi/2$, the level-0 lines of $p_m^*(\boldsymbol{\xi})$ cannot be wrapped in two opposite sectors of an acute angle. In this case, there is no obviously preferred direction for general $\nabla^m u$, and we consider its anisotropic ratio equal to 1. Therefore, we define the anisotropic ratio of $\nabla^m u$ as

$$\mathcal{S}_m = \max(1, \cot(\frac{1}{2}(\overline{\beta} - \underline{\beta}))).$$

Clearly, the larger $\mathcal{S}_m$ is, the stronger the anisotropic behavior of $\nabla^m u(\boldsymbol{x})$.

Lemma 2.1. *Let $p_m(\boldsymbol{\xi})$ be the polynomial defined in* (5)*, and let $\mathcal{D}_m$ be the magnitude of $\nabla^m u$. Then*

$$(\sin\frac{\pi}{2m})^m \mathcal{D}_m \le \sup_{\|\boldsymbol{\xi}\|=1} |p_m(\boldsymbol{\xi})| \le \mathcal{D}_m.$$

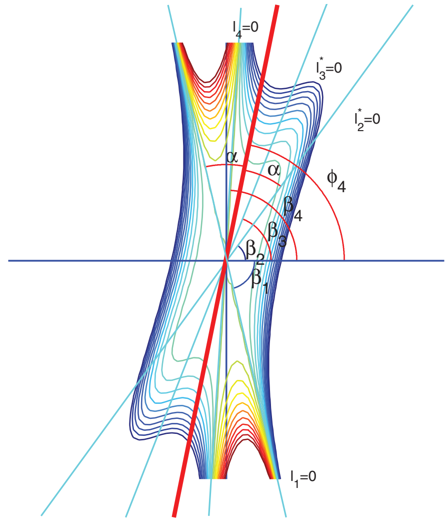

FIGURE 1. Contour plot of $p_4(\boldsymbol{\xi})$ for some $\nabla^4 u$ at a point. Here $p_4(\boldsymbol{\xi})$ is the product of two linear factors $\ell_1(\boldsymbol{\xi})$ and $\ell_4(\boldsymbol{\xi})$, and a quadratic factor $q_1(\boldsymbol{\xi})$, whose associated polynomial is $q_1^*(\boldsymbol{\xi}) = \ell_2(\boldsymbol{\xi})\ell_3(\boldsymbol{\xi})$. $[\underline{\beta}, \bar{\beta}] = [\beta_2, \beta_1 + \pi]$. The bold line with angle ϕ_4 is the orientation of $\nabla^4 u$, and the anisotropic ratio $\mathcal{S}_4 = \cot\alpha$.

Proof. Note that for $\|\boldsymbol{\xi}\| = 1$, $|\ell_i(\boldsymbol{\xi})| \le 1$, and $|q_i(\boldsymbol{\xi})| \le 1$. We have from (6) that

$$\sup_{\|\boldsymbol{\xi}\|=1} |p_m(\boldsymbol{\xi})| \le \mathcal{D}_m.$$

To show the left hand side inequality, we note that there always exists a direction $\boldsymbol{\xi}_0 = [\cos t_0, \sin t_0]^T$ which is at least a $\pi/(2m)$ angle away from all the level-0 lines of $p_m^*(\boldsymbol{\xi})$, i.e.,

$$\frac{\pi}{2m} \le |t_0 - \beta_i| \le \pi - \frac{\pi}{2m}, \qquad \forall i = 1, 2, \cdots, m.$$

Hence we have for all $1 \le i \le m$,

$$|\ell_i(\boldsymbol{\xi}_0)| = \sin(|t_0 - \beta_i|) \ge \sin(\frac{\pi}{2m}).$$

Recall that for each quadratic factor q_i of p_m, there are two linear factors $\ell_{i'}(\boldsymbol{\xi})$ and $\ell_{i''}(\boldsymbol{\xi})$ of $p_m^*(\boldsymbol{\xi})$ such that $q_i^*(\boldsymbol{\xi}) = \ell_{i'}(\boldsymbol{\xi})\ell_{i''}(\boldsymbol{\xi})$. Thus

$$|q_i(\boldsymbol{\xi}_0)| \ge |q_i^*(\boldsymbol{\xi}_0)| = |\ell_{i'}(\boldsymbol{\xi}_0)| \cdot |\ell_{i''}(\boldsymbol{\xi}_0)| \ge \sin^2(\frac{\pi}{2m}).$$

Therefore we have

$$\sup_{|\boldsymbol{\xi}|=1} |p_m(\boldsymbol{\xi})| \ge (\sin\frac{\pi}{2m})^m \mathcal{D}_m.$$

□

Lemma 2.2. *Let $\mathcal{D}_m$ be the magnitude of $\nabla^m u$. Then there exist positive constants c_1 and c_2 depending on m only such that*

$$c_1 \mathcal{D}_m \le \sum_{i+j=m} |\partial_{x^i y^j} u| \le c_2 \mathcal{D}_m.$$

Proof. By the previous lemma, we need only to show that $\sum_{i+j=m} |\partial_{x^i y^j} u|$ is equivalent to $\sup_{\|\boldsymbol{\xi}\|=1} |p_m(\boldsymbol{\xi})|$. Note that

$$p_m(\boldsymbol{\xi}) = \frac{1}{m!} \sum_{i=0}^{m} \binom{m}{i} \frac{\partial^m u}{\partial x^i \partial y^{m-i}} \; \xi^i \eta^{m-i}; \tag{10}$$

we have

$$\sup_{\|\boldsymbol{\xi}\|=1} |p_m(\boldsymbol{\xi})| \leq \frac{2^m}{m!} \max_{0 \leq i \leq m} |\partial_{x^i y^{m-i}} u| \leq \frac{2^m}{m!} \sum_{i+j=m} |\partial_{x^i y^j} u|.$$

On the other hand, we see from (10) that

$$\partial_{x^i y^j} u = \partial_{\xi^i \eta^j} p_m(\boldsymbol{\xi}) = \mathcal{D}_m \partial_{\xi^i \eta^j} \left(\Pi_{k=1}^{m_1} \ell_k \; \Pi_{k=1}^{m_2} q_k \right).$$

Note that $\ell_k(\boldsymbol{\xi})$ and $q_k(\boldsymbol{\xi})$ are linear and quadratic functions of $\boldsymbol{\xi}$, respectively. For any $\|\boldsymbol{\xi}\| = 1$, we have that all $|\ell_k(\boldsymbol{\xi})|$, $|q_k(\boldsymbol{\xi})|$, and the derivative of ℓ_k, are no more than 1, and that all the first and second order partial derivatives of q_k are no more than 2. Because $\partial_{\xi^i \eta^j}[\Pi_{k=1}^{m_1} \ell_k \Pi_{k=1}^{m_2} q_k]$ is the sum of at most m^m terms of the products of them (with the derivatives of q_k's counted twice), we have

$$|\partial_{x^i y^j} u| \leq m^m \mathcal{D}_m,$$

which completes the proof of this lemma. □

Lemma 2.3. *Let $\mathcal{D}_m$, ϕ_m, and S_m be the magnitude, orientation angle, and anisotropic ratio of $\nabla^m u$, respectively. Then there exists a constant c depending only on m such that*

$$|p_m(\boldsymbol{\xi})| \;\leq\; c \; \mathcal{D}_m (\boldsymbol{\xi} \cdot Q_m \boldsymbol{\xi})^{m/2}, \qquad \forall \boldsymbol{\xi} \in \mathcal{R}^2, \tag{11}$$

where

$$Q_m = R_{\phi_m} \begin{bmatrix} \frac{1}{S_m^2} & 0 \\ 0 & 1 \end{bmatrix} R_{\phi_m}^T, \tag{12}$$

and R_{ϕ_m} is the matrix of rotation by angle ϕ_m counter-clockwise.

Proof. Because both sides of the above inequality are homogeneous polynomials of the same degree, we only have to prove the inequality for all $\|\boldsymbol{\xi}\| = 1$. Also, since $\mathcal{D}_m$ and $\mathcal{S}_m$ are invariant under the rotation of xy-coordinates, we only have to prove this lemma for functions with $\phi_m = 0$. Namely, if the orientation of $\nabla^m u$ is along the x-axis, then

$$|p_m(\boldsymbol{\xi})| \leq \; c \; \mathcal{D}_m (\xi^2/\mathcal{S}_m^2 + \eta^2)^{m/2}. \tag{13}$$

Recall the decomposition (6) of $p_m(\boldsymbol{\xi})$. If $\mathcal{S}_m = 1$, the above inequality is obviously true. When $\mathcal{S}_m > 1$, all the level-0 lines of $p_m^*(\boldsymbol{\xi})$ are within sector $[-\alpha, \alpha]$ with $\alpha = \arctan(1/\mathcal{S}_m) \in [0, \frac{\pi}{4})$. This implies that for each linear factor $\ell_i(\boldsymbol{\xi}) = \xi \sin \beta_i - \eta \cos \beta_i$ of $p_m(\boldsymbol{\xi})$ in (6) we have $|\beta_i| \leq \alpha$. To bound the linear factor ℓ_i, we study for $\boldsymbol{\xi} = [\cos t, \sin t]^T$ the ratio

$$G_1(t) \;\equiv\; \frac{[\ell_i(\boldsymbol{\xi})]^2}{\cos^2 t / \mathcal{S}_m^2 + \sin^2 t} \;=\; \cos^2 \beta_i \; \frac{(\tan \beta_i - \tan t)^2}{1/\mathcal{S}_m^2 + \tan^2 t}, \qquad \forall t \in [0, 2\pi].$$

The maximum value of $G_1(t)$ is attained when $\tan t = -1/(\mathcal{S}_m^2 \; \tan \beta_i)$ with

$$\max_{t \in [0, 2\pi]} G_1(t) \;=\; \cos^2 \beta_i \cdot (\mathcal{S}_m^2 \tan^2 \beta_i + 1) \;\leq\; 2,$$

where we have used $|\tan\beta_i| \le \tan\alpha = 1/\mathcal{S}_m$. Hence we conclude that

$$|\ell_i(\boldsymbol{\xi})| \;\le\; 2\;(\xi^2/\mathcal{S}_m^2+\eta^2)^{1/2}, \qquad \forall i = 1,2,\cdots,m_1.$$

For the quadratic factors $q_i(\boldsymbol{\xi})$, $i = 1,\cdots,m_2$, in (6), we have from (7) that

$$q_i(\boldsymbol{\xi}) \;=\; \sin^2\delta_i\,\cos^2(t-\theta_i) + \cos^2\delta_i\,\sin^2(t-\theta_i),$$

where $\delta_i = \arctan(\sqrt{\lambda_1^{(i)}/\lambda_2^{(i)}})$. Its associated polynomial is

$$q_i^*(\boldsymbol{\xi}) \;=\; \sin^2\delta_i\,\cos^2(t-\theta_i) - \cos^2\delta_i\,\sin^2(t-\theta_i) \;=\; -\sin(t-\theta_i-\delta_i)\,\sin(t-\theta_i+\delta_i).$$

By the definition of the anisotropic ratio, we have

$$|\theta_i \pm \delta_i| \le \alpha.$$

To bound the quadratic factor $q_i(\boldsymbol{\xi})$, we define for $\boldsymbol{\xi} = [\cos t, \sin t]^T$

$$\begin{aligned} G_2(t) &\equiv \frac{q_i(\boldsymbol{\xi})}{\cos^2 t/\mathcal{S}_m^2 + \sin^2 t} \\ &= \cos^2\alpha\;\frac{\sin^2\delta_i\,\cos^2(t-\theta_i)+\cos^2\delta_i\,\sin^2(t-\theta_i)}{\sin^2\alpha\,\cos^2 t+\cos^2\alpha\,\sin^2 t}, \quad \forall t\in[0,2\pi]. \end{aligned}$$

We study the supremum of G_2 for all $0 \le \alpha \le \pi/4$, $0 \le \delta_i \le \alpha$, $|\theta_i \pm \delta_i| \le \alpha$, and $0 \le t \le 2\pi$. Let's first fix $\alpha \in [0,\pi/4]$, $|\theta| \le \alpha$, and $t \in [0,2\pi]$, and look for the maximum of G_2 for $\delta_i \in [0, \alpha - |\theta_i|]$. Since G_2 is a linear function of $\sin^2\delta_i$, its maximum can be attained at either $\delta_i = 0$ or $\delta_i = \alpha - |\theta_i|$. For the case $\delta_i = 0$, we have

$$\begin{aligned} G_2(t) \;&=\; \cos^2\alpha\;\frac{\sin^2(t-\theta_i)}{\sin^2\alpha\cos^2 t+\cos^2\alpha\sin^2 t} \\ &\le\; 2\cos^2\alpha\;\frac{\sin^2\theta_i\cos^2 t+\cos^2\theta_i\sin^2 t}{\sin^2\alpha\cos^2 t+\cos^2\alpha\sin^2 t}. \end{aligned}$$

The right hand side of the above inequality is a rational function of $\sin^2 t$. Its maximum is attained at either $\sin^2 t = 0$ or $\sin^2 t = 1$. In both cases, we have $G_2(t) \le 2$ by the fact $|\theta_i| \le \alpha$.

In the case $\delta_i = \alpha - |\theta_i|$, we have

$$\begin{aligned} G_2(t) \;\le\; \cos^2\alpha \Bigg(&\frac{\sin^2(t-\theta_i)}{\sin^2\alpha\cos^2 t+\cos^2\alpha\sin^2 t} \\ &+\frac{2\sin^2(\alpha-|\theta_i|)(\cos^2\theta_i\cos^2 t+\sin^2\theta_i\sin^2 t)}{\sin^2\alpha\cos^2 t+\cos^2\alpha\sin^2 t}\Bigg). \end{aligned}$$

The first term in the right hand side of the above inequality can be bounded by 2 as shown before. The second term is again a rational function of $\sin^2 t$, whose maximum is less than or equal to 2 by the fact $|\theta_i| \le \alpha \le \pi/4$.

Combining the above two cases, we conclude that for each quadratic factor q_i,

$$|q_i(\boldsymbol{\xi})| \;\le\; 4\;(\xi^2/S_m^2+\eta^2), \qquad \forall\boldsymbol{\xi}\in\mathcal{R}^2,$$

which completes the proof of the lemma. □

Remark 2.1. Lemma 2.3 is the key to characterizing the anisotropic behavior of higher order derivatives. Note that the right hand side of (11) is the $\frac{m}{2}$-th power of a quadratic function of $\boldsymbol{\xi}$, whose level curves are concentric ellipses. We may alternatively define the magnitude, orientation, and anisotropic ratio of $\nabla^m u$ by the size, orientation, and eccentricity of the largest possible ellipse enclosed in the curve $|p_m(\boldsymbol{\xi})| = 1$. This idea can be generalized to define the anisotropic measures of higher order derivatives in three dimensions.

Remark 2.2. Lemmas 2.1 and 2.2 indicate that the magnitude of $\nabla^m u$ is equivalent to its largest m-th order directional derivative. In particular, all the mixed m-th order partial derivatives $\partial_{x^i y^{m-i}} u(\boldsymbol{x})$ can be bounded by the largest m-th order directional derivative of u at $\boldsymbol{x}$.

3. Interpolation error estimates

We first recall some classical results for the interpolation error estimates. Let $\{\mathcal{T}_N\}$ be a family of triangulations for a given polygonal domain Ω, where N represents the total number of elements. $\{\mathcal{T}_N\}$ is called regular if each element is shape regular, i.e., $\forall \tau \in \mathcal{T}_N$, $diam(\tau) \leq c|\tau|^{1/2}$, or equivalently, the minimum internal angles of every $\tau \in \mathcal{T}_N$ is bounded from below by a positive constant. Let k be a positive integer. Denote by $\mathcal{P}_k$ the set of polynomials of $\boldsymbol{x}$ of total degree $\leq k$. Let Π_k be an interpolation operator which preserves $\mathcal{P}_k$ on each element. It is well-known (see, e.g., Thm 3.1.5 of [12]) that for $0 \leq m \leq k$ and $p, q \in [1, \infty]$,

$$|u - \Pi_k u|_{m,p,\tau} \leq \ c \ |\tau|^{(k+1-m)/2+1/p-1/q} |u|_{k+1,q,\tau}, \tag{14}$$

provided that

$$W^{k+1,q}(\tau) \hookrightarrow C^s(\tau), \qquad \text{and} \qquad W^{k+1,q}(\tau) \hookrightarrow W^{m,p}(\tau), \tag{15}$$

where s is the greatest order of the partial derivatives occuring in the definition of Π_k.

If we further assume that $\{\mathcal{T}_N\}$ is quasi-uniform, i.e., all $\tau \in \mathcal{T}_N$ are shape regular and

$$\max_{\tau \in \mathcal{T}_N} |\tau| \leq c \min_{\tau \in \mathcal{T}_N} |\tau|,$$

then we have globally

$$\Big(\sum_{\forall \tau \in \mathcal{T}_N} |u - \Pi_k u|^p_{m,p,\tau} \Big)^{1/p} \leq c N^{-(k+1-m)/2} |u|_{k+1,p,\Omega}. \tag{16}$$

To derive the anisotropic error estimates, we introduce a Riemannian metric M on Ω. We extend the above classical results to the case where the triangulation $\{\mathcal{T}_N\}$ is quasi-uniform under metric M. Without loss of generality, we assume

$$M(\boldsymbol{x}) = \mu \ R_\psi \begin{bmatrix} \frac{1}{r} & 0 \\ 0 & r \end{bmatrix} R_\psi^T, \tag{17}$$

where $\mu > 0, r \geq 1$, and ψ are smooth functions of $\boldsymbol{x}$, and R_ψ is the matrix of rotation by angle ψ. For each element τ in the triangulation, we define $\bar{\mu}$ to be the mean value of μ on τ, i.e.,

$$\bar{\mu} = \frac{1}{|\tau|} \int_\tau \mu(\boldsymbol{y}) \, d\boldsymbol{y}.$$

Similarly $\bar{\psi}$ and $\bar{r}$ denote the mean values of ψ and r, respectively. Let $\boldsymbol{x}_c$ be the center of τ. Define an affine transform $\tilde{\boldsymbol{x}} = F_\tau^{-1}(\boldsymbol{x} - \boldsymbol{x}_c)$ with

$$F_\tau = \bar{\mu}^{-1/2} R_{\bar{\psi}} \begin{bmatrix} \sqrt{\bar{r}} & 0 \\ 0 & \frac{1}{\sqrt{\bar{r}}} \end{bmatrix}.$$

We call a family of triangulations $\{\mathcal{T}_N\}$ quasi-uniform under metric M, if

(i) $\forall \tau \in \mathcal{T}_N$, $\tilde{\tau} = F_\tau^{-1}(\tau - \boldsymbol{x}_c)$ is shape regular; and

(ii) $\max_{\tau\in\mathcal{T}_N} |\tilde{\tau}| \le c \min_{\tau\in\mathcal{T}_N} |\tilde{\tau}|$.

We first list a lemma about the magnitude of the higher order derivatives in the transformed coordinate $\tilde{\boldsymbol{x}}$. Define $\tilde{u}(\tilde{\boldsymbol{x}}) = u(\boldsymbol{x}_c + F_\tau \tilde{\boldsymbol{x}})$ and denote by $\tilde{\nabla}$ the gradient operator with respect to $\tilde{\boldsymbol{x}}$. Clearly

$$\tilde{\nabla} \;=\; (\frac{\partial \boldsymbol{x}}{\partial \tilde{\boldsymbol{x}}})^T \;\nabla \;=\; F_\tau^T \;\nabla.$$

Lemma 3.1. *For each element $\tau \in \mathcal{T}_N$, denote by $\mathcal{D}_m$, ϕ_m, and $\mathcal{S}_m$ the magnitude, orientation angle, and anisotropic ratio of $\nabla^m u$ at a point $\boldsymbol{x} \in \tau$, respectively. Then we have for the magnitude $\tilde{\mathcal{D}}_m$ of $\tilde{\nabla}^m \tilde{u}$ at $\tilde{\boldsymbol{x}} = F_\tau^{-1}(\boldsymbol{x} - \boldsymbol{x}_c)$ that*

$$\tilde{\mathcal{D}}_m \le c\,\mathcal{D}_m\,(\bar{\mu} S_m)^{-m/2}\{\cos^2(\phi_m - \bar{\psi})\;(\frac{\bar{r}}{S_m} + \frac{S_m}{\bar{r}}) + \sin^2(\phi_m - \bar{\psi})\;(\bar{r} S_m)\}^{m/2}.$$

Proof. Let

$$(18)\quad p_m(\boldsymbol{\xi}) = \frac{1}{m!}(\boldsymbol{\xi}\cdot\nabla)^m u(\boldsymbol{x}), \qquad \text{and} \qquad \tilde{p}_m(\boldsymbol{\xi}) = \frac{1}{m!}(\boldsymbol{\xi}\cdot\tilde{\nabla})^m \tilde{u}(\tilde{\boldsymbol{x}}).$$

Then we have

$$\tilde{p}_m(\boldsymbol{\xi}) = \frac{1}{m!}(\boldsymbol{\xi}\cdot F_\tau^T \nabla)^m u(\boldsymbol{x}) = p_m(F_\tau \boldsymbol{\xi}).$$

By Lemma 2.3,

$$|\tilde{p}_m(\boldsymbol{\xi})| \;\le\; c\,\mathcal{D}_m\,(\boldsymbol{\xi}\cdot F_\tau^T Q_m F_\tau\;\boldsymbol{\xi})^{m/2},$$

where Q_m is the matrix defined in (12) by ϕ_m and S_m. By Lemma 2.1, we have

$$\tilde{\mathcal{D}}_m \le c \sup_{\|\boldsymbol{\xi}\|=1} |\tilde{p}_m(\boldsymbol{\xi})| \;\le\; c\,\mathcal{D}_m\,[\sup_{\|\boldsymbol{\xi}\|=1} \boldsymbol{\xi}\cdot F_\tau^T Q_m F_\tau\;\boldsymbol{\xi}]^{m/2}.$$

Expanding $\boldsymbol{\xi}\cdot F_\tau^T Q_m F_\tau\;\boldsymbol{\xi}$, we get

$$\begin{aligned}
\boldsymbol{\xi}\cdot F_\tau^T Q_m F_\tau\;\boldsymbol{\xi} &= \bar{\mu}^{-1}\begin{bmatrix} \sqrt{\bar{r}}\xi \\ \eta/\sqrt{\bar{r}} \end{bmatrix}\cdot R_{\phi_m-\bar{\psi}}\begin{bmatrix} 1/S_m^2 & 0 \\ 0 & 1 \end{bmatrix}\cdot R^T_{\phi_m-\bar{\psi}}\begin{bmatrix} \sqrt{\bar{r}}\xi \\ \eta/\sqrt{\bar{r}} \end{bmatrix} \\
&= (\bar{\mu} S_m)^{-1}\{\cos^2(\phi_m-\bar{\psi})\cdot(\frac{\bar{r}}{S_m}\xi^2 + \frac{S_m}{\bar{r}}\eta^2) + \sin^2(\phi_m-\bar{\psi}) \\
&\qquad \cdot(\frac{1}{\bar{r}S_m}\eta^2 + \bar{r}S_m\xi^2) + 2\xi\eta\sin(\phi_m-\bar{\psi})\cos(\phi_m-\bar{\psi})(\frac{1}{S_m} - S_m)\} \\
&\le 2(\bar{\mu} S_m)^{-1}\{\cos^2(\phi_m-\bar{\psi})(\frac{\bar{r}}{S_m} + \frac{S_m}{\bar{r}}) + \sin^2(\phi_m-\psi)\;(\bar{r}S_m)\},
\end{aligned}$$

where in the last step above we used the facts $\|\boldsymbol{\xi}\| = 1$, $\bar{r} \ge 1$, and $S_m \ge 1$. The conclusion of this lemma follows from the above inequality. $\square$

Similarly, we have an inequality to bound the magnitude $\mathcal{D}_m$ of $\nabla^m u(\boldsymbol{x})$ in terms of the magnitude $\tilde{\mathcal{D}}_m$ of $\tilde{\nabla}^m \tilde{u}(\tilde{\boldsymbol{x}})$.

Lemma 3.2. *For each element $\tau \in \mathcal{T}_N$, we have*

$$\mathcal{D}_m \;\le\; c\,(\bar{\mu}\bar{r})^{m/2}\tilde{\mathcal{D}}_m, \qquad \forall \boldsymbol{x} \in \tau.$$

Proof. Let $p_m(\boldsymbol{\xi})$ and $\tilde{p}_m(\boldsymbol{\xi})$ be the polynomials defined by $\nabla^m u(\boldsymbol{x})$ and $\tilde{\nabla}^m \tilde{u}(\tilde{\boldsymbol{x}})$ as in (18), respectively. Note that

$$|\tilde{p}_m(\boldsymbol{\xi})| \le \tilde{\mathcal{D}}_m\,(\boldsymbol{\xi}\cdot\boldsymbol{\xi})^{m/2},$$

therefore

$$|p_m(\boldsymbol{\xi})| = |\tilde{p}_m(F_\tau^{-1}\boldsymbol{\xi}))| \le \tilde{\mathcal{D}}_m \left[\boldsymbol{\xi} \cdot (F_\tau F_\tau^T)^{-1}\boldsymbol{\xi}\right]^{m/2}.$$

The conclusion of this lemma follows immediately from Lemma 2.1 and that

$$\sup_{\|\boldsymbol{\xi}\|=1} \boldsymbol{\xi} \cdot (F_\tau F_\tau^T)^{-1}\boldsymbol{\xi} \le \bar{\mu}\bar{r}. \qquad \square$$

Next, we make an assumption on the metric M. Let $M(\boldsymbol{x})$ be decomposed in the form (17). We assume that there exists a positive number δ such that for all $\boldsymbol{x} \in \Omega$ and any neighborhood $\mathcal{N}_{\boldsymbol{x}}$ of $\boldsymbol{x}$ with radius (under metric-M) less than δ,

$$\left\{ \begin{array}{l} c\bar{\mu} \le \mu(\boldsymbol{y}) \le c\bar{\mu}, \\ c\bar{r} \le r(\boldsymbol{y}) \le c\bar{r}, \\ |\psi(\boldsymbol{y}) - \bar{\psi}| \le c\bar{r}^{-1}, \end{array} \right. \quad \forall \boldsymbol{y} \in \mathcal{N}_{\boldsymbol{x}}, \tag{19}$$

where $\bar{\mu}, \bar{r}$, and $\bar{\psi}$ are the mean values of μ, r, and ψ on $\mathcal{N}_{\boldsymbol{x}}$, respectively. This assumption is basically to require the continuity of M. Indeed, if $\mu(\boldsymbol{x}), r(\boldsymbol{x})$, and $\psi(\boldsymbol{x})$ are uniformly continuous over Ω, then the above assumption holds trivially.

In mesh adaptation, μ and r in metric M will be defined in terms of the magnitude and anisotropic ratio of $\nabla^{k+1}u$. In general $\mathcal{D}_{k+1}$ can be zero and $\mathcal{S}_{k+1}$ can be ∞. In order to have (19) satisfied, we introduce two functions that level-off the magnitude and anisotropic ratio of $\nabla^{k+1}u$. More precisely, Let $d_* > 0$ and $S^* \ge 1$ be parameters. Define

$$\begin{array}{l} D_{k+1}(\boldsymbol{x}) = \max(d_*, \mathcal{D}_{k+1}(\boldsymbol{x})), \\ S_{k+1}(\boldsymbol{x}) = \min(S^*, \mathcal{S}_{k+1}(\boldsymbol{x})). \end{array}$$

When d_* is small and S^* is large, D_{k+1} and S_{k+1} are almost identical to $\mathcal{D}_{k+1}$ and $\mathcal{S}_{k+1}$, respectively. When $\nabla^{k+1}u$ is uniformly continuous, we have relations similar to assumption (19) that hold for D_{k+1}, S_{k+1}, and ϕ_{k+1}, i.e.,

$$\left\{ \begin{array}{l} c\bar{D}_{k+1} \le D_{k+1}(\boldsymbol{y}) \le c\bar{D}_{k+1}, \\ c\bar{S}_{k+1} \le S_{k+1}(\boldsymbol{y}) \le c\bar{S}_{k+1}, \\ |\phi_{k+1}(\boldsymbol{y}) - \bar{\phi}_{k+1}| \le c[\bar{S}_{k+1}]^{-1}. \end{array} \right. \quad \forall \boldsymbol{y} \in \mathcal{N}_{\boldsymbol{x}}, \tag{20}$$

Now we state the main theorem of this paper.

Theorem 3.1. *Let M be a Riemannian metric on Ω satisfying assumption* (19). *$\{\mathcal{T}_N\}$ is a family of triangulation of Ω that is quasi-uniform under metric M. Let k be a positive integer, and let Π_k be an interpolation operator that preserves $\mathcal{P}_k$. For $0 \le m \le k$ and $p \in [1, \infty]$ satisfying* (15)*, we have*

$$\begin{aligned} \Big(\sum_{\tau\in\mathcal{T}_N} |u - \Pi_k u|^p_{m,p,\tau}\Big)^{1/p} \le cN^{-(k+1-m)/2} \ &\Big\{\int_\Omega \mu\Big\}^{(k+1-m)/2} \\ &\cdot \Big\{\int_\Omega (\mu\bar{r})^{mp/2} \Big(\frac{V}{\mu S_{k+1}}\Big)^{(k+1)p/2} |D_{k+1}|^p\Big\}^{1/p}, \end{aligned} \tag{21}$$

where

$$V = \cos^2(\phi_{k+1} - \bar{\psi})\Big(\frac{r}{S_{k+1}} + \frac{S_{k+1}}{r}\Big) + \sin^2(\phi_{k+1} - \bar{\psi})(rS_{k+1}). \tag{22}$$

Furthermore, among all the Riemannian metrics, the optimal bound of the above estimate is attained when M is defined to be

$$
\begin{aligned}
M_{k+1,m,p} \equiv (S_{k+1})^{-\frac{(k+1-m)p}{(k+1-m)p+2}} (D_{k+1})^{\frac{2p}{(k+1-m)p+2}} R_{\phi_{k+1}} \\
\cdot \begin{bmatrix} \sqrt{\frac{k+1+m}{k+1-m}} \frac{1}{S_{k+1}} & 0 \\ 0 & \sqrt{\frac{k+1-m}{k+1+m}} S_{k+1} \end{bmatrix} R^T_{\phi_{k+1}}.
\end{aligned} \tag{23}
$$

If $\{\mathcal{T}_N\}$ is quasi-uniform under $M_{k+1,m,p}$, we have

$$
\begin{aligned}
&(\sum_{\tau\in\mathcal{T}_N} |u-\Pi_k u|^p_{m,p,\tau})^{1/p} \\
&\qquad \le cN^{-(k+1-m)/2} \|(S_{k+1})^{-(k+1-m)/2} D_{k+1}\|_{L^{2/(k+1-m+2/p)}(\Omega)}.
\end{aligned} \tag{24}
$$

Proof. Consider an element $\tau \in \mathcal{T}_N$. Denote by $\bar{\mu}, \bar{r}, \bar{\psi}$ the mean values of μ, r, and ψ on τ, respectively. It follows from Lemma 3.2 that

$$
\begin{aligned}
|u-\Pi_k u|^p_{m,p,\tau} &= \int_{\tilde{\tau}} |\nabla^m (u-\Pi_k u)(\tilde{\boldsymbol{x}})|^p \cdot \det(\frac{\partial \boldsymbol{x}}{\partial \tilde{\boldsymbol{x}}})\ d\tilde{\boldsymbol{x}} \\
&\le c(\bar{\mu}\bar{r})^{mp/2}\ \frac{|\tau|}{|\tilde{\tau}|}\ \int_{\tilde{\tau}} |\tilde{\nabla}^m (\tilde{u}-\tilde{\Pi}_k \tilde{u})(\tilde{\boldsymbol{x}})|^p\ d\tilde{\boldsymbol{x}}.
\end{aligned}
$$

Because $\tilde{\tau}$ is shape regular, we have from the classical error estimate (14) that

$$
|u-\Pi_k u|^p_{m,p,\tau} \le c|\tilde{\tau}|^{(k+1-m)p/2}\ (\bar{\mu}\bar{r})^{mp/2}\ \|\tilde{\nabla}^{k+1}\tilde{u}\|^p_{\infty,\tilde{\tau}} \cdot |\tau|,
$$

which implies by Lemma 3.1 that

$$
\begin{aligned}
|u-\Pi_k u|^p_{m,p,\tau} &\le c|\tilde{\tau}|^{(k+1-m)p/2} (\bar{\mu}\bar{r})^{mp/2} \\
&\quad \cdot (\max_{\forall \boldsymbol{x}\in\tau} \frac{V(\boldsymbol{x})}{\bar{\mu} S_{k+1}(\boldsymbol{x})})^{-(k+1)p/2} \cdot \|\nabla^{k+1} u\|^p_{\infty,\tau} \cdot |\tau|.
\end{aligned} \tag{25}
$$

Summing up the above inequality for all $\tau \in \mathcal{T}_N$, we have from assumptions (19) and (20) that

$$
\begin{aligned}
\sum_{\tau\in\mathcal{T}_N} |u-\Pi_k u|^p_{m,p,\tau} &\le c(\max_{\tau\in\mathcal{T}_N} |\tilde{\tau}|)^{(k+1-m)p/2} \sum_{\tau\in\mathcal{T}_N} (\bar{\mu}\bar{r})^{mp/2} \\
&\quad \cdot [\max_{\forall \boldsymbol{x}\in\tau} \frac{V(\boldsymbol{x})}{\bar{\mu} S_{k+1}(\boldsymbol{x})}]^{(k+1)p/2} \cdot \|\nabla^{k+1} u\|^p_{\infty,\tau} \cdot |\tau| \\
&\le c(\max_{\tau\in\mathcal{T}_N} |\tilde{\tau}|)^{(k+1-m)p/2} \int_\Omega (\mu r)^{mp/2} [\frac{V}{\mu S_{k+1}}]^{(k+1)p/2} |D_{k+1}|^p.
\end{aligned} \tag{26}
$$

Now we estimate $|\tilde{\tau}|$. By the assumption that $\{\mathcal{T}_N\}$ is quasi-uniform under metric M, the sizes of all $\tilde{\tau} = F_\tau^{-1}(\tau - \boldsymbol{x}_c)$ are of the same order. Hence

$$
\max_{\tau\in\mathcal{T}_N} |\tilde{\tau}| \le cN^{-1} \sum_{\tau\in\mathcal{T}_N} |\tilde{\tau}| = cN^{-1} \sum_{\tau\in\mathcal{T}_N} \bar{\mu}|\tau| = cN^{-1} \int_\Omega \mu.
$$

Putting the above inequality into (26), we have estimate (21).

Next, we consider for what metric M the bound at the right hand side of (21) is the smallest. Clearly, the optimal metric requires $\psi(\boldsymbol{x}) = \phi_{k+1}(\boldsymbol{x})$. In this case

$$
V(\boldsymbol{x}) = \frac{r}{S_{k+1}} + \frac{S_{k+1}}{r},
$$

and the integrand of the second integral in the right hand side of (21) becomes

$$
\begin{aligned}
&(\mu r)^{mp/2}(\tfrac{V}{\mu S_{k+1}})^{(k+1)p/2}|D_{k+1}|^p \\
&= (\mu)^{-(k+1-m)p/2}(S_{k+1})^{-(k+1)p} \cdot \left(\tfrac{(r^2+S_{k+1}^2)^{k+1}}{r^{k+1-m}}\right)^{p/2} \cdot |D_{k+1}|^p.
\end{aligned} \tag{27}
$$

It is easy to identify that when

$$r = \sqrt{\frac{k+1-m}{k+1+m}} S_{k+1},$$

the right hand side of (27) attains the minimum

$$\frac{[2(k+1)]^{(k+1)p/2}}{(k+1+m)^{(k+1+m)p/4}(k+1-m)^{(k+1-m)p/4}}(\mu S_{k+1})^{-(k+1-m)p/2} \cdot |D_{k+1}|^p.$$

Now estimate (21) becomes

$$
\begin{aligned}
(\sum_{\tau \in \mathcal{T}_N} |u - \Pi_k u|_{m,p,\tau}^p)^{1/p} \le cN^{-(k+1-m)/2} \ &\{\int_\Omega \mu\}^{(k+1-m)/2} \\
&\cdot \{\int_\Omega (\mu S_{k+1})^{-(k+1-m)p/2} |D_{k+1}|^p\}^{1/p}.
\end{aligned} \tag{28}
$$

To determine the $\mu(\boldsymbol{x})$ for the optimal bound of the above estimate, let

$$q = \frac{1}{2}(k+1-m)p,$$

and let

$$f = \mu^{q/(q+1)}, \quad g = (\mu S_{k+1})^{-q/(q+1)}(D_{k+1})^{p/(q+1)}.$$

Recall Hölder's inequality,

$$\int fg \le \|f\|_{L^{(q+1)/q}} \cdot \|g\|_{L^{q+1}}, \qquad \forall f, g,$$

and that the equality holds iff $f^{(q+1)/q}$ is a multiple of g^{q+1}. Then we have

$$
\begin{aligned}
&\{\int_\Omega \mu\}^{(k+1-m)/2} \ \{\int_\Omega (\mu S_{k+1})^{-(k+1-m)p/2} |D_{k+1}|^p\}^{1/p} \\
&= \left(\{\int_\Omega |f|^{(q+1)/q}\}^{q/(q+1)} \ \{\int_\Omega |g|^{q+1}\}^{1/(q+1)}\right)^{(q+1)/p} \\
&\ge \left(\int_\Omega |fg|\right)^{(q+1)/p} \\
&= \left(\int_\Omega (S_{k+1})^{-q/(q+1)} |D_{k+1}|^{p/(q+1)}\right)^{(q+1)/p} \\
&= \|(S_{k+1})^{-(k+1-m)/2} D_{k+1}\|_{L^{2/(k+1-m+2/p)}(\Omega)},
\end{aligned}
$$

and the equality is attained when

$$
\begin{aligned}
\mu(\boldsymbol{x}) \ &= (S_{k+1})^{-q/(q+1)}(D_{k+1})^{p/(q+1)} \\
&= (S_{k+1})^{-\frac{(k+1-m)p}{(k+1-m)p+2}}(D_{k+1})^{\frac{2p}{(k+1-m)p+2}}.
\end{aligned}
$$

This completes the proof of the theorem. □

Remark 3.1. Theorem 3.1 is a generalization of the classical result (16) for quasi-uniform triangulations. Indeed, when M is constant, the fact that $\{\mathcal{T}_N\}$ is quasi-uniform under metric M implies that it is quasi-uniform in the usual sense. In this case, $\mu(\boldsymbol{x})$ and $r(\boldsymbol{x})$ are constant. Therefore $V(\boldsymbol{x}) \leq cS_{k+1}$, and (21) is reduced to (16).

Remark 3.2. A triangulation family $\{\mathcal{T}_N\}$ being quasi-uniform under the optimal metric $M_{k+1,m,p}$ can be characterized by two features: (i) For each element τ, the mapping $\tilde{\boldsymbol{x}} = F_\tau^{-1}(\boldsymbol{x} - \boldsymbol{x}_c)$ not only transforms τ into a shape regular triangle, but also makes the level-0 lines of $\tilde{p}_{k+1}(\boldsymbol{\xi})$ for $\tilde{\nabla}^{k+1}\tilde{u}$ evenly distributed across the entire plane. In other words, $\{\mathcal{T}_N\}$ makes $\tilde{\nabla}^{k+1}\tilde{u}$ isotropic on each element. (ii) Estimate (25) for the $W^{m,p}(\tau)$-seminorm of the interpolation error is of the same magnitude on every element τ. In other words, $\{\mathcal{T}_N\}$ makes the $W^{m,p}$-seminorm of the error over every element evenly distributed. Therefore, the optimal metric can also be considered as derived from the so-called equidistribution principle. This principle has been used extensively to justify the selection of optimal or nearly optimal meshes; see, e.g., [10, 11, 16, 17].

Remark 3.3. The condition $\psi = \phi_{k+1}$ for a triangle aligned with $\nabla^{k+1}u$ can be relaxed to $|\psi - \phi_{k+1}| \leq c[S_{k+1}]^{-1}$. All the conclusions requiring $\psi = \phi_{k+1}$ still hold under this weak sense of alignment.

We illustrate the conclusion of Theorem 3.1 in several special cases.

Case 1: $k = 1, m = 0$. In this case Π_k is the linear interpolation operator. $\nabla^2 u(\boldsymbol{x})$ is the Hessian of u at $\boldsymbol{x}$. The magnitude and anisotropic ratio of $\nabla^2 u$ are respectively

$$\mathcal{D}_2 = \frac{1}{2}(|\lambda_1| + |\lambda_2|), \qquad \mathcal{S}_2 = \sqrt{|\frac{\lambda_2}{\lambda_1}|},$$

where $|\lambda_1| \leq |\lambda_2|$ are the two eigenvalues of $\nabla^2 u$. Neglecting the cut-off for D_2 and S_2, we have

$$\begin{array}{rcl} \frac{1}{2}|\lambda_2| \leq & D_2 & \leq |\lambda_2|, \\ \frac{1}{2}\sqrt{\det(\nabla^2 u)} \leq & \frac{D_2}{S_2} & \leq \sqrt{\det(\nabla^2 u)}. \end{array}$$

Inequality (21) becomes

$$\|u - \Pi_1 u\|_{L^p(\Omega)} \leq cN^{-1} \ \{\int_\Omega \mu\} \ \{\int_\Omega (\frac{V}{\mu})^p \ (\sqrt{\det(\nabla^2 u)})^p\}^{1/p}.$$

The optimal metric for this case is

$$\begin{aligned} M_{2,0,p} &= (\det(\nabla^2 u))^{p/(2p+2)} R_{\phi_2} \cdot \begin{bmatrix} \sqrt{|\frac{\lambda_1}{\lambda_2}|} & 0 \\ 0 & \sqrt{|\frac{\lambda_2}{\lambda_1}|} \end{bmatrix} R_{\phi_2}^T \\ &= (\det(\nabla^2 u))^{-1/(2p+2)} \cdot \mathrm{abs}(\nabla^2 u), \end{aligned}$$

where

$$\mathrm{abs}(\nabla^2 u) = R_{\phi_2} \cdot \begin{bmatrix} |\lambda_1| & 0 \\ 0 & |\lambda_2| \end{bmatrix} R_{\phi_2}^T.$$

When $\{\mathcal{T}_N\}$ is quasi-uniform under this metric, the error estimate is

$$\|u - \Pi_1 u\|_{L^p(\Omega)} \leq cN^{-1} \ \| \ \sqrt{\det(\nabla^2 u)} \ \|_{L^{p/(p+1)}(\Omega)}.$$

This is exactly the conclusion in [10] in two dimensions. Also, if a triangulation is quasi-uniform under metric $M_{2,0,p}$, the aspect ratio of each element should be about $\sqrt{|\lambda_2/\lambda_1|}$. This aspect ratio coincides with that obtained based on the exact error formulas for linear interpolation of quadratic functions on a triangle [8]. Note that in [8] the notation for λ_1 and λ_2 are switched.

Case 2: $k=1, m=1$. In this case the optimal metric is

$$M_{2,1,p} = |\lambda_1|^{-1/(p+2)}|\lambda_2|^{(p-1)/(p+2)} \cdot R_{\phi_2} \begin{bmatrix} \sqrt{3}|\lambda_1| & 0 \\ 0 & \frac{|\lambda_2|}{\sqrt{3}} \end{bmatrix} R_{\phi_2}^T,$$

and the error estimate is

$$|u-\Pi_1 u|_{1,p,\Omega} \le cN^{-1/2} \| \, |\lambda_1|^{1/4}|\lambda_2|^{3/4} \|_{L^{2p/(p+2)}(\Omega)}.$$

In particular, when $p=2$, the optimal metric is

$$M_{2,1,2} = |\lambda_1|^{-1/4}|\lambda_2|^{1/4} \cdot R_{\phi_2} \begin{bmatrix} \sqrt{3}|\lambda_1| & 0 \\ 0 & \frac{|\lambda_2|}{\sqrt{3}} \end{bmatrix} R_{\phi_2}^T.$$

When $\{\mathcal{T}_N\}$ is quasi-uniform under $M_{2,1,2}$, the error estimate is

$$|u-\Pi_1 u|_{H^1(\Omega)} \le cN^{-1/2} \| \, |\lambda_1|^{1/4}|\lambda_2|^{3/4} \, \|_{L^1(\Omega)}.$$

This implies that the H^1-seminorm of the linear interpolation error is about the smallest when each element is align with the orientation of $\nabla^2 u$, the aspect ratio is equal to $S_2/\sqrt{3} \approx 0.577\sqrt{|\lambda_2/\lambda_1|}$, and the area is be proportional to

$$[\det(M_{2,1,2})]^{-1/2} = |\lambda_1|^{-1/4}|\lambda_2|^{-3/4}.$$

It is shown in [8] that the H^1-seminorm of the linear interpolation error on a general element τ is minimum when τ is aligned with $\nabla^2 u$ and takes the aspect ratio $c\sqrt{|\lambda_2/\lambda_1|}$ with $c \in [0.849, 1.178]$. The minimum H^1-error on τ is

$$\|\nabla(u-\Pi_1 u)\|_{L^2(\tau)} \approx c|\lambda_1|^{1/4}|\lambda_2|^{3/4} \, |\tau|.$$

The best aspect ratio predicted in this paper is slightly smaller than that in [8], but both of them are in the same order. The quasi-uniformity under metric $M_{2,1,2}$ again implies the equidistribution of the H^1-error over every element.

Case 3: $k=2, p=2$. In this case Π_k is the operator for quadratic Lagrange interpolation. The optimal metric is

$$M_{3,0,2} = (S_3)^{-3/4}(D_3)^{1/2} R_{\phi_3} \cdot \begin{bmatrix} \frac{1}{S_3} & 0 \\ 0 & S_3 \end{bmatrix} R_{\phi_3}^T$$

for the L^2-norm of the error, and

$$M_{3,1,2} = (S_3)^{-2/3}(D_3)^{2/3} R_{\phi_3} \cdot \begin{bmatrix} \frac{\sqrt{2}}{S_3} & 0 \\ 0 & \frac{S_3}{\sqrt{2}} \end{bmatrix} R_{\phi_3}^T$$

for the H^1-seminorm of the error. These imply that to minimize the L^2-norm of the error, the aspect ratio of each element should be equal to S_3, and the area should be proportional to $(S_3)^{3/4}(D_3)^{-1/2}$. To minimize the H^1-seminorm of the error, the aspect ratio of each element should be $S_3/\sqrt{2} \approx 0.707 S_3$, and the area should be proportional to $(S_3)^{2/3}(D_3)^{-2/3}$. It is shown in [9] that for quadratic interpolation, the minimum values of both the L^2-norm and the H^1-seminorm of the interpolation error on an element τ are attained when the aspect ratio of τ is about the anisotropic ratio of $\nabla^3 u$. With this aspect ratio, the L^2-norm and H^1-seminorm on τ are proportional to $S_3^{-3/2} D_3 |\tau|^2$ and $(S_3)^{-1} D_3 |\tau|^{3/2}$, respectively. When triangulation $\{\mathcal{T}_N\}$ is quasi-uniform under $M_{3,0,2}$ or $M_{3,1,2}$, we have respectively that

$$\|u - \Pi_2 u\|_{L^2(\Omega)} \leq cN^{-3/2} \|(S_3)^{-3/2} D_3\|_{L^{1/2}(\Omega)}$$

and

$$|u - \Pi_2 u|_{H^1(\Omega)} \leq cN^{-1} \|(S_3)^{-1} D_3\|_{L^{2/3}(\Omega)}.$$

4. Numerical results

In this section, we present a numerical example to compare various norms of the interpolation errors using triangulations that are quasi-uniform under different metrics $M_{k+1,m,p}$ developed in the previous section. We consider linear and quadratic interpolations of the following function on $\Omega = \{(x, y) : \ r = \sqrt{x^2 + y^2} \leq 1, \ x \geq 0, y \geq 0\}$:

$$u(x, y) = u(r) = e^{-10000(r-0.5)^2} + r^3. \tag{29}$$

We first study the magnitudes, anisotropic ratios, and orientations of $\nabla^2 u$ and $\nabla^3 u$. By elementary calculation, we see that for each $\boldsymbol{x}$ with $|\boldsymbol{x}| = r$,

$$\begin{aligned} p_2(\boldsymbol{\xi}) &= \tfrac{1}{2!} (\boldsymbol{\xi} \cdot \nabla)^2 u(\boldsymbol{x}) \\ &= \tfrac{1}{2} [u_{rr} r^{-2} (\boldsymbol{x} \cdot \boldsymbol{\xi})^2 + u_r r^{-3} (\boldsymbol{x} \times \boldsymbol{\xi})^2]. \end{aligned} \tag{30}$$

It can be verified by the definition of anisotropic measures that

$$\begin{cases} \mathcal{D}_2 &= \tfrac{1}{2} (|u_{rr}| + |r^{-1} u_r|), \\ \mathcal{S}_2 &= \sqrt{\left| \dfrac{r u_{rr}}{u_r} \right|}. \end{cases} \tag{31}$$

The orientation of $\nabla^2 u(\boldsymbol{x})$ is the angular direction at $\boldsymbol{x}$.

The anisotropic measures of $\nabla^3 u$ can be determined similarly. Indeed, let $\boldsymbol{x} = r[\cos z, \sin z]^T$, and $\boldsymbol{\xi} = [\cos t, \sin t]^T$. It can be verified that

$$\begin{aligned} p_3(\boldsymbol{\xi}) &= \frac{1}{3!} (\boldsymbol{\xi} \cdot \nabla)^3 u(\boldsymbol{x}) \\ &= \frac{1}{3!} [u_{rrr} \cos^3(t - z) + 3 u_{rr} r^{-1} \sin^2(t - z) \cos(t - z) \\ &\quad - 3 u_r r^{-2} (\frac{1}{2} \sin(2t) \sin(2z) \cos(t - z) + \cos(t + z) \sin(t + z) \sin(t - z))]. \end{aligned} \tag{32}$$

Since our main concern is the region around $r = 0.5$ where $\nabla^3 u$ is large, we have for this region

$$p_3(\boldsymbol{\xi}) \approx \frac{1}{3!} \cos(t - z) [u_{rrr} \cos^2(t - z) + 3 u_{rr} r^{-1} \sin^2(t - z)]. \tag{33}$$

From the above expression we see that

$$\begin{cases} \mathcal{D}_3 \approx \tfrac{1}{3!} (|u_{rrr}| + 3 |r^{-1} u_{rr}|), \\ \mathcal{S}_3 \approx \sqrt{\left| \dfrac{r u_{rrr}}{3 \ u_{rr}} \right|}. \end{cases} \tag{34}$$

The orientation of $\nabla^3 u$ is also the angular direction.

We make a cut-off of the anisotropic ratio as follows:

$$S_{k+1} = \min(1000, \mathcal{S}_{k+1}), \qquad k = 1, 2.$$

This means the largest aspect ratio allowed is 1000. We do not level-off the magnitude of $\nabla^{k+1}u$, i.e., we define $D_{k+1} = \mathcal{D}_{k+1}$ directly, because it is already bounded away from 0. Using S_2, D_2 and S_3, D_3, together with the angular direction for orientation, we can define the optimal metrics $M_{2,m,p}$ and $M_{3,m,p}$ according to (23).

Next, we describe how to generate a triangulation of Ω that is quasi-uniform under a metric $M_{k+1,m,p}$. Since we wish to have a more precise control of the geometric features of each element, we choose to "manually" create the anisotropic triangulations. For more general domain and applications, readers may resort to some robust algorithms, e.g., [5, 6], to generate the anisotropic mesh under a given Riemannian metric. To start our process, we note that u is a function of the radial variable only. We place all the grid points along $N_r + 1$ circles of radii: $0 = r_0 < r_1 < \cdots < r_{N_r} = 1$. On each circle $r = r_j$, we distribute evenly $N_a(r_j)$ points. By the fact that the orientation of $\nabla^{k+1}u$ is the angular direction (at least this is the case around $r = 0.5$ where the interpolation error mostly comes from), we see that the distance between two neighboring grids on $r = r_j$ should be much greater than $r_{j+1} - r_j$ or $r_j - r_{j-1}$. Therefore, we connect all the grid points on the same circle $r = r_j$, and connect each grid on r_j with the closest grid on $r = r_{j+1}$. In this way the longest sides of the triangles in the annulus $r_j \le r \le r_{j+1}$ are approximately

$$h_{long} \approx \frac{\pi}{2}\frac{r_j}{N_a(r_j)},$$

and the heights over the longest side for these triangles are

$$h_{short} \approx r_{j+1} - r_j.$$

Thus the aspect ratios and the areas of such triangles are

$$\begin{aligned} r_{asp} &\approx h_{long}/h_{short} \approx \tfrac{\pi}{2}\frac{r_j}{N_a(r_j)(r_{j+1} - r_j)}, \\ |\tau| &\approx \tfrac{1}{2}h_{long}h_{short} \approx \tfrac{\pi}{2}\frac{r_j(r_{j+1} - r_j)}{N_a(r_j)}, \end{aligned}$$

respectively. In order to make the triangulation quasi-uniform under metric $M_{k+1,m,p}$, we require that

$$\begin{aligned} r_{asp} &\approx \sqrt{\tfrac{k+1-m}{k+1+m}}S_{k+1}, \\ |\tau| &\approx (S_{k+1})^{-\frac{(k+1-m)p}{(k+1-m)p+2}}(D_{k+1})^{\frac{2p}{(k+1-m)p+2}}, \end{aligned} \tag{35}$$

which results in a system of non-linear equations for $r_j,\ 1 \le j \le N_r - 1$, as follows:

$$(S_{k+1})^{\frac{1}{(k+1-m)p+2}}(D_{k+1})^{\frac{p}{(k+1-m)p+2}}(r_{j+1} - r_j) = const, \qquad \forall j.$$

Once the distribution of r_j is solved, the number of points $N_a(r_j)$ is calculated according to either the condition for the area or the condition for the aspect ratio of the elements.

We consider specifically the linear ($k = 1$) and the quadratic ($k = 2$) interpolation of u, and we measure the errors in L^1, L^2, and L^∞ norms (i.e., $m = 0$, $p = 1, 2, \infty$), as well as in H^1-seminorm (i.e., $m = 1$, $p = 2$). We show in Figures 2 and 3 the anisotropic meshes generated in the above manner. Note that in these graphs the total number of elements are all about $4,000$. We list in Tables 1 and 2 the linear and quadratic interpolation errors in various norms. It is clearly seen that in all the cases except the L^1-error of quadratic interpolation, the smallest $W^{m,p}$-norm of the interpolation error is obtained when the mesh is generated according to the optimal metric $M_{k+1,m,p}$. For the case of quadratic interpolation, the meshes generated with metrics $M_{3,0,1}$ and $M_{3,0,2}$ are close to each other, and the L^1-errors based on $M_{3,0,1}$ mesh are close to the smallest ones.

TABLE 1. Various norms of the *linear* interpolation error for function (29) based on meshes that are quasi-uniform under different metrics $M_{2,m,p}$. N_v and N_e represent respectively the total number of nodes and elements in the triangulation.

	N_v	N_e	Metric	$\|e\|_{L^1}$	$\|e\|_{L^2}$	$\|e\|_{L^\infty}$	$\|e\|_{H^1}$
$N_e \approx 1000$	533	968	$M_{2,0,1}$	**1.49339e-3**	4.67063e-3	6.23914e-2	2.12481e+0
	539	986	$M_{2,0,2}$	1.56953e-3	**3.07417e-3**	5.41552e-2	1.92157e+0
	559	1050	$M_{2,0,\infty}$	5.86828e-3	9.22595e-3	**3.26411e-2**	1.41867e+0
	565	1067	$M_{2,1,2}$	4.79373e-3	4.45060e-3	4.37467e-2	**1.21329e+0**
$N_e \approx 4000$	2115	4036	$M_{2,0,1}$	**3.16475e-4**	8.73823e-4	1.19351e-2	1.00562e+0
	2072	3967	$M_{2,0,2}$	3.72421e-4	**6.41605e-4**	6.10854e-3	7.91351e-1
	2089	4048	$M_{2,0,\infty}$	1.23867e-3	1.64885e-3	**4.79847e-3**	6.29452e-1
	2080	4046	$M_{2,1,2}$	3.53340e-3	5.57094e-3	2.05686e-2	**5.82440e-1**
$N_e \approx 16000$	8151	15920	$M_{2,0,1}$	**7.75799e-5**	2.09533e-4	2.49726e-3	4.98888e-1
	8105	15863	$M_{2,0,2}$	9.24885e-5	**1.59773e-4**	1.50450e-3	3.89114e-1
	8368	16480	$M_{2,0,\infty}$	2.85356e-4	3.64991e-4	**1.18945e-3**	3.08177e-1
	8070	15918	$M_{2,1,2}$	6.74756e-4	8.69082e-4	2.22446e-3	**2.89580e-1**

TABLE 2. Various norms of the *quadratic* interpolation error for function (29) based on meshes that are quasi-uniform under different metrics $M_{3,m,p}$. N_v and N_e represent respectively the total number of nodes and elements in the triangulation.

	N_v	N_e	Metric	$\|e\|_{L^1}$	$\|e\|_{L^2}$	$\|e\|_{L^\infty}$	$\|e\|_{H^1}$
$N_e \approx 1000$	576	1063	$M_{3,0,1}$	3.84101e-5	1.99156e-4	3.53582e-3	2.54690e-1
	555	1033	$M_{3,0,2}$	**3.70225e-5**	**1.35188e-4**	2.63480e-3	1.83796e-1
	536	1015	$M_{3,0,\infty}$	1.01811e-4	1.73567e-4	**1.41021e-3**	1.54345e-1
	556	1057	$M_{3,1,2}$	1.30927e-4	1.95614e-4	2.01960e-3	**1.25994e-1**
$N_e \approx 4000$	2012	3858	$M_{3,0,1}$	5.23449e-6	2.62015e-5	4.56166e-4	6.37554e-2
	2040	3931	$M_{3,0,2}$	**4.29827e-6**	**1.53680e-5**	2.56361e-4	4.33876e-2
	2126	4139	$M_{3,0,\infty}$	1.12890e-5	1.88816e-5	**1.71215e-4**	3.61567e-2
	2066	4029	$M_{3,1,2}$	2.30966e-5	3.43089e-5	2.00585e-4	**3.27660e-2**
$N_e \approx 16000$	8202	16068	$M_{3,0,1}$	5.86753e-7	2.87826e-6	4.76983e-5	1.49075e-2
	8154	16017	$M_{3,0,2}$	**5.42702e-7**	**1.87424e-6**	2.70740e-5	1.10013e-2
	8047	15880	$M_{3,0,\infty}$	1.55820e-6	2.52208e-6	**2.29942e-5**	9.39049e-3
	8027	15854	$M_{3,1,2}$	2.78704e-6	3.96082e-6	2.39793e-5	**8.29941e-3**

5. DISCUSSIONS

In the previous sections we introduced some anisotropic measures for the higher order derivative $\nabla^{k+1}u$ a function u. The magnitude of $\nabla^{k+1}u$ is equivalent to its usual Euclidean norm. The orientation is the direction along which the absolute value of the $k+1$-th directional derivative is about the smallest, while along its perpendicular direction it is about the largest. The anisotropic ratio measures the strength of the anisotropic behavior of $\nabla^{k+1}u$. These quantities are invariant under translation and rotation of xy-coordinates. They correspond to the size, orientation, and aspect ratio for triangular elements. In terms of these anisotropic measures, we derive an anisotropic error estimate for the interpolation over triangulations that are quasi-uniform under a given Riemannian metric M. It is identified from this estimate that among triangulations of a fixed number of elements the interpolation error is nearly the minimum when the elements are aligned with the orientation of $\nabla^{k+1}u$, the aspect ratios are about the anisotropic ratio of $\nabla^{k+1}u$, and the areas make the error over each element be evenly distributed.

There are a couple of immediate applications of Theorem 3.1 for anisotropic mesh generation and refinement. For instance, there have been a number of studies on the so-called mesh quality measures, which quantifies the optimality of an anisotropic mesh for various considerations [4, 16]. The right hand side of estimate (21) is a natural "quality measure" if the interpolation error is the main concern. This measure decreases when the elements become more aligned with the orientation of $\nabla^{k+1}u$, their aspect ratios approach the anisotropic ratio of $\nabla^{k+1}u$, and the error over each element becomes more evenly distributed. Of course, these three geometric features are related, and the overall mesh quality depends on the interaction among them. A nice feature of the quality measure based on (21) would be that it has a clear geometric interpretation and it is an upper bound of the interpolation error itself, while many other measures are ad hoc and not directly related to the error. Another application of this work is in the moving mesh method or the r-refinement for solving partial differential equations. In the moving mesh method, it is required to define a monitor function, or the target mesh metric [7]. Interpolation error is one of the frequently used error indicators for defining the monitor functions. The optimal metric $M_{k+1,m,p}$ provided in Theorem 3.1 can be naturally used for such a purpose. These applications are currently under investigation and will be presented elsewhere.

There are a number of unresolved issues related to this work. First, the error estimate stated in Theorem 3.1 is optimal with respect to the order of N. But we are not clear if it is optimal with respect to the order of anisotropic ratio S_{k+1}. In the case of linear interpolation, it has been shown by Chen, Sun, and Xu [10] that the optimal error bound in L^p-norm is sharp for the class of functions with non-vanishing Hessians on general triangular elements. However, we are not clear whether a similar conclusion is true for the estimate of the $W^{m,p}$-error of Π_k with $m \geq 1$ or $k \geq 2$. Second, in Theorem 3.1 we only assume the triangulations are quasi-uniform under a given metric. Therefore, the conclusion holds for the case when the largest internal angles of the elements approach π. However, if we assume in addition that the triangulations satisfy the maximum angle condition [3], can we improve estimate (21) and select a metric leading to even smaller interpolation errors? There are some preliminary studies for this issue. For instance, for minimizing the H^1-seminorm of linear interpolation errors, if we restrict all triangles to be acute isosceles, then the optimal aspect ratio will be $(S_2)^2$, and the H^1-error can be a factor of S_2 smaller than that by using the aspect ratio close to S_2; see Section 4 of [8]. Shewchuk [20] observed the same phenomenon for general anisotropic triangles satisfying the maximum angle condition. He describes this type of case as "super-accuracy". However, for quadratic interpolations, it seems the optimal aspect ratio is always in the magnitude of S_3 according to the study of model problems reported in [9]. We are not clear if the so called super-accuracy phenomenon does exist for interpolations of degree $k \geq 2$ or not.

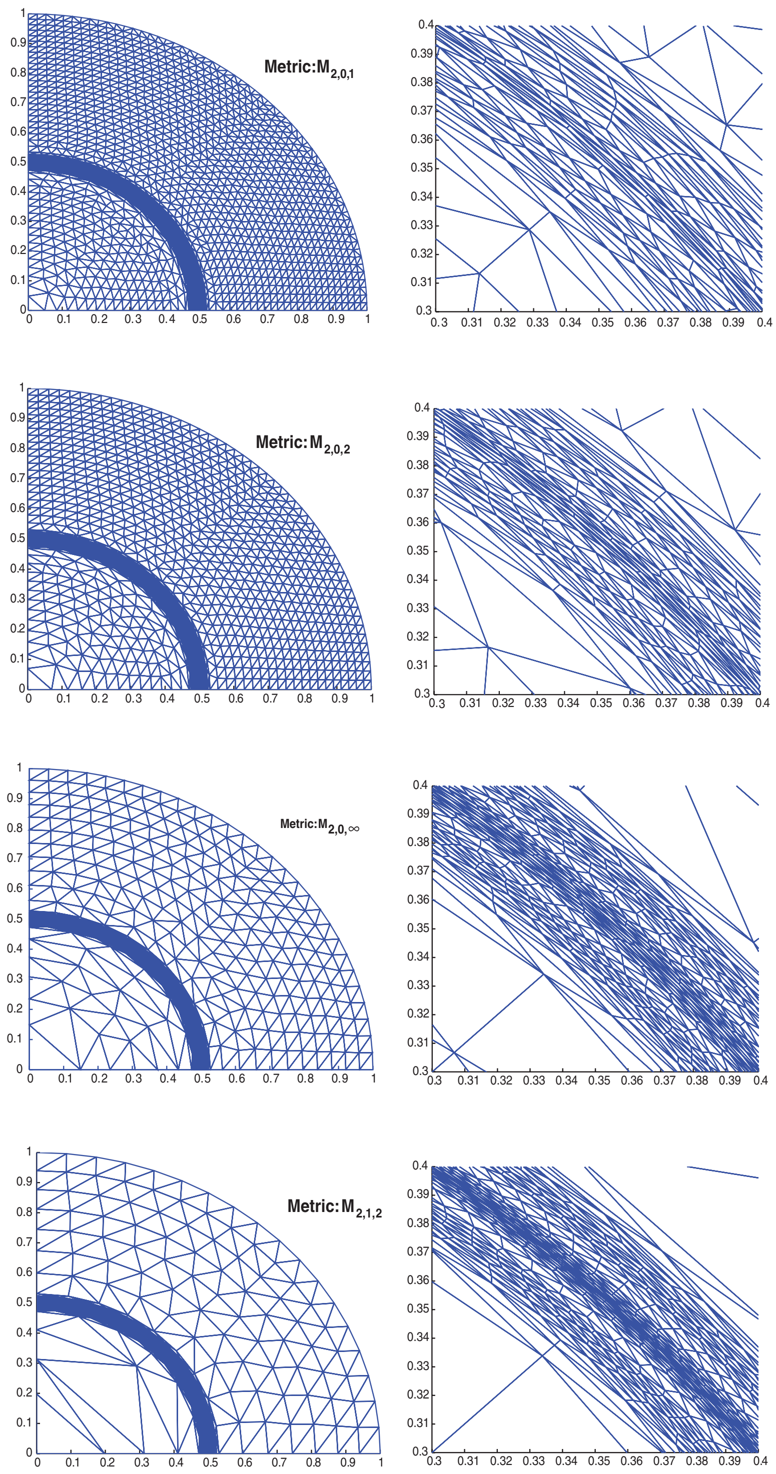

FIGURE 2. Triangulations that are generated according to different metrics $M_{2,m,p}$ and their close looks.

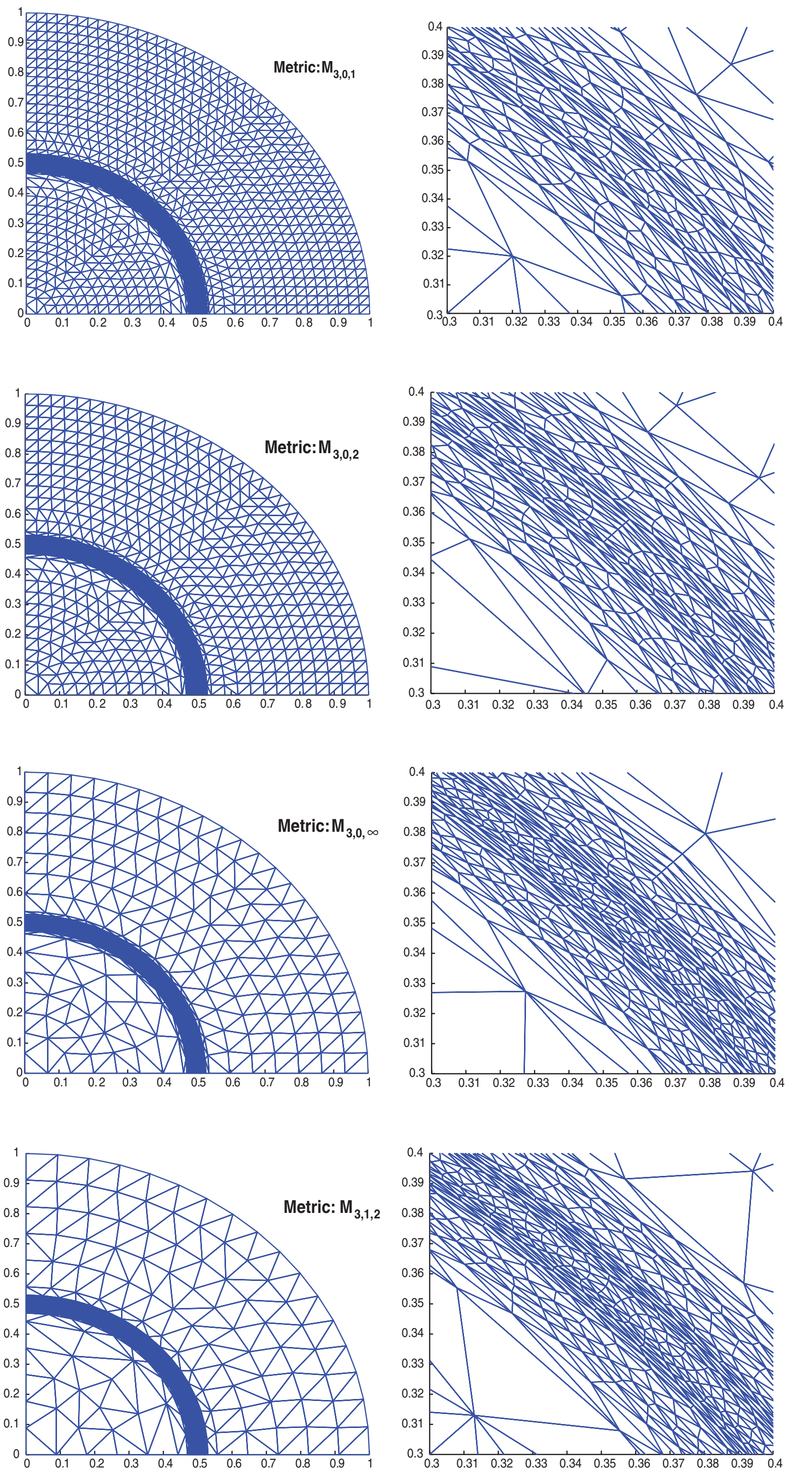

FIGURE 3. Triangulations that are generated according to different metrics $M_{3,m,p}$ and their close looks.

References

[1] D. Ait-Ali-Yahia, G. Baruzzi, W.G. Habashi, M. Fortin, J. Dompierre, and M. Vallet, Anisotropic mesh adaptation: towards user-independent, mesh-independent and solver-independent CFD. II. Structured grids. *Internat. J. Numer. Methods Fluids*, 39(2002), 657–673. MR1911881 (2003d:76101)

[2] T. Apel, *Anisotropic Finite Elements: Local Estimates and Applications*, Book series: Advances in Numerical Mathematics, Stuttgart, Teubner, 1999. MR1716824 (2000k:65002)

[3] I. Babuška and A. K. Aziz, On the angle condition in the finite element method, *SIAM J. Numer. Anal.* 13 (1976), 214-226. MR0455462 (56:13700)

[4] M. Berzins, A solution-based triangular and tetrahedral mesh quality indicators, *SIAM J. Sci. Comp.* 19(1998), 2051-2060. MR1638037 (99d:65341)

[5] H. Borouchaki, P.L. George, F. Hecht, P. Laug, and E. Saltel, Delaunay mesh generation governed by metric specifications. I. algorithms. *Finite Elem. Anal. Des.* 25(1997), 61-83. MR1442300 (98c:65163)

[6] H. Borouchaki, P.L. George, and B. Mohammadi, Delaunay mesh generation governed by metric specifications. II. applications. *Finite Elem. Anal. Des.* 25(1997), 85-109. MR1442301 (98c:65164)

[7] W. Cao, W. Huang, and R.D. Russell, A study of monitor functions for two-dimensional adaptive mesh generation, *SIAM J. Sci. Comput.* 20(1999), 1978-1994. MR1694650 (2000c:65113)

[8] W. Cao, On the error of linear interpolation and the orientation, aspect ratio, and internal angles of a triangle, *SIAM J. of Numer. Anal.* 43(2005), 19-40. MR2177954 (2006k:65023)

[9] W. Cao, Anisotropic measure of third order derivatives and the quadratic interpolation error on triangular elements, to appear in *SIAM J. Sci. Comput.*, 2007.

[10] L. Chen, P. Sun, and J. Xu, Optimal anisotropic meshes for minimizing interpolation errors in L^p-norm, *Math. of Comput.* 76(2007), 179-204. MR2261017

[11] L. Chen and J. Xu, Optimal Delaunay triangulations, *J. Comput. Math.* 22(2004), 299-308. MR2058939 (2005c:41042)

[12] P. G. Ciarlet, The Finite Element Methods for Elliptic Problems, SIAM Classics in Applied Mathematics, SIAM, Philadelphia, 2002. MR1930132

[13] E. F. D'Azevedo and R. B. Simpson, On optimal triangular meshes for minimizing the gradient error, *Numer. Math.* 59(1991), 321-348. MR1113194 (92c:65095)

[14] J. Dompierre, M. Vallet, Y. Bourgault, M. Fortin, and W.G. Habashi, Anisotropic mesh adaptation: towards user-independent, mesh-independent and solver-independent CFD. III. Unstructured meshes. *Internat. J. Numer. Methods Fluids* 39(2002), 675-702. MR1911882 (2003d:76105)

[15] W.G. Habashi, J. Dompierre, Y. Bourgault, D. Ait-Ali-Yahia, M. Fortin, and M. Vallet, Anisotropic mesh adaptation: towards user-independent, mesh-independent and solver-independent CFD. I. General principles. *Internat. J. Numer. Methods Fluids* 32(2000), 725-744. MR1911901 (2003d:76106)

[16] W. Huang, Measuring mesh qualities and application to variational mesh, *SIAM J. of Sci. Comput.*, 26(2005), 1643-1666. MR2142589 (2006b:65187)

[17] W. Huang and W. Sun, Variational mesh adaptation II: Error estimates and monitor functions, *J. Comput. Phys.* 184 (2003), 619-648. MR1959407

[18] E. J. Nadler, Piecewise linear approximation on triangulations of a planar region, Ph.D. Thesis, Division of Applied Mathematics, Brown University, Providence, RI, May 1985.

[19] S. Rippa, Long and thin triangles can be good for linear interpolation, *SIAM J. Numer. Anal.* 29 (1992), 257-270. MR1149097 (92j:65007)

[20] J. R. Shewchuk, What is a good linear finite element? Interpolation, conditioning, anisotropy, and quality measure. *preprint, Dept. of Electronic Engineering and Computer Sciences, University of California at Berkeley*, 2002.

[21] R. B. Simpson, Anisotropic mesh transformations and optimal error control, *Applied Numer. Math.* 14(1994), 183-198. MR1273824

[22] O. C. Zienkiewicz and J. Wu, Automatic directional refinement in adaptive analysis of compressible flows, *Int. J. Numer. Methods Engrg.* 37(1994), 2189-2210. MR1285070 (95c:76065)

Department of Mathematics, University of Texas at San Antonio, San Antonio, Texas 78249

E-mail address: wcao@math.utsa.edu