COMMUNICATIONS IN COMPUTATIONAL PHYSICS
Vol. **1**, No. 2, pp. 276-310

Commun. Comput. Phys.
April 2006

Mathematical Principles of Anisotropic Mesh Adaptation[†]

Weizhang Huang[*]

Department of Mathematics, University of Kansas, Lawrence, KS 66045, U.S.A.

Received 8 September 2005; Accepted (in revised version) 27 November 2005

Abstract. Mesh adaptation is studied from the mesh control point of view. Two principles, equidistribution and alignment, are obtained and found to be necessary and sufficient for a complete control of the size, shape, and orientation of mesh elements. A key component in these principles is the monitor function, a symmetric and positive definite matrix used for specifying the mesh information. A monitor function is defined based on interpolation error in a way with which an error bound is minimized on a mesh satisfying the equidistribution and alignment conditions. Algorithms for generating meshes satisfying the conditions are developed and two-dimensional numerical results are presented.

Key words: Mesh adaptation; anisotropic mesh; equidistribution; alignment; error analysis; finite element.

1 Introduction

Many partial differential equations (PDEs) arising from science and engineering have a common feature that they have a small portion of the physical domain where small node separations are required to resolve large solution variations. Examples include problems having boundary layers, shock waves, ignition fronts, and sharp interfaces in fluid dynamics, the combustion and heat transfer theory, and groundwater hydrodynamics. Numerical solution of these PDEs using a uniform mesh can be formidable when the systems involve more than two spatial dimensions since the number of mesh nodes required may become large. To improve efficiency and accuracy of numerical solution it is natural to put more mesh nodes in regions of large solution variation than the rest of the physical domain.

[*]Correspondence to: Weizhang Huang, Department of Mathematics, University of Kansas, Lawrence, KS 66045, U.S.A. Email: huang@math.ku.edu

[†]This work was supported in part by the NSF under grant DMS-0410545.

With this basic idea of mesh adaptation, the number of mesh nodes required can be much smaller; thus significant economies are gained.

Essential to mesh adaptation is the ability to control the *size*, *shape*, and *orientation* of mesh elements throughout the domain. Traditionally, research has been concentrated on isotropic mesh adaptation where mesh elements are adjusted only in size according to an error estimate or indicator while their shape is kept close to being equilateral; e.g., see books [2,15,27,52] and references therein. However, isotropic meshes often tend to use too many elements in regions of large solution error. This is especially true when problems have an anisotropic feature that the solution changes more significantly in one direction than the others. Full benefits of mesh adaptation can only be taken by simultaneously adjusting the size, shape, and orientation of mesh elements according to the behavior of the physical solution. This often results in an anisotropic mesh, a mesh having elements of large aspect ratio.

The well-known equidistribution principle [11, 20] has been playing an important role in mesh adaptation. It entails finding a mesh which evenly distributes an error density among the mesh elements. The principle has been serving as a guideline in developing mesh adaptation strategies, and most existing adaptive mesh algorithms are more or less related to it. Unfortunately, it is known [49] that the equidistribution principle is insufficient to determine an anisotropic mesh in multi-dimensions. Great effort has been made in the last decade to develop multi-dimensional generalizations of the equidistribution principle and/or other principles for anisotropic mesh adaptation; e.g., see [4, 19, 32, 37, 41, 42, 50]. Given a physical domain $\Omega \subset \Re^n$ $(n \geq 1)$, an adaptive mesh thereon can be generated as the image of a logical or computational mesh under a coordinate transformation $x = x(\xi) : \Omega_c \to \Omega$, where Ω_c is the computational domain artificially chosen for the purpose of mesh generation. Denote by $\boldsymbol{J} = (\partial x)/(\partial \xi)$ the Jacobian matrix of the coordinate transformation and $J = \det(\boldsymbol{J})$ its determinant. Motivated by a discrete constrained optimization problem, Steinberg and Roache [50] define $x = x(\xi)$ by minimizing the functional $\int_{\Omega_c} J^2 d\xi$ subject to the global implicit constraint $\int_{\Omega_c} J d\xi = |\Omega|$, with intention to keep element volume constant. These ideas of relating mesh adaptation functionals to equidistribution and using global implicit constraints are studied extensively by Knupp and Robidoux [42]. Upon studying linear interpolation error on triangular elements, D'Azevedo and Simpson [19] suggest that the coordinate transformation be chosen to minimize the gradient of interpolation error and thus to satisfy

$$\boldsymbol{J}^T H(v)^T H(v) \boldsymbol{J} = cI, \qquad \text{in } \Omega_c \tag{1.1}$$

where c is a constant and $H(v)$ denotes the Hessian of a function v. Huang and Sloan [37] choose the coordinate transformation such that the function, when transformed into the new coordinate, has the same change rate at every point and in every direction. This results in

$$\boldsymbol{J}^T \left(I + \nabla v \nabla v^T \right) \boldsymbol{J} = cI \qquad \text{in } \Omega_c \tag{1.2}$$

for some constant c. Huang [32] generalizes the ideas of [19, 37] to the case with an arbitrary $n \times n$ symmetric and positive definite matrix $M = M(x)$ (named *a monitor*

function) and requires the coordinate transformation to satisfy

$$\boldsymbol{J}^T M \boldsymbol{J} = \left(\frac{\sigma}{|\Omega_c|}\right)^{\frac{2}{n}} I, \qquad \text{in } \Omega_c, \tag{1.3}$$

where $\sigma = \int_\Omega \rho(x)dx$, $\rho = \sqrt{\det(M)}$, and $|\Omega_c|$ denotes the volume of Ω_c. Here, the constant $(\sigma/|\Omega_c|)^{2/n}$ results from compatibility and is obtained by taking determinant of (1.3) and integrating it over Ω_c. It is shown in [32] that equation (1.3) is equivalent to the conditions

$$J\rho = \frac{\sigma}{|\Omega_c|}, \tag{1.4}$$

$$\frac{1}{n}\text{tr}\left(\boldsymbol{J}^T M \boldsymbol{J}\right) = \det\left(\boldsymbol{J}^T M \boldsymbol{J}\right)^{\frac{1}{n}}. \tag{1.5}$$

As will be seen in §2, condition (1.4) is a multi-dimensional generalization of the equidistribution principle while (1.5) is an alignment condition characterizing the shape and orientation of mesh elements. A functional is constructed in [32] based on (1.4) and (1.5) for determining the coordinate transformation needed in variational mesh adaptation. Other works include Baines [4] and Knupp et al. [39,41]. Baines shows, using an algebraic identity, that the least squares minimization of the residual of the divergence of a vector field is equivalent to a least squares measure of equidistribution of the residual. Knupp et al. [39,41] define the coordinate transformation by specifying the inverse Jacobian matrix in the least squares sense:

$$\int_\Omega \|\boldsymbol{J}^{-1} - S(x)\|_F^2 dx \tag{1.6}$$

where $\|\cdot\|_F$ is the Frobenius matrix norm and $S = S(x)$ is the user-specified Jacobian matrix.

Despite this progress, the basic mathematical principles in anisotropic mesh adaptation are not fully understood. For instance, it is unclear how the existing conditions such as (1.4) and (1.5) and (1.6) are related to the control of mesh elements. It is neither clear if these conditions, particularly the specification of the Jacobian matrix, over-determine the coordinate transformation in the light of control of the size, shape, and orientation of mesh elements.

The objective of this paper is to present an in-depth mathematical study of anisotropic mesh adaptation from the mesh control point of view. Interestingly, using the singular value decomposition of the Jacobian matrix we re-discover the equidistribution and alignment conditions (1.4) and (1.5), or (2.2) and (2.3) and show that they are necessary and sufficient for a complete control of the size, shape, and orientation of mesh elements throughout the physical domain. A key component in these conditions is the monitor function prescribed by the user for specifying the mesh information. The monitor function can be defined based on error estimates and/or geometric and physical considerations. Indeed, it is defined in Section 3 in such a way that an interpolation error bound is minimized

on a mesh satisfying equidistribution and alignment conditions (1.4) and (1.5). It is also shown that the conditions can serve as guidelines for developing algorithms for generating adaptive anisotropic meshes.

An outline of the paper is as follows. In Section 2 the equidistribution and alignment conditions are derived and measures for how closely they are satisfied by an existing mesh are developed. In Section 3 the monitor function is defined based on interpolation error. Algorithms for generating adaptive meshes satisfying the the equidistribution and alignment conditions are studied in Section 4. Numerical results are presented in Section 5 for a selection of two-dimensional examples. Conclusions are drawn in Section 6.

2 Basic principles of mesh adaptation

2.1 Equidistribution and alignment

We consider mesh adaptation as a mathematical equivalent of the determination of a coordinate transformation. That is, an adaptive mesh is viewed as the image of a uniform reference mesh under a coordinate transformation $x = x(\xi)$ from the computational domain Ω_c to the physical domain Ω. In this continuous point of view mesh elements can be idealized as ellipsoids, for which the size, shape, and orientation can easily be defined. Indeed, for an ellipsoid the size is its volume, the shape is determined by the ratios between the lengths of the principal axes, and the orientation is controlled by the principal directions.

We study mesh control or control of the size, shape, and orientation of mesh elements from the continuous point of view. Our tool is the singular value decomposition (SVD) of a matrix. Let $\boldsymbol{J}$ be the Jacobian matrix of $x = x(\xi) : \Omega_c \to \Omega$, i.e., $\boldsymbol{J}(\xi) = \frac{\partial x}{\partial \xi}(\xi)$. By linearizing the coordinate transformation about an arbitrary point, $\xi_0 \in \Omega_c$, we have

$$x(\xi) = x(\xi_0) + \boldsymbol{J}(\xi_0)(\xi - \xi_0) + \mathcal{O}(|\xi - \xi_0|^2).$$

Obviously, the behavior of the coordinate transformation around ξ_0 is determined by $\boldsymbol{J}(\xi_0)$. Moreover, since a uniform mesh is used on Ω_c, the element size, shape, and orientation in a neighborhood of ξ_0 are fully determined by $\boldsymbol{J}(\xi_0)$. Denote the SVD of $\boldsymbol{J}(\xi_0)$ by

$$\boldsymbol{J}(\xi_0) = U\Sigma V^T,$$

where U and V are orthogonal matrices and $\Sigma = \mathrm{diag}(\sigma_1, ..., \sigma_n)$, with the σ_i's being the singular values of $\boldsymbol{J}(\xi_0)$. The mapping of a computational mesh element (e_c, a ball) into a physical element (e, an ellipsoid) under $x(\xi) = x(\xi_0) + \boldsymbol{J}(\xi_0)(\xi - \xi_0)$ and the roles of U, V, and Σ in the mapping are illustrated in Fig. 1. From the figure we can observe that the orientation of e is determined by the left singular vectors $U = [u_1, ..., u_n]$, its size and shape are controlled by the singular values $\Sigma = \mathrm{diag}(\sigma_1, ..., \sigma_n)$, and V plays no role in determining the size, shape, and orientation of e. Notice that u_i and σ_i^{-2} form the eigenvectors and eigenvalues of matrix $\boldsymbol{J}^{-T}\boldsymbol{J}^{-1}$, viz.,

$$\boldsymbol{J}^{-T}\boldsymbol{J}^{-1} = U\Sigma^{-2}U^T,$$

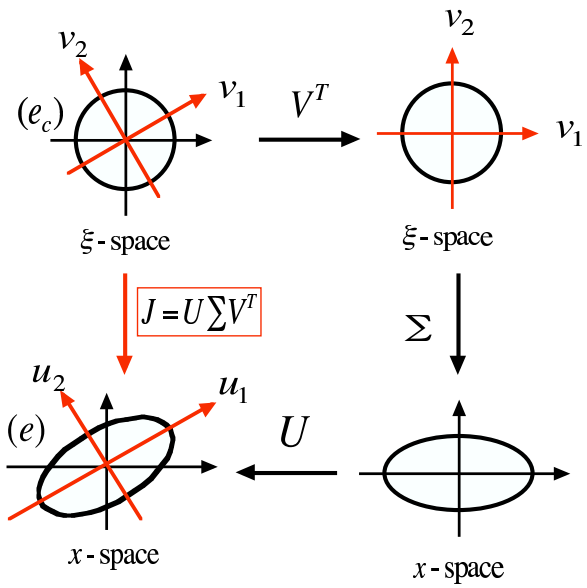

Figure 1: A computational element (e_c, a ball) is mapped into a physical mesh element (e, an ellipsoid) under $x = x(\xi_0) + \boldsymbol{J}(\xi_0)(\xi - \xi_0) : \Omega_c \to \Omega$. The figure shows the two-dimensional case with $U = [u_1, u_2]$ and $V = [v_1, v_2]$ and with $x(\xi_0)$ and ξ_0 being taken to be zero.

and their determination is equivalent to determination of the matrix $\boldsymbol{J}^{-T}\boldsymbol{J}^{-1}$. Thus, a complete control of the element can be realized by specifying the matrix. One choice is

$$\boldsymbol{J}^{-T}\boldsymbol{J}^{-1} = \left(\frac{\sigma}{|\Omega_c|}\right)^{-\frac{2}{n}} M(x), \tag{2.1}$$

where $M = M(x)$, the so-called *monitor function*, is a user-specified, $n \times n$ symmetric and positive definite matrix and

$$\sigma = \int_\Omega \rho(x)dx, \quad \rho = \sqrt{\det(M)}.$$

(Note that equation (2.1) is equivalent to equation (1.3).) We assume for the moment that $M(x)$ is given. It is shown in §3 that the monitor function can be chosen based on interpolation error estimates. The function $\rho = \rho(x)$ will hereafter be referred to as *the adaptation function*.

When the coordinate transformation satisfies relation (2.1), the element size, shape, and orientation are completely determined by $M = M(x)$ throughout the domain. To see this, we first take the determinant of both sides of (2.1). It follows that

$$J^{-2} = \left(\frac{\sigma}{|\Omega_c|}\right)^{-2} \rho^2$$

or

$$\rho J = \frac{\sigma}{|\Omega_c|}, \tag{2.2}$$

where $J = \det(\boldsymbol{J})$. Equation (2.2) is a multi-dimensional generalization of the well-known equidistribution principle [20]. It implies that J, characterizing the size of mesh elements, is small in the region of large ρ and is large in the region of small ρ.

Next, we rewrite (2.1) into equation (1.3). From it we can clearly see that relation (2.1) requires all of the eigenvalues of matrix $\boldsymbol{J}^T M\boldsymbol{J}$ to be the same and equal to the constant $(\sigma/|\Omega_c|)^{\frac{2}{n}}$. The requirement can be satisfied by asking the eigenvalues to be equal to each other and one of their products to be constant. The second condition is fulfilled in the equidistribution condition (2.2), upon observing that $(\rho J)^2 = \det(\boldsymbol{J}^T M\boldsymbol{J})$ is equal to the product of the eigenvalues. On the other hand, the first condition can be formulated as follows using the arithmetic-geometric mean inequality, which states that the arithmetic mean of any n positive numbers is greater than or equal to their geometric mean, with equality if and only if they are equal to each other. Noticing that the sum and product of the eigenvalues of a matrix are equal to its trace and determinant, respectively, we conclude that the equation

$$\frac{1}{n}\mathrm{tr}\left(\boldsymbol{J}^T M\boldsymbol{J}\right) = \det\left(\boldsymbol{J}^T M\boldsymbol{J}\right)^{\frac{1}{n}} \tag{2.3}$$

implies that the eigenvalues of $\boldsymbol{J}^T M\boldsymbol{J}$ are equal to each other. After some simple algebraic manipulation it can be shown that (2.3) is equivalent to

$$\boldsymbol{J}^{-T}\boldsymbol{J}^{-1} = \theta(x)M(x) \tag{2.4}$$

for a scalar but undetermined function $\theta = \theta(x)$. Equation (2.4) implies that the eigenvectors and the ratios between (or relative magnitude of) the eigenvalues of matrix $\boldsymbol{J}^{-T}\boldsymbol{J}^{-1}$ are determined by monitor function $M = M(x)$. Recalling that the eigenvectors and relative magnitude of the eigenvalues of matrix $\boldsymbol{J}^{-T}\boldsymbol{J}^{-1}$ characterize the shape and orientation of the mesh elements, we can conclude that the shape and orientation, or the alignment in short, of the mesh elements are controlled by the monitor function $M = M(x)$ through condition (2.3). For this reason, condition (2.3) will hereafter be referred to as *the alignment condition*.

In summary, we have seen that given a monitor function $M = M(x)$, *the equidistribution and alignment conditions* (2.2) *and* (2.3) *are necessary and sufficient for a complete control of the size, shape, and orientation of mesh elements throughout the physical domain.*

2.2 Measures of equidistribution and alignment

It is useful to know how closely the equidistribution and alignment conditions (2.2) and (2.3) are satisfied by an existing coordinate transformation for a given monitor function $M = M(x)$. We use the quality measures defined in [34]. Indeed, from (2.2) we define *the equidistribution measure* as

$$Q_{eq}(x) = \frac{\rho J|\Omega_c|}{\sigma}. \tag{2.5}$$

It can be deduced immediately from the definition that

$$Q_{eq}(x) > 0 \quad \text{and} \quad \frac{1}{|\Omega_c|}\int_{\Omega_c} Q_{eq} d\xi = 1.$$

Thus, $\max_x Q_{eq}(x) = 1$ if and only if the coordinate transformation satisfies the equidistribution condition (2.2). Moreover, the larger $\max_x Q_{eq}$ is, the more the function ρJ oscillates and in this sense, the farther the coordinate transformation is from satisfying (2.2).

On the other hand, from (2.3) we define *the alignment measure* as

$$Q_{ali}(x) = \left[\frac{\operatorname{tr}\left(\boldsymbol{J}^T M \boldsymbol{J}\right)}{n \det\left(\boldsymbol{J}^T M \boldsymbol{J}\right)^{\frac{1}{n}}}\right]^{\frac{n}{2(n-1)}}. \tag{2.6}$$

The arithmetic-geometric mean inequality implies

$$Q_{ali}(x) \geq 1,$$

with $Q_{ali}(x) = 1$ if and only if the alignment condition (2.3) is satisfied exactly. The larger $Q_{ali}(x)$ is, the more the eigenvalues of $\boldsymbol{J}^T M \boldsymbol{J}$ are different from each other and therefore the farther the alignment condition (2.3) is from being satisfied.

Note that the definition (2.6) is based on the Jacobian matrix $\boldsymbol{J}$. The alignment measure can also be defined based on its inverse, $\boldsymbol{J}^{-1}$, viz.,

$$\hat{Q}_{ali}(x) = \left[\frac{\operatorname{tr}\left(\boldsymbol{J}^{-1} M^{-1} \boldsymbol{J}^{-T}\right)}{n \det\left(\boldsymbol{J}^{-1} M^{-1} \boldsymbol{J}^{-T}\right)^{\frac{1}{n}}}\right]^{\frac{n}{2(n-1)}}. \tag{2.7}$$

Similarly,

$$\hat{Q}_{ali}(x) \geq 1,$$

and $\hat{Q}_{ali}(x) = 1$ if and only if the alignment condition (2.3) is satisfied exactly.

We observed that the quantities Q_{ali} and $\hat{Q}_{ali}$ and the alignment condition (2.3) are invariant if M is multiplied by a scalar function $\theta = \theta(x)$, i.e., $M \to \theta(x) M$. This invariance property is used in several places in the rest of the paper but without mentioning it explicitly.

In practice it is often useful to know how skewed the mesh elements are. This purely geometric property can be measured using the alignment measures defined above but with $M(x)$ being taken to be the identity matrix. This leads to *the geometric measures*,

$$Q_{geo}(x) = \left[\frac{\operatorname{tr}\left(\boldsymbol{J}^T \boldsymbol{J}\right)}{n \det\left(\boldsymbol{J}^T \boldsymbol{J}\right)^{\frac{1}{n}}}\right]^{\frac{n}{2(n-1)}}, \tag{2.8}$$

$$\hat{Q}_{geo}(x) = \left[\frac{\operatorname{tr}\left(\boldsymbol{J}^{-1} \boldsymbol{J}^{-T}\right)}{n \det\left(\boldsymbol{J}^{-1} \boldsymbol{J}^{-T}\right)^{\frac{1}{n}}}\right]^{\frac{n}{2(n-1)}}. \tag{2.9}$$

Like the alignment measures, they have the property

$$Q_{geo}(x) \geq 1, \quad \hat{Q}_{geo}(x) \geq 1,$$

and $Q_{geo}(x) = 1$ or $\hat{Q}_{geo}(x) = 1$ if and only if the mesh elements are equilateral. Moreover, it has been shown in [34] that they are equivalent to the aspect ratio of the mesh elements.

The relations between the alignment and geometric measures can be developed as follows. Denote the eigen-decomposition of M by

$$M = Q\text{diag}(\lambda_1, ..., \lambda_n)Q^T.$$

Let

$$(Q^T\boldsymbol{J})^T = [a_1, ..., a_n].$$

Then,

$$\text{tr}\left(\boldsymbol{J}^T M\boldsymbol{J}\right) = \text{tr}\left((Q^T\boldsymbol{J})^T\text{diag}(\lambda_1, ..., \lambda_n)(Q^T\boldsymbol{J})\right) = \sum_i \lambda_i\|a_i\|^2,$$

where $\|\cdot\|$ denotes the l_2 norm of a vector or a matrix. On the other hand,

$$\text{tr}\left(\boldsymbol{J}^T\boldsymbol{J}\right) = \text{tr}\left((Q^T\boldsymbol{J})^T(Q^T\boldsymbol{J})\right) = \sum_i \|a_i\|^2 = \sum_i \lambda_i^{-1}\left(\lambda_i\|a_i\|^2\right).$$

Thus,

$$\text{tr}\left(\boldsymbol{J}^T\boldsymbol{J}\right) \leq \max_i \lambda_i^{-1} \cdot \sum_i \lambda_i\|a_i\|^2 = \|M^{-1}\| \cdot \text{tr}\left(\boldsymbol{J}^T M\boldsymbol{J}\right),$$

and

$$\text{tr}\left(\boldsymbol{J}^T\boldsymbol{J}\right) \geq \min_i \lambda_i^{-1} \cdot \sum_i \lambda_i\|a_i\|^2 = \frac{1}{\max_i \lambda_i} \cdot \sum_i \lambda_i\|a_i\|^2 = \|M\|^{-1} \cdot \text{tr}\left(\boldsymbol{J}^T M\boldsymbol{J}\right).$$

Combining these results, we get

$$\|M\|^{-1} \cdot \text{tr}\left(\boldsymbol{J}^T M\boldsymbol{J}\right) \leq \text{tr}\left(\boldsymbol{J}^T\boldsymbol{J}\right) \leq \|M^{-1}\| \cdot \text{tr}\left(\boldsymbol{J}^T M\boldsymbol{J}\right).$$

From the definitions of the alignment and geometric measures it follows that

$$Q_{ali}^{\frac{2(n-1)}{n}} \frac{\|M\|^{-1}}{\rho^{-\frac{2}{n}}} \leq Q_{geo}^{\frac{2(n-1)}{n}} \leq Q_{ali}^{\frac{2(n-1)}{n}} \frac{\|M^{-1}\|}{\rho^{-\frac{2}{n}}}, \tag{2.10}$$

where we recall that $\rho = \sqrt{\det(M)}$. Similarly we have

$$\hat{Q}_{ali}^{\frac{2(n-1)}{n}} \frac{\|M^{-1}\|^{-1}}{\rho^{\frac{2}{n}}} \leq \hat{Q}_{geo}^{\frac{2(n-1)}{n}} \leq \hat{Q}_{ali}^{\frac{2(n-1)}{n}} \frac{\|M\|}{\rho^{\frac{2}{n}}}. \tag{2.11}$$

3 Monitor function

From the previous section we have seen that the monitor function $M = M(x)$ plays a role of controlling element size, shape, and orientation through the equidistribution and alignment conditions (2.2) and (2.3). Generally speaking, the monitor function can be defined based on error estimates, physical or geometric considerations; e.g. see [14, 23, 36, 43]. In this section we define the monitor function in a way that the interpolation error has a lower upper-bound on a coordinate transformation satisfying (2.2) and (2.3). The same approach has been employed in [34, 38], and the obtained results are similar, except for the anisotropic case where the gradient of the linear interpolation error is used. For this case, the present result requires that the mesh satisfy the equidistribution and alignment conditions (2.2) and (2.3) whereas those in [34, 38] require that the mesh satisfy the two conditions as well as the geometric condition (3.18). Generally speaking, a mesh cannot simultaneously satisfy all three conditions since the alignment condition is contradictory to the geometric one. Thus, the present result is an improvement over those in [34,38] for the anisotropic case with gradient error of linear interpolation.

3.1 Interpolation error estimates

In this subsection we describe the error estimates for polynomial preserving interpolation on simplicial elements. We begin by introducing some notation. Assume that an affine family of triangulations $\{\mathcal{T}_h\}$ is given on the physical domain Ω. Then for each element K, there exists an invertible affine mapping $F_K : \hat{K} \to K$ such that $K = F_K(\hat{K})$, where $\hat{K}$ is the reference element *chosen to be equilateral and have the unitary volume*. In the literature this type of element is often referred to as simplicial elements. The norm and semi-norm of the Sobolev space $W^{m,p}(K)$ are denoted by $\|\cdot\|_{W^{m,p}(K)}$ and $|\cdot|_{W^{m,p}(K)}$, respectively. The *scaled* semi-norm of $W^{m,p}(K)$ is defined as $\langle\cdot\rangle_{W^{m,p}(K)} \equiv |K|^{-1/p}|\cdot|_{W^{m,p}(K)}$, where $|K|$ is the volume of K. Obviously, $\langle v\rangle_{W^{m,p}(K)}$ is an L^p average of $v^{(m)}$ on K. Hereafter, C denotes a generic positive constant.

The element-wise error estimates in the following theorem are developed in [34] using the standard theory of interpolation for finite elements (e.g., see [17]).

Theorem 3.1. *Let $(\hat{K}, \hat{P}, \hat{\Sigma})$ be a finite element, where $\hat{K}$ is the reference element, $\hat{P}$ is a finite-dimensional linear space of functions defined on $\hat{K}$, and $\hat{\Sigma}$ is a set of degrees of freedom. Let s be the greatest order of partial derivatives occurring in $\hat{\Sigma}$. For some integers m, k, and l: $0 \le m \le l \le k+1$, and some numbers p, $q \in [1, \infty]$, if*

$$W^{l,p}(\hat{K}) \hookrightarrow C^s(\hat{K}), \tag{3.1}$$

$$W^{l,p}(\hat{K}) \hookrightarrow W^{m,q}(\hat{K}), \tag{3.2}$$

$$P_k(\hat{K}) \subset \hat{P} \subset W^{m,q}(\hat{K}), \tag{3.3}$$

then there exists a constant $C = C(\hat{K}, \hat{P}, \hat{\Sigma})$ such that, for all affine-equivalent finite

elements (K, P_K, Σ_K),

$$|v - \Pi_{k,K} v|_{W^{m,q}(K)} \leq C \|(F_K^{'})^{-1}\|^m \cdot |\det(F_K^{'})|^{\frac{1}{q}} \cdot \langle v \rangle_{W^{l,p}(K)} \tag{3.4}$$

for $v \in W^{l,p}(K)$,

$$|v - \Pi_{k,K} v|_{W^{m,q}(K)} \leq C \|(F_K^{'})^{-1}\|^m \cdot \|(F_K^{'})\|^l \cdot |\det(F_K^{'})|^{\frac{1}{q}} \cdot \left\langle \mathrm{tr}\left((F_K^{'})^T \nabla v \nabla v^T F_K^{'}\right) \right\rangle^{\frac{1}{2}}_{L^{\frac{p}{2}}(K)} \tag{3.5}$$

for $v \in W^{1,p}(K)$, *and*

$$\begin{aligned} & |v - \Pi_{k,K} v|_{W^{m,q}(K)} \\ \leq \; & C \|(F_K^{'})^{-1}\|^m \cdot \|F_K^{'}\|^{l-2} \cdot |\det(F_K^{'})|^{\frac{1}{q}} \cdot \left\langle \mathrm{tr}\left((F_K^{'})^T \left| H\left(D^{l-2} v\right)\right| F_K^{'}\right) \right\rangle_{L^p(K)} \end{aligned} \tag{3.6}$$

for $v \in W^{l,p}(K)$ *with* $l \geq 2$. *Here,* $\Pi_{k,K}$ *denotes the* k^{th}*-degree* P_K*-interpolation operator on* K; $\det(\cdot)$ *and* $\mathrm{tr}(\cdot)$ *denote the determinant and trace of a matrix, respectively;*

$$|H(D^{l-2} v)| \equiv \sum_{i_1, \ldots, i_{l-2}} \left| H\left(\frac{\partial^{l-2} v}{\partial x_{i_1} \cdots \partial x_{i_{l-2}}} \right) \right|; \tag{3.7}$$

$H(\cdot)$ *is the Hessian of a function; and* $|H(\cdot)| \equiv Q \mathrm{diag}(|\lambda_1|, ..., |\lambda_n|) Q^T$ *for a given eigen-decomposition* $H(\cdot) = Q\, \mathrm{diag}(\lambda_1, ..., \lambda_n)\, Q^T$.

Remark 3.1. The numbers l and p are related to the regularity of the considered functions. The theorem holds for k^{th} degree interpolation polynomials with $k \geq 0$ and thus for high order simplicial finite elements. Since the elements are affine, only the Jacobian of the map F_K, $F_K^{'}$, appears in the error estimates. The sufficient conditions for (3.1) and (3.2) can be derived from the Embedding Theorem for Sobolev spaces (e.g., see [1]). For the widely used case of Lagrange interpolation ($s = 0$), the conditions $0 \leq m \leq l \leq k+1$, $1 \leq q \leq p$, and $l > n/p$ for $p > 1$ or $l \geq n$ for $p = 1$ are sufficient for (3.1) and (3.2) to hold.

Remark 3.2. The estimate (3.4) is isotropic whereas the estimates (3.5) and (3.6) are anisotropic. This is because function derivatives are directly coupled in the latter two estimates with $F_K^{'}$, the Jacobian matrix which characterizes the size, shape, and orientation of K.

Remark 3.3. The estimate (3.6) still holds if (3.7) is replaced with a "smaller" matrix

$$|H(D^{l-2} v)| \equiv \bigcap_{i_1, \ldots, i_{l-2}} \left| H\left(\frac{\partial^{l-2} v}{\partial x_{i_1} \cdots \partial x_{i_{l-2}}} \right) \right|, \tag{3.8}$$

where $\cap$ denotes a certain matrix intersection operator. The purpose of introducing such an intersection operator is to approximate the biggest ellipsoid contained in the intersection

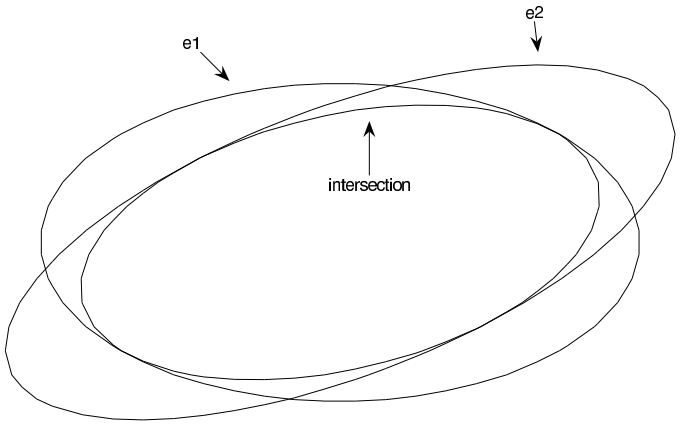

Figure 2: Illustration of the intersection of two ellipses.

of all the ellipsoids corresponding to the symmetric and semi-positive definite matrices $S_1, ..., S_m$. An approach is given in [16] to find an approximation of the biggest ellipsoid. We propose here a slightly different approach. We first consider two matrices S_1 and S_2, with S_1 being non-singular. In this case, there exists a non-singular matrix P such that $P^T S_1 P = I$ and $P^T S_2 P = \Sigma \equiv \text{diag}(\sigma_1, ..., \sigma_n)$. The intersection of S_1 and S_2 can be defined as

$$S_1 \cap S_2 = P^{-T} \text{diag}(\max\{1, \sigma_1\}, ..., \max\{1, \sigma_n\}) P^{-1}.$$

It is easy to show that the ellipsoid corresponding to matrix $S_1 \cap S_2$ is the biggest ellipsoid inscribed in the intersection of the ellipsoids corresponding to S_1 and S_2. The intersection is illustrated in Fig. 2. When both S_1 and S_2 are singular, the intersection can be defined by perturbating S_1 with a small positive number. For the case with more than two matrices, the intersection can be defined by adding a matrix each time. For instance, $S_1 \cap S_2 \cap S_3 = (S_1 \cap S_2) \cap S_3$. For more than two matrices, the ellipsoid corresponding to matrix $\bigcap_i S_i$ is contained in the intersection of the ellipsoids corresponding to $S_1, ..., S_m$ but it is unclear that the ellipsoid is the biggest inscribed one.

3.2 Continuous form of error estimates

We now proceed with the continuous form of the error estimates. A continuous form of the estimates is needed for defining the monitor function for the equidistribution and alignment conditions described in the previous section. We first define a global coordinate transformation. To this end, we assume that an affine, quasi-uniform mesh $\mathcal{T}_{c,h}$ can be defined on Ω such that it has the same connectivity as $\mathcal{T}_h$ does. In the following, the domain Ω will be viewed as the "computational" domain, i.e., $\Omega_c \equiv \Omega$, when associated with $\mathcal{T}_{c,h}$. The corresponding coordinate will be denoted by ξ. Then, a piecewise linear,

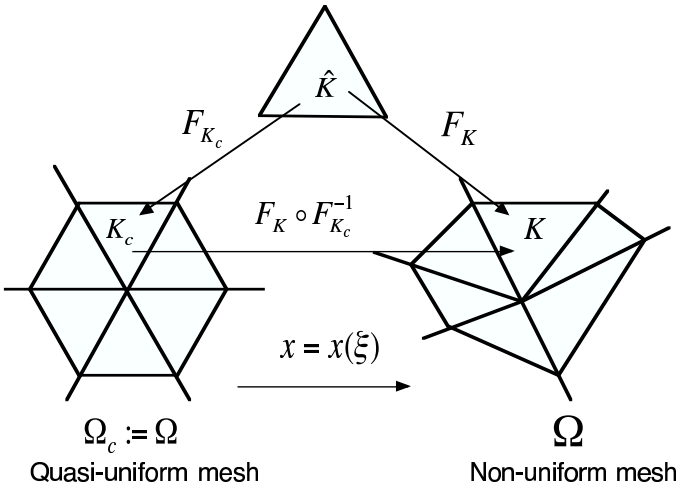

Figure 3: The definition of the piecewise linear coordinate transformation $x = x(\xi)$ is illustrated.

global coordinate transformation $x = x(\xi) : \Omega_c \to \Omega$ can be defined as

$$x(\xi) := F_K(F_{K_c}^{-1}(\xi)), \quad \forall \xi \in K_c, \forall K_c \in \mathcal{T}_{c,h}, \tag{3.9}$$

where K and K_c are the corresponding elements on Ω and Ω_c and $F_K : \hat{K} \to K$ and $F_{K_c} : \hat{K}_c \to K$ are linear mappings. The definition is illustrated in Fig. 3.

The assumption that $\mathcal{T}_{c,h}$ is quasi-uniform and has the same connectivity as $\mathcal{T}_h$ implies

$$|\det(F_{K_c}^{'})| = \mathcal{O}\left(N^{-1}\right), \quad \|F_{K_c}^{'}\| = \mathcal{O}\left(N^{-\frac{1}{n}}\right), \quad \forall K_c \in \mathcal{T}_{c,h}$$

where N is the number of elements in $\mathcal{T}_h$. From the definition of $x = x(\xi)$, we have

$$F_K^{'} = \boldsymbol{J} \cdot F_{K_c}^{'}, \quad |K| = \mathcal{O}\left(N^{-1} J\right), \quad \|F_K^{'}\| = \mathcal{O}\left(N^{-\frac{1}{n}} \|\boldsymbol{J}\|\right), \tag{3.10}$$

where $\boldsymbol{J} = (\partial x)/(\partial \xi)$ and $J = \det(\boldsymbol{J})$.

Using these relations, taking the q^{th} power on both sides of (3.4), (3.5), and (3.6), and summing them over all the elements, we obtain

$$|v - \Pi_k v|_{W^{m,q}(\Omega)}^q \le C N^{-\frac{(l-m)q}{n}} \sum_K |K| \cdot \|\boldsymbol{J}^{-1}\|^{mq} \cdot \|\boldsymbol{J}\|^{lq} \cdot \left(\frac{1}{|K|} \int_K \|D^l v\|_p^p dx\right)^{\frac{q}{p}} \tag{3.11}$$

for $v \in W^{l,p}(\Omega)$,

$$|v - \Pi_k v|_{W^{m,q}(\Omega)}^q \le C N^{-\frac{(1-m)q}{n}} \sum_K |K| \cdot \|\boldsymbol{J}^{-1}\|^{mq} \cdot \left(\frac{1}{|K|} \int_K \left(\mathrm{tr}\left(\boldsymbol{J}^T \nabla v \nabla v^T \boldsymbol{J}\right)\right)^{\frac{p}{2}} dx\right)^{\frac{q}{p}} \tag{3.12}$$

for $v \in W^{1,p}(\Omega)$, and

$$|v - \Pi_k v|^q_{W^{m,q}(\Omega)} \leq CN^{-\frac{(l-m)q}{n}} \sum_K |K| \cdot \|\boldsymbol{J}^{-1}\|^{mq} \cdot \|\boldsymbol{J}\|^{(l-2)q} \times \left(\frac{1}{|K|}\int_K \left(\mathrm{tr}\left(\boldsymbol{J}^T \left|H\left(D^{l-2}v\right)\right| \boldsymbol{J}\right)\right)^p dx\right)^{\frac{q}{p}} \tag{3.13}$$

for $v \in W^{l,p}(\Omega)$ with $l \geq 2$. In (3.11), $\|\cdot\|_p$ denotes the l_p matrix norm.

Then one can easily see that the right-hand-side term of (3.11) has the limit

$$CN^{-\frac{(l-m)q}{n}} \int_\Omega \|\boldsymbol{J}^{-1}\|^{mq} \cdot \|\boldsymbol{J}\|^{lq} \cdot \|D^l v\|^q_p dx$$

as $\max_K \mathrm{diam}(K) \to 0$. For notational simplicity, this asymptotical bound is denoted by

$$|v - \Pi_k v|^q_{W^{m,q}(\Omega)} \overset{<}{\to} CN^{-\frac{(l-m)q}{n}} \int_\Omega \|\boldsymbol{J}^{-1}\|^{mq} \cdot \|\boldsymbol{J}\|^{lq} \cdot \|D^l v\|^q_p dx \tag{3.14}$$

for $v \in W^{l,p}(\Omega)$. Similarly, from (3.12) and (3.13) we have

$$|v - \Pi_k v|^q_{W^{m,q}(\Omega)} \overset{<}{\to} CN^{-\frac{(1-m)q}{n}} \int_\Omega \|\boldsymbol{J}^{-1}\|^{mq} \left(\mathrm{tr}\left(\boldsymbol{J}^T \nabla v \nabla v^T \boldsymbol{J}\right)\right)^{\frac{q}{2}} dx \tag{3.15}$$

for $v \in W^{1,p}(\Omega)$ and

$$|v - \Pi_k v|^q_{W^{m,q}(\Omega)} \overset{<}{\to} CN^{-\frac{(l-m)q}{n}} \int_\Omega \|\boldsymbol{J}^{-1}\|^{mq} \cdot \|\boldsymbol{J}\|^{(l-2)q} \cdot \left(\mathrm{tr}\left(\boldsymbol{J}^T \left|H\left(D^{l-2}v\right)\right| \boldsymbol{J}\right)\right)^q dx \tag{3.16}$$

for $v \in W^{l,p}(\Omega)$ with $l \geq 2$.

The estimates (3.14), (3.15), and (3.16) have been obtained for the piecewise linear coordinate transformation $x = x(\xi)$ defined in (3.9) over simplicial mesh elements. Strictly speaking, these results can only be used for generation of meshes with simplicial elements. Nevertheless, we note that the bounds given in (3.14), (3.15), and (3.16) are asymptotic in the limit of $\max_K \mathrm{diam}(K) \to 0$. It is reasonable to expect that these asymptotic bounds also hold for other types of elements and coordinate transformations. For this reason, the monitor functions defined based on (3.14), (3.15), and (3.16) will be used for generating all types of adaptive mesh.

3.3 Monitor function

The monitor function is defined for three cases, the isotropic case and the anisotropic cases with $l = 1$ and $l = 2$.

3.3.1 The isotropic case

In this case, the monitor function $M = M(x)$ is defined based on the estimate (3.14). To be able to define a positive M, we regularize (3.14) with a to-be-determined parameter $\alpha_{iso} > 0$; viz.,

$$\begin{aligned} & |v - \Pi_k v|^q_{W^{m,q}(\Omega)} \\ \stackrel{<}{\to} \; & CN^{-\frac{(l-m)q}{n}} \int_\Omega \|\boldsymbol{J}^{-1}\|^{mq} \cdot \|\boldsymbol{J}\|^{lq} \cdot \left(\alpha_{iso} + \|D^l v\|_p\right)^q dx \\ = \; & CN^{-\frac{(l-m)q}{n}} \alpha^q_{iso} \int_\Omega \|\boldsymbol{J}^{-1}\|^{mq} \cdot \|\boldsymbol{J}\|^{lq} \cdot \left(1 + \frac{1}{\alpha_{iso}}\|D^l v\|_p\right)^q dx. \end{aligned}$$

From the definitions of the geometric measures, (2.8) and (2.9) it follows that

$$\begin{aligned} \|\boldsymbol{J}\| \le \|\boldsymbol{J}\|_F &= \sqrt{\mathrm{tr}(\boldsymbol{J}^T\boldsymbol{J})} = \sqrt{n} Q_{geo}^{\frac{n-1}{n}} \cdot |J|^{\frac{1}{n}}, \\ \|\boldsymbol{J}^{-1}\| \le \|\boldsymbol{J}^{-1}\|_F &= \sqrt{\mathrm{tr}(\boldsymbol{J}^{-1}\boldsymbol{J}^{-T})} = \sqrt{n} \hat{Q}_{geo}^{\frac{n-1}{n}} \cdot |J|^{-\frac{1}{n}}. \end{aligned}$$

Inserting these results into the above estimate, we get

$$\begin{aligned} & |v - \Pi_k v|^q_{W^{m,q}(\Omega)} \\ & \quad \stackrel{<}{\to} CN^{-\frac{(l-m)q}{n}} \alpha^q_{iso} \int_\Omega \hat{Q}_{geo}^{\frac{mq(n-1)}{n}} \cdot Q_{geo}^{\frac{lq(n-1)}{n}} \cdot |J|^{\frac{q(l-m)}{n}} \cdot \left(1 + \frac{1}{\alpha_{iso}}\|D^l v\|_p\right)^q dx. \end{aligned} \tag{3.17}$$

The monitor function is defined such that the bound given in (3.17) *is made as low as possible on a coordinate transformation satisfying equidistribution and alignment conditions* (2.2) *and* (2.3). To this end, we note that the geometric measures Q_{geo} (≥ 1) and $\hat{Q}_{geo}$ (≥ 1) are involved in the bound given in (3.17). To make the bound as low as possible, we should require that a coordinate transformation satisfying (2.2) and (2.3) also satisfy $Q_{geo} = 1$ and $\hat{Q}_{geo} = 1$; or equivalently,

$$\frac{1}{n}\mathrm{tr}\left(\boldsymbol{J}^T\boldsymbol{J}\right) = \det\left(\boldsymbol{J}^T\boldsymbol{J}\right)^{\frac{1}{n}}. \tag{3.18}$$

Comparing this with the alignment condition (2.3) yields

$$M = \theta(x) I \tag{3.19}$$

for some scalar function $\theta = \theta(x)$. The error bound becomes

$$|v - \Pi_k v|^q_{W^{m,q}(\Omega)} \stackrel{<}{\to} CN^{-\frac{(l-m)q}{n}} \alpha^q_{iso} \int_\Omega |J|^{\frac{q(l-m)}{n}} \cdot \left(1 + \frac{1}{\alpha_{iso}}\|D^l v\|_p\right)^q dx. \tag{3.20}$$

Next we determine the monitor function in the form (3.19) through the adaptation function $\rho = \sqrt{\det(M)}$. The tool is the following theorem on the optimality of equidistributing meshes (or coordinate transformations).

Theorem 3.2. *Given a real number $s > 0$ and a positive function $\rho = \rho(x)$ defined on Ω, then*

$$I[\xi] \equiv \int_\Omega \rho(|J|\rho)^s dx \geq \sigma \left(\frac{\sigma}{|\Omega_c|}\right)^s, \tag{3.21}$$

where $\sigma = \int_\Omega \rho dx$, for all invertible coordinate transformations $\xi = \xi(x) : \Omega \to \Omega_c$. The lower bound is attained on a mesh satisfying equidistribution condition (2.2).

Proof. Inequality (3.21) follows from the inequality

$$\left(\int_\Omega \frac{\rho}{\sigma}(|J|\rho)^s dx\right)^{\frac{1}{s}} \geq \left(\int_\Omega \frac{\rho}{\sigma}(|J|\rho)^{-1} dx\right)^{-1} = \frac{\sigma}{|\Omega_c|}$$

for any $s > 0$. It is easy to verify that a mesh satisfying the equidistribution condition (2.2) gives the lower bound. ■

From Theorem 3.2, it is clear that the right-hand-side term of (3.20) attains its lowest bound on an equidistributing mesh if the adaptation function $\rho = \rho(x)$ is defined such that

$$|J|^{\frac{q(l-m)}{n}} \cdot \left(1 + \frac{1}{\alpha_{iso}}\|D^l v\|_p\right)^q = \rho(|J|\rho)^s$$

for some number $s > 0$. Comparing the exponents of $|J|$ on both sides, we obtain $s = q(l-m)/n$ and

$$\rho = \rho_{iso} \equiv \left(1 + \frac{1}{\alpha_{iso}}\|D^l v\|_p\right)^{\frac{nq}{n+q(l-m)}}. \tag{3.22}$$

From (3.19) and the definition $\rho = \sqrt{\det(M)}$, the monitor function is determined as

$$M = M_{iso} \equiv \left(1 + \frac{1}{\alpha_{iso}}\|D^l v\|_p\right)^{\frac{2q}{n+q(l-m)}} I. \tag{3.23}$$

It remains to define α_{iso}, which is often referred to as the intensity parameter in the context of mesh adaptation since it controls the intensity of mesh concentration. It is suggested in [32] that α_{iso} be chosen such that (i) the monitor function M_{iso} is invariant under the scaling transformation of v and (ii) $\sigma \equiv \int_\Omega \rho_{iso} dx \leq C$ for some constant C. For the current situation,

$$\begin{aligned}
\sigma &\equiv \int_\Omega \rho_{iso}(x) dx \\
&\leq C_1 \int_\Omega \left[1 + \alpha_{iso}^{-\frac{nq}{n+q(l-m)}} \|D^l v\|_p^{\frac{nq}{n+q(l-m)}}\right] dx \\
&= C_1 \left[|\Omega| + \alpha_{iso}^{-\frac{nq}{n+q(l-m)}} \int_\Omega \|D^l v\|_p^{\frac{nq}{n+q(l-m)}} dx\right].
\end{aligned}$$

By choosing

$$\alpha_{iso} = \left[\frac{1}{|\Omega|}\int_\Omega \|D^l v\|_p^{\frac{nq}{n+q(l-m)}}\, dx\right]^{\frac{n+q(l-m)}{nq}}, \tag{3.24}$$

we have $\sigma \le 2C_1|\Omega|$ and $\rho_{iso}(x)$ and $M_{iso}(x)$ are invariant under the scaling transformation of v. Moreover, it is shown in [38] that this choice of α_{iso} concentrates about 50% of the mesh points in the regions of large ρ_{iso}. The mesh concentration can be adjusted by modifying the definition into

$$\alpha_{iso} = \left[\frac{(1-\beta)}{\beta\,|\Omega|}\int_\Omega \|D^l v\|_p^{\frac{nq}{n+q(l-m)}}\, dx\right]^{\frac{n+q(l-m)}{nq}}, \tag{3.25}$$

where $\beta \in (0,1)$ indicates the percentage of the mesh points concentrated in the regions of large ρ_{iso}. Numerical experience shows that $\beta \in [0.5, 0.8]$ works well for most cases.

For the monitor function (3.23) with α_{iso} given in (3.25), Theorem 3.2 and inequality (3.21) imply that the interpolation error has a bound on a mesh satisfying equidistribution and alignment conditions (2.2) and (2.3) as

$$|v - \Pi_k v|_{W^{m,q}(\Omega)} \overset{<}{\to} CN^{-\frac{(l-m)}{n}}\, |v|_{W^{l,\frac{nq}{n+q(l-m)}}(\Omega)}, \tag{3.26}$$

where $\sigma \le C$ and (3.25) have been used.

On a general mesh, from the definitions for Q_{eq} (2.5) and ρ (3.22) we can rewrite (3.17) as

$$|v - \Pi_k v|_{W^{m,q}(\Omega)} \overset{<}{\to} C\,N^{-\frac{(l-m)}{n}}\, Q_{mesh,iso}\, |v|_{W^{l,\frac{nq}{n+q(l-m)}}(\Omega)}, \tag{3.27}$$

where the overall mesh quality measure, $Q_{mesh,iso}$, is defined by

$$Q_{mesh,iso} = \left[\frac{1}{\sigma}\int_\Omega \left(\hat{Q}_{geo}^{\frac{m(n-1)}{n}} \cdot Q_{geo}^{\frac{l(n-1)}{n}} \cdot Q_{eq}^{\frac{(l-m)}{n}}\right)^q \rho_{iso} dx\right]^{\frac{1}{q}}, \tag{3.28}$$

which is the weighted L^q norm (with weight function ρ_{iso}) of $\hat{Q}_{geo}^{\frac{m(n-1)}{n}}\, Q_{geo}^{\frac{l(n-1)}{n}}\, Q_{eq}^{\frac{(l-m)}{n}}$.

3.3.2 The anisotropic case with $l = 1$

For this case, conditions (3.1) and (3.2) require that $s = 0$ (with s being the highest order of derivatives appearing in the interpolation) and $p > n$. Moreover, it is common practice to use $m = 0$ for $l = 1$. For this reason, we restrict our attention to the situation $m = 0$ in this subsection.

The monitor function is defined based on estimate (3.15) with $m = 0$, i.e.,

$$\|v - \Pi_k v\|_{L^q(\Omega)}^q \overset{<}{\to} CN^{-\frac{q}{n}} \int_\Omega \left(\mathrm{tr}\left(\boldsymbol{J}^T \nabla v \nabla v^T \boldsymbol{J}\right)\right)^{\frac{q}{2}} dx, \quad \forall v \in W^{1,p}(\Omega).$$

This is regularized with an intensity parameter $\alpha_{ani,1} > 0$ as

$$\|v-\Pi_k v\|^q_{L^q(\Omega)} \overset{<}{\to} CN^{-\frac{q}{n}}\alpha^q_{ani,1}\int_\Omega \left(\mathrm{tr}\left(\boldsymbol{J}^T\left[I+\frac{1}{\alpha^2_{ani,1}}\nabla v\nabla v^T\right]\boldsymbol{J}\right)\right)^{\frac{q}{2}}dx, \quad \forall v\in W^{1,p}(\Omega).$$

The derivations of the monitor function M and the intensity parameter $\alpha_{ani,1}$ for the current case are similar to the isotropic case in the previous subsection, except that we now require that a mesh satisfying the equidistribution and alignment conditions (2.2) and (2.3) also satisfy

$$\frac{1}{n}\mathrm{tr}\left(\boldsymbol{J}^T\left[I+\frac{1}{\alpha^2_{ani,1}}\nabla v\nabla v^T\right]\boldsymbol{J}\right) = \det\left(\boldsymbol{J}^T\left[I+\frac{1}{\alpha^2_{ani,1}}\nabla v\nabla v^T\right]\boldsymbol{J}\right)^{\frac{1}{n}}, \tag{3.29}$$

instead of the geometric condition (3.18). The results are recorded as follows:

$$\rho = \rho_{ani,1} \equiv \left(1+\frac{1}{\alpha^2_{ani,1}}\|\nabla v\|^2\right)^{\frac{q}{2(n+q)}}, \tag{3.30}$$

$$M = M_{ani,1} \equiv \left(1+\frac{1}{\alpha^2_{ani,1}}\|\nabla v\|^2\right)^{-\frac{1}{n+q}}\left[I+\frac{1}{\alpha^2_{ani,1}}\nabla v\nabla v^T\right], \tag{3.31}$$

$$\alpha_{ani,1} = \left[\frac{(1-\beta)}{\beta\,|\Omega|}\int_\Omega \|\nabla v\|^{\frac{q}{n+q}}dx\right]^{\frac{n+q}{q}}, \tag{3.32}$$

where β indicates roughly the percentage of the mesh points concentrated in the regions of large $\rho_{ani,1}$. It is recommended that β be chosen from $[0.5, 0.8]$. On a general mesh the interpolation error has a bound given by

$$\|v-\Pi_k v\|_{L^q(\Omega)} \overset{<}{\to} CN^{-\frac{1}{n}}Q_{mesh,ani,1}\cdot|v|_{W^{1,\frac{q}{n+q}}(\Omega)}, \quad \forall v\in W^{1,p}(\Omega), \tag{3.33}$$

where the overall mesh quality measure is defined by

$$Q_{mesh,ani,1} = \left[\frac{1}{\sigma}\int_\Omega\left(Q_{ali}^{\frac{n-1}{n}}\cdot Q_{eq}^{\frac{1}{n}}\right)^q\rho_{ani,1}dx\right]^{\frac{1}{q}}. \tag{3.34}$$

It is noted that $Q_{mesh,ani,1} = 1$ on a mesh satisfying the equidistribution and alignment conditions (2.2) and (2.3).

3.3.3 The anisotropic case with $l = 2$

For simplicity we consider in this subsection only the case $l = 2$. The procedure can straightforwardly be used for the general case with $l \geq 2$.

The monitor function is defined based on the estimate (3.16) with $l = 2$, i.e.,

$$|v - \Pi_k v|^q_{W^{m,q}(\Omega)} \overset{<}{\to} CN^{-\frac{(2-m)q}{n}} \int_\Omega \|\boldsymbol{J}^{-1}\|^{mq} \cdot \left(\mathrm{tr}\left(\boldsymbol{J}^T |H(v)| \boldsymbol{J}\right)\right)^q dx$$

for $v \in W^{2,p}(\Omega)$. The regularized form with an intensity parameter $\alpha_{ani,2} > 0$ reads

$$\begin{aligned} &|v - \Pi_k v|^q_{W^{m,q}(\Omega)} \\ &\overset{<}{\to} CN^{-\frac{(2-m)q}{n}} \alpha^q_{ani,2} \int_\Omega \|\boldsymbol{J}^{-1}\|^{mq} \cdot \left(\mathrm{tr}\left(\boldsymbol{J}^T \left[I + \frac{1}{\alpha_{ani,2}} |H(v)|\right] \boldsymbol{J}\right)\right)^q dx. \end{aligned} \tag{3.35}$$

Note that for the current case, m can take the values $m = 0$ and $m = 1$. When $m = 1$, the error bound given above involves $\|\boldsymbol{J}^{-1}\|^{mq}$ and the term containing the matrix trace. As we have seen in the previous subsections, the former is associated with the geometric measure or condition (3.18) while the latter is linked to the alignment measure or

$$\frac{1}{n}\mathrm{tr}\left(\boldsymbol{J}^T \left[I + \frac{1}{\alpha^2_{ani,1}} \nabla v \nabla v^T\right] \boldsymbol{J}\right) = \det\left(\boldsymbol{J}^T \left[I + \frac{1}{\alpha_{ani,2}} |H(v)|\right] \boldsymbol{J}\right)^{\frac{1}{n}} \tag{3.36}$$

(Also see condition (3.29)). Since a mesh cannot simultaneously satisfy both the geometric condition (3.18) and the alignment condition (3.36), we define the monitor function by requiring that a mesh satisfying the equidistribution and alignment conditions (2.2) and (2.3) also satisfy (3.36). Comparison of (3.36) with (2.3) gives

$$M = \theta(x) \left[I + \frac{1}{\alpha_{ani,2}} |H(v)|\right] \tag{3.37}$$

for some scalar function $\theta = \theta(x)$. With the so-defined monitor function, definition of Q_{ali} in Eq. (2.6) yields

$$\mathrm{tr}\left(\boldsymbol{J}^T \left[I + \frac{1}{\alpha^2_{ani,1}} \nabla v \nabla v^T\right] \boldsymbol{J}\right) = Q_{ali}^{\frac{2(n-1)}{n}} \, n \, J^{\frac{2}{n}} \det\left(I + \frac{1}{\alpha_{ani,2}} |H(v)|\right)^{\frac{1}{n}}. \tag{3.38}$$

It is noted that when $m > 0$, the factor $\|\boldsymbol{J}^{-1}\|$ appears in the bound (3.35). This factor is bounded in [34] using the geometric quality measure Q_{geo} (see Eq. (2.8)); namely,

$$\|\boldsymbol{J}^{-1}\| \le \|\boldsymbol{J}^{-1}\|_F = Q_{geo}^{\frac{(n-1)}{n}} \sqrt{n} \, J^{\frac{1}{n}}.$$

The final bound for the interpolation error then involves both Q_{ali} and Q_{geo}, whose minimization will require that the mesh satisfy the alignment and geometric conditions (2.3) and (3.18). Generally speaking, such a mesh does not exist since the conditions are contradictory to each other. To avoid this difficulty, we bound $\|\boldsymbol{J}^{-1}\|$ with the alignment

measure (instead of the geometric measure). For a monitor function in the form (3.37), from (2.11) it is not difficult to get

$$\hat{Q}_{geo}^{\frac{2(n-1)}{n}} \le \hat{Q}_{ali}^{\frac{2(n-1)}{n}} \frac{\left\| I + \frac{1}{\alpha_{ani,2}} |H(v)| \right\|}{\det\left(I + \frac{1}{\alpha_{ani,2}} |H(v)| \right)^{\frac{1}{n}}}.$$

Hence,

$$\begin{aligned} \|\boldsymbol{J}^{-1}\| &\le \|\boldsymbol{J}^{-1}\|_F = \sqrt{\mathrm{tr}(\boldsymbol{J}^{-1}\boldsymbol{J}^{-T})} \\ &= \sqrt{n}\hat{Q}_{geo}^{\frac{n-1}{n}} \cdot |J|^{-\frac{1}{n}} \\ &\le \sqrt{n}\,|J|^{-\frac{1}{n}} \hat{Q}_{ali}^{\frac{n-1}{n}} \frac{\left\| I + \frac{1}{\alpha_{ani,2}} |H(v)| \right\|^{\frac{1}{2}}}{\det\left(I + \frac{1}{\alpha_{ani,2}} |H(v)| \right)^{\frac{1}{2n}}}. \end{aligned} \tag{3.39}$$

Thus, from (3.39), the definition of the alignment measure, and the error bound (3.35) we have

$$\begin{aligned} |v - \Pi_k v|^q_{W^{m,q}(\Omega)} \overset{<}{\sim} CN^{-\frac{(2-m)q}{n}} \alpha_{ani,2}^q \int_\Omega \hat{Q}_{ali}^{\frac{mq(n-1)}{n}} \cdot Q_{ali}^{\frac{2q(n-1)}{n}} \cdot |J|^{\frac{(2-m)q}{n}} \times \\ \left\| I + \frac{1}{\alpha_{ani,2}} |H(v)| \right\|^{\frac{mq}{2}} \det\left(I + \frac{1}{\alpha_{ani,2}} |H(v)| \right)^{\frac{(2-m)q}{2n}} dx. \end{aligned} \tag{3.40}$$

This error bound is greater than those obtained in [34, 35] for the case $m = 1$, but it has the advantage that the effect of element skewness has been taken into account in (3.40) through inequality (3.39). The procedure used in the previous subsections to define M can straightforwardly be applied to (3.40). We obtain

$$\rho = \rho_{ani,2} \equiv \left\| I + \frac{1}{\alpha_{ani,2}} |H(v)| \right\|^{\frac{mnq}{2(n+(2-m)q)}} \cdot \det\left(I + \frac{1}{\alpha_{ani,2}} |H(v)| \right)^{\frac{(2-m)q}{2(n+(2-m)q)}} \tag{3.41}$$

$$M = M_{ani,2} \equiv \rho_{ani,2}^{\frac{2}{n}} \cdot \det\left(I + \frac{1}{\alpha_{ani,2}} |H(v)| \right)^{-\frac{1}{n}} \left[I + \frac{1}{\alpha_{ani,2}} |H(v)| \right], \tag{3.42}$$

and the intensity parameter $\alpha_{ani,2}$ has to be defined implicitly through the equation

$$\int_\Omega \rho_{ani,2}(x)dx = \frac{|\Omega|}{1-\beta}, \tag{3.43}$$

where, once again, β indicates roughly the percentage of the mesh points concentrated in the regions of large $\rho_{ani,2}$. It is recommended that β be chosen from $[0.5, 0.8]$. It can be shown that the so-defined $\alpha_{ani,2}$ satisfies

$$\alpha_{ani,2} \le C \left[\int_\Omega \mathrm{tr}\left(|H(v)| \right)^{\frac{nq}{n+(2-m)q}} dx \right]^{\frac{n+(2-m)q}{nq}} \le C|v|_{W^{2,\frac{nq}{n+(2-m)q}}(\Omega)}. \tag{3.44}$$

The interpolation error on a general mesh is bounded by

$$|v - \Pi_k v|_{W^{m,q}(\Omega)} \lesssim C N^{-\frac{(2-m)}{n}} \alpha_{ani,2} Q_{mesh,ani,2}, \tag{3.45}$$

where the overall quality measure is defined as

$$Q_{mesh,ani,2} = \left[\frac{1}{\sigma} \int_\Omega \left(\hat{Q}_{ali}^{\frac{m(n-1)}{n}} \cdot Q_{ali}^{\frac{2(n-1)}{n}} \cdot Q_{eq}^{\frac{(2-m)}{n}} \right)^q \rho_{ani,2} dx \right]^{\frac{1}{q}}. \tag{3.46}$$

with $Q_{mesh,ani,2} = 1$ for a mesh satisfying equidistribution and alignment conditions (2.2) and (2.3) with the monitor function defined in (3.42).

3.3.4 Remark on the computation of monitor functions

The formulas of the monitor function developed in this section depend on several factors, including the function regularity (through parameters l and p), the dimension of space (n), and the norm used to measure interpolation error (m and q). Moreover, the formulas involve derivatives (of order $l \leq k+1$) of the physical solution, which is unknown in general. Fortunately, adaptive computation is often carried out in an iterative fashion and approximations of the nodal values of the physical solution on the current mesh are always available. A gradient recovery technique such as those of Zienkiewicz and Zhu [58,59] and Zhang and Naga [57] can then be used for computing the needed derivatives.

4 Adaptive anisotropic mesh generation

In this section we study the use of the principles discussed in §2 in the design of algorithms for generating meshes that satisfy equidistribution and alignment conditions (2.2) and (2.3). We focus on two types of mesh adaptation method, variational and refinement methods, for steady state problems.

4.1 Variational mesh adaptation

In the variational approach of mesh adaptation, adaptive meshes are generated as images of a computational mesh under a coordinate transformation from the computational domain to the physical domain. The coordinate transformation is determined by the so-called adaptation functional which is commonly designed to measure the difficulty in the numerical approximation of the physical solution. The functional often involves mesh properties and employs a monitor function to control mesh concentration.

The key to the development of variational methods is to formulate the adaptation functional. Generally speaking, an error bound such as (3.14) - (3.16) cannot be used directly as an adaptation functional since they are highly nonlinear and non-convex and their direct minimization often leads to a nasty optimization problem. Adaptation functionals of most existing variational methods have been developed based on geometric and

physical considerations; e.g., see [9, 10, 23, 41, 54] and the books [24, 40, 45, 53], and the references therein. It is worth mentioning that adaptation functionals for steady problems can also be used for solving time dependent problems, particularly in the context of moving meshes; see [5, 13, 36, 44, 51, 56].

In this section we derive a functional for the inverse coordinate transformation $\xi = \xi(x) : \Omega \to \Omega_c$ based on the equidistribution and alignment conditions (2.2) and (2.3). The functional has been proposed in [33] using the isotropy (or conformal) and equidistribution arguments.

We first consider the alignment condition (2.3). Denote the eigenvalues of matrix $\boldsymbol{J}^{-1}M^{-1}\boldsymbol{J}^{-T}$ by $\lambda_1, ..., \lambda_n$. By the arithmetic-geometric mean inequality, the desired coordinate transformation (which should satisfy (2.3) as closely as possible) can be obtained by minimizing the difference between the two sides of the inequality

$$\left(\prod_i \lambda_i\right)^{\frac{1}{n}} \le \frac{1}{n}\sum_i \lambda_i. \tag{4.1}$$

From

$$\sum_i \lambda_i = \mathrm{tr}\left(\boldsymbol{J}^{-1}M^{-1}\boldsymbol{J}^{-T}\right) = \sum_i (\nabla\xi_i)^T M^{-1}\nabla\xi_i,$$
$$\prod_i \lambda_i = \det\left(\boldsymbol{J}^{-1}M^{-1}\boldsymbol{J}^{-T}\right) = \frac{1}{(J\rho)^2},$$

where $\rho = \sqrt{\det(M)}$, (4.1) can be rewritten into

$$\left(\frac{1}{J\rho}\right)^{\frac{2}{n}} \le \frac{1}{n}\sum_i (\nabla\xi_i)^T M^{-1}\nabla\xi_i$$

or

$$n^{\frac{n\gamma}{2}}\frac{\rho}{(J\rho)^\gamma} \le \rho\left(\sum_i (\nabla\xi_i)^T M^{-1}\nabla\xi_i\right)^{\frac{n\gamma}{2}}$$

for any real number $\gamma \ge 1$. Globally, the coordinate transformation should be defined as a minimizer of the functional

$$\int_\Omega \left[\rho\left(\sum_i (\nabla\xi_i)^T M^{-1}\nabla\xi_i\right)^{\frac{n\gamma}{2}} - n^{\frac{n\gamma}{2}}\frac{\rho}{(J\rho)^\gamma}\right] dx. \tag{4.2}$$

On the other hand, it is known [30, 33] that the inequality

$$\int_\Omega \frac{\rho}{J\rho}dx = \int_{\Omega_c} d\xi \le \left(\int_\Omega \frac{\rho}{(J\rho)^\gamma}dx\right)^{\frac{1}{\gamma}}$$

holds for any number $\gamma > 1$, with equality if and only if $J\rho \equiv$ constant. Hence, the desired coordinate transformation according to the equidistribution condition (2.2) should be defined as a minimizer of the functional

$$\int_\Omega \frac{\rho}{(J\rho)^\gamma} dx. \tag{4.3}$$

The functional balancing the equidistribution and alignment conditions can then be obtained by weighing (4.2) and (4.3), viz.,

$$I[\xi] = \theta \int_\Omega \rho \left(\sum_i (\nabla \xi_i)^T M^{-1} \nabla \xi_i \right)^{\frac{n\gamma}{2}} dx + (1-2\theta) n^{\frac{n\gamma}{2}} \int_\Omega \frac{\rho}{(J\rho)^\gamma} dx, \tag{4.4}$$

where $\theta \in [0,1]$ and $\gamma > 1$. This functional has been derived in [33] from slightly different considerations. Numerical experiments in [12,33] have shown that $\theta = 0.1$ and $\gamma = 2$ work well for the tested problems.

4.2 Unstructured mesh adaptation

Refinement or h-version adaptation has been the most popular approach in use for generating unstructured meshes in finite element computation because of its reliability and conceptual simplicity. Typically with the approach, a local error estimate is computed and then local minimization tools are employed to generate the needed adaptive mesh; e.g., see [16, 21, 22, 25, 27].

For anisotropic mesh adaptation, the common practice in the refinement approach is to generate the needed anisotropic mesh as a quasi-uniform one in the metric determined by a tensor (or a matrix-valued function) that specifies the size, shape, and orientation of the mesh elements throughout the domain. Examples of anisotropic meshing strategies include the Delaunay triangulation method [6,7,16,47], the advancing front method [26], the bubble mesh method [55], and the method combining local modification with smoothing or node movement [3,8,22,29]. About a dozen computer codes, mostly in two dimensions, have been developed or modified with the anisotropic mesh option; e.g., see the meshing software survey by Owen [46]. Among these meshing strategies and computer codes, the metric tensor is commonly defined based on the Hessian of the physical solution, motivated mainly by the results of D'Azevedo [18] and D'Azevedo and Simpson [19] on linear interpolation for quadratic functions on triangles; e.g., see [16, 28, 29, 48]. A number of researchers (e.g., see [16, 25, 29, 48]) even use the Hessian directly as the metric tensor. The Hessian can give correct information for element shape and orientation but a multiplicative scalar function is needed to be added to give a correct distribution of element size, cf. the metric tensor given in (4.10) or (4.13).

Recently, a metric tensor was developed in [35] based on interpolation error estimates. The strategy used therein is similar to that used in the previous section for defining the monitor function. The main difference is for the case with $m > 0$ where the effect of

element skewness has now been taken into account (cf. inequality (3.39)) and the overall mesh quality measure (3.46) does not involve the geometric measure Q_{geo} or $\hat{Q}_{geo}$. This is important since it means that the mesh is required to satisfy only the equidistribution and alignment conditions (2.2) and (2.3), and such a mesh is more likely to exist (than the one satisfying (2.2), (2.3), and (3.18)).

Given a monitor function $M = M(x)$, it is straightforward to define the metric tensor, denoted by $\mathcal{M} = \mathcal{M}(x)$. This is because they play the same role in mesh adaptation: to specify the size, shape, and orientation of the mesh elements throughout the physical domain. To be specific, we consider in the following the definition of the metric tensor for a C++ mesh generation code, BAMG (Bidimensional Anisotropic Mesh Generator), developed by Hecht [31]. BAMG is a Delaunay-type triangulator which allows the user to supply a metric tensor defined at the vertices of a background mesh. The user-defined metric tensor should be given such that the elements of the desired mesh are isotropic and have a unitary volume in the metric. Once the metric is given, BAMG employs five local minimization tools, edge suppression, vertex suppression, vertex addition, edge swapping, and vertex reallocation (barycentering step), to generate the needed anisotropic mesh.

Since both M and $\mathcal{M}$ play the same role in mesh adaptation, it is reasonable to assume that they are related by

$$\mathcal{M}(x) = cM(x) \tag{4.5}$$

for some constant c. The unitary volume condition can be written mathematically as

$$\int_K \sqrt{\det(\mathcal{M}(x))}dx = 1, \quad \forall K \in \mathcal{T}_h. \tag{4.6}$$

Combining (4.5) and (4.6) gives

$$c^{\frac{n}{2}} \int_K \rho(x)dx = 1,$$

where $\rho = \sqrt{\det(M)}$. Summing it over all the elements, we get

$$c = \left(\frac{N}{\sigma}\right)^{\frac{2}{n}},$$

where N is the number of elements and $\sigma = \int_\Omega \rho dx$. Inserting this into (4.5) leads to

$$\mathcal{M}(x) = \left(\frac{N}{\sigma}\right)^{\frac{2}{n}} M(x). \tag{4.7}$$

For the monitor functions defined in §3, we then have

$$\mathcal{M}_{iso} = \left(\frac{N}{\sigma}\right)^{\frac{2}{n}} \left(1 + \frac{1}{\alpha_{iso}}\|D^l v\|_p\right)^{\frac{2q}{n+q(l-m)}} I \tag{4.8}$$

for the isotropic case,

$$\mathcal{M}_{ani,1} = \left(\frac{N}{\sigma}\right)^{\frac{2}{n}} \left(1 + \frac{1}{\alpha_{ani,1}^2}\|\nabla v\|^2\right)^{-\frac{1}{n+q}} \left[I + \frac{1}{\alpha_{ani,1}^2}\nabla v \nabla v^T\right] \tag{4.9}$$

for the anisotropic case with $l = 1$, and

$$\mathcal{M}_{ani,2} = \left(\frac{N}{\sigma}\right)^{\frac{2}{n}} \rho_{ani,2}^{\frac{2}{n}} \cdot \det\left(I + \frac{1}{\alpha_{ani,2}}|H(v)|\right)^{-\frac{1}{n}} \left[I + \frac{1}{\alpha_{ani,2}}|H(v)|\right] \tag{4.10}$$

for the anisotropic case with $l = 2$. Note that $\sigma = \int_\Omega \rho_{iso} dx$, $\int_\Omega \rho_{ani,1} dx$, and $\int_\Omega \rho_{ani,2} dx$ for the above three cases, respectively.

The metric tensor can also be given in terms of a prescribed error level. We take the isotropic case as an example. From (3.27) and (3.25), we can see that the error bound is asymptotically proportional to $N^{-(l-m)/n}\alpha_{iso}$ provided that $Q_{mesh,iso} \le C$ for some constant C. (Recall that the desired mesh is the one with $Q_{mesh,iso} = 1$.) Setting

$$N^{-\frac{l-m}{n}}\alpha_{iso} = \epsilon_0,$$

where ϵ_0 is a prescribed level for the global error $|v - \Pi_k v|_{W^{m,q}(\Omega)}$, we get

$$N = \left(\frac{\alpha_{iso}}{\epsilon_0}\right)^{\frac{n}{l-m}}.$$

Inserting it into (4.8) yields

$$\mathcal{M}_{iso} = \left(\frac{1}{\sigma}\left(\frac{\alpha_{iso}}{\epsilon_0}\right)^{\frac{n}{l-m}}\right)^{\frac{2}{n}} \left(1 + \frac{1}{\alpha_{iso}}\|D^l v\|_p\right)^{\frac{2q}{n+q(l-m)}} I \tag{4.11}$$

for the isotropic case. Similarly, we have

$$\mathcal{M}_{ani,1} = \left(\frac{1}{\sigma}\left(\frac{\alpha_{ani,1}}{\epsilon_0}\right)^{n}\right)^{\frac{2}{n}} \left(1 + \frac{1}{\alpha_{ani,1}^2}\|\nabla v\|^2\right)^{-\frac{1}{n+q}} \left[I + \frac{1}{\alpha_{ani,1}^2}\nabla v \nabla v^T\right] \tag{4.12}$$

for the anisotropic case with $l = 1$, and

$$\mathcal{M}_{ani,2} = \left(\frac{1}{\sigma}\left(\frac{\alpha_{ani,2}}{\epsilon_0}\right)^{\frac{n}{2-m}}\right)^{\frac{2}{n}} \rho_{ani,2}^{\frac{2}{n}} \cdot \det\left(I + \frac{1}{\alpha_{ani,2}}|H(v)|\right)^{-\frac{1}{n}} \left[I + \frac{1}{\alpha_{ani,2}}|H(v)|\right] \tag{4.13}$$

for the anisotropic case with $l = 2$. Note that ϵ_0 in (4.12) and (4.13) is a prescribed level of global error $\|v - \Pi_k v\|_{L^q(\Omega)}$ and $|v - \Pi_k v|_{W^{m,q}(\Omega)}$, respectively.

It is worth mentioning that the remark on the computation of the monitor function in §3.3.4 also holds for the computation of the metric tensor.

Finally we point out that, as for the monitor functions defined in the previous section, the metric tensors defined here are the same as those of [35], except for the anisotropic case $l = 2$ and $m = 1$. The present metric tensor is defined using a bound on the gradient of linear interpolation error which is based on a more realistic mesh requirement – with the mesh closely satisfying the equidistribution and alignment conditions (2.2) and (2.3).

5 Numerical examples

In this section we present some numerical results for four two-dimensional examples. Adaptive meshes are generated in an iterative fashion using BAMG (the Bidimensional Anisotropic Mesh Generator developed by Hecht [31]) with the metric tensor defined in §4.2. To be more specific, we assume that a current mesh is available. The nodal values of the physical solution, or their approximations as typically in the numerical solution of PDEs, are then obtained. This is followed by the computation of the metric tensor on the current mesh, where a gradient recovery technique is employed to calculate the solution derivatives needed in the definition of the metric tensor; see the remark made in §3.3.4. The adaptive mesh is finally obtained using BAMG with the computed metric tensor defined on the current mesh (which serves as the background mesh). The process is repeated twenty times.

We recall that the metric tensor defined in §4.2 is different from that defined in [35] for anisotropic cases with $m \geq 1$. For this reason, we focus our discussion on the isotropic and anisotropic cases with linear interpolation ($k = 1$ and $l = 2$) and the L^2 measure of gradient of the interpolation error ($m = 1$ and $p = q = 2$). The interested reader is referred to [34, 35] for more numerical results in anisotropic and variational mesh adaptation.

The parameter β used in the definitions of α_{iso} (3.25) and $\alpha_{ani,2}$ (3.43) is taken as $\beta = 0.75$ in all of the computation.

Example 5.1. This example is to generate an adaptive mesh for the analytical function

$$v(x, y) = \tanh(60y) - \tanh(60(x - y) - 30) \qquad \text{in } \Omega = (0, 1)^2. \tag{5.1}$$

This function simulates the interaction of a boundary layer along the x-axis with an oblique shock wave along the straight line $y = x - 0.5$.

Fig. 4 shows typical isotropic and anisotropic meshes obtained with metric tensors (4.8) and (4.10). The meshes have almost the same number of vertices (or elements) but the linear interpolation error in either H^1 seminorm or L^2 norm is much smaller on the anisotropic mesh than on the isotropic one. (In fact, the L^2 norm of the error is an order of magnitude smaller on the anisotropic mesh.) The aspect ratio of some anisotropic elements is about $\|Q_{geo}\|_\infty = 17$.

The advantage of using an adaptive mesh over a uniform one, or an anisotropic mesh over an isotropic one, can be seen from Fig. 5 where the linear interpolation error is depicted as function of the number of elements. Fig. 6 shows the geometric, alignment,

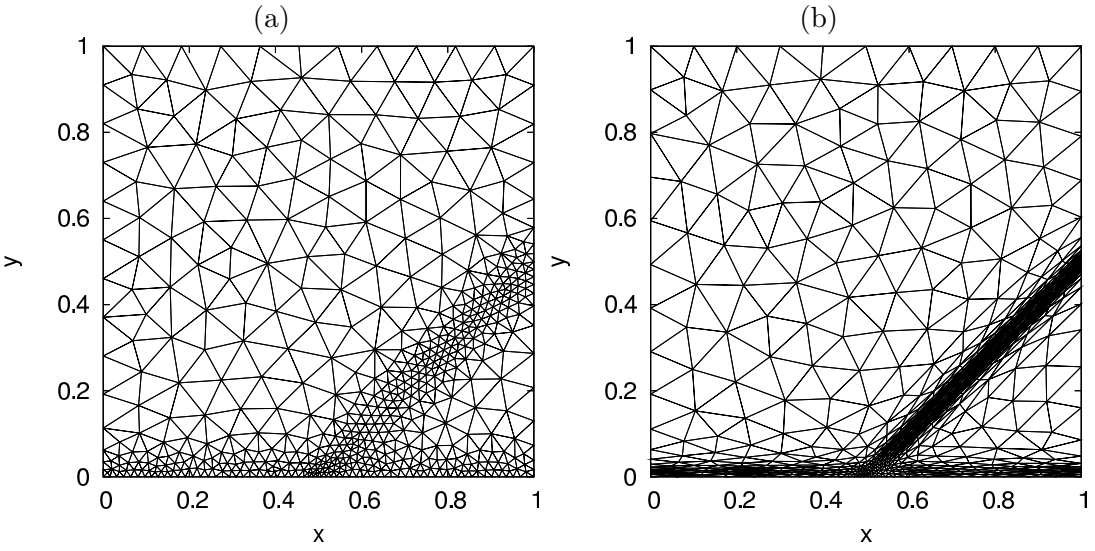

Figure 4: Example 5.1. (a) An isotropic mesh is obtained with metric tensor $\mathcal{M}_{iso}$ (4.8): $nbv = 567$, $nbt = 1022$, $|e|_{H^1} = 3.4$, $\|e\|_{L^2} = 1.6e-2$, $\|Q_{geo}\|_\infty = 1.5$. (b) An anisotropic mesh is obtained with metric tensor $\mathcal{M}_{ani,2}$ (4.10): $nbv = 566$, $nbt = 1039$, $|e|_{H^1} = 0.83$, $\|e\|_{L^2} = 1.4e-3$, $\|Q_{geo}\|_\infty = 17$, $\|Q_{ali}\|_\infty = 1.5$.

and overall qualities for isotropic and anisotropic meshes. On can see that the mesh quality does not deteriorate as the mesh is refined. Moreover, a large value of $\|Q_{geo}\|_\infty$ for the anisotropic mesh indicates that some elements have large aspect ratio and are highly skewed. Nevertheless, the overall mesh quality measure, which measures the effect of mesh quality on interpolation error (cf. bound (3.45)), maintains small. This is not surprising since the definition of the measure, (3.46), does not involve Q_{geo}. It instead contains the alignment and equidistribution measures which are small for the current situation (see Fig. 6(a)).

In Fig. 7(a) the actual number (nbt) of elements is plotted against the prescribed number (N) of elements and in Fig. 7(b) the H^1 seminorm of interpolation error ($|e|_{H^1}$) is plotted against the prescribed level of the error (ϵ_0). We can see that $nbt \propto N$ for sufficiently large N and $|e|_{H^1} \propto \epsilon_0$ for sufficiently small ϵ_0. This shows that the metric tensor defined in §4.2 provides an effective control on adaptive mesh generation.

Example 5.2. The second example is to solve a nonlinear boundary value problem (BVP) for the differential equation

$$-0.005(v_{xx} + v_{yy}) + vv_x + vv_y = f \qquad \text{in } \Omega = (0,1)^2. \tag{5.2}$$

The function f and the Dirichlet boundary condition are chosen such that the BVP has the exact solution

$$v(x,y) = \left(1.0 + e^{\frac{x+y-0.85}{0.01}}\right)^{-1}. \tag{5.3}$$

The PDE is discretized on a triangular mesh using linear elements. It is emphasized that

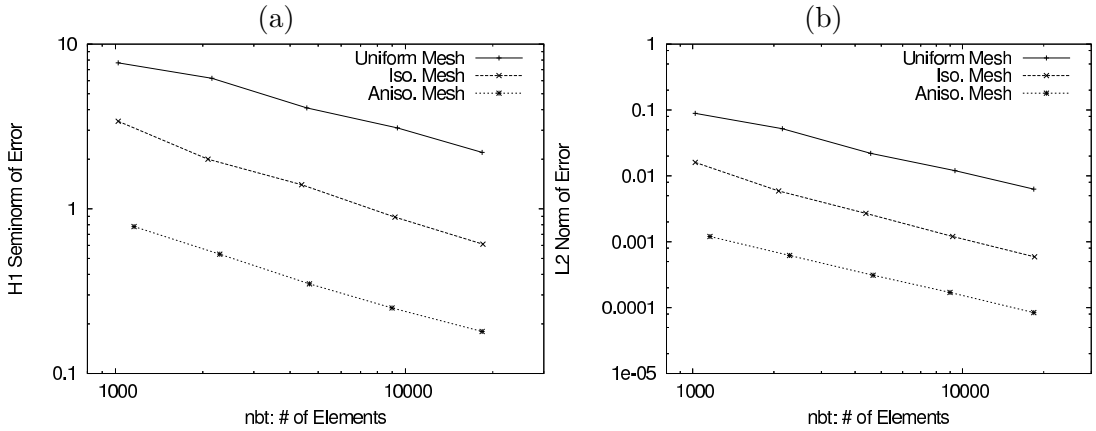

Figure 5: Example 5.1. The H^1 seminorm and L^2 norm of linear interpolation error are shown against the number of elements.

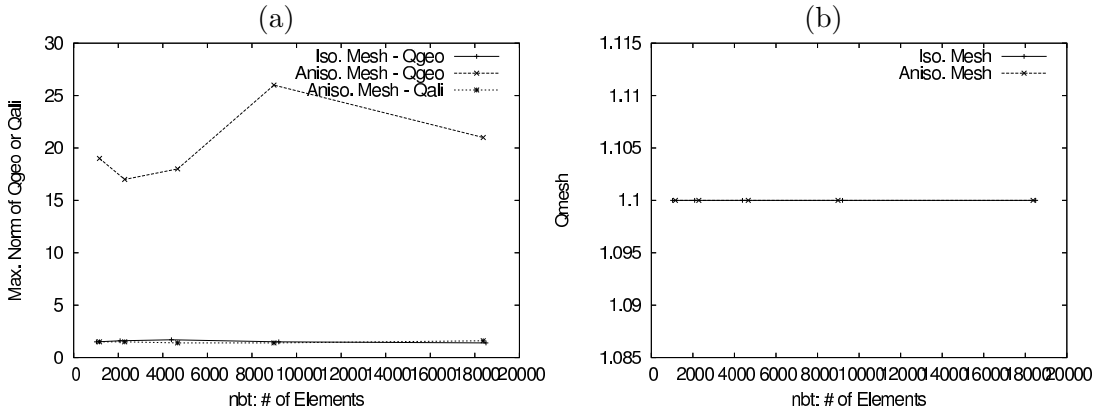

Figure 6: Example 5.1. The geometric, alignment, and overall mesh quality measures are shown as function of the number of elements for isotropic and anisotropic meshes.

although the metric tensor has been derived in §4.2 using interpolation error estimates (a priori), our computation is completely a posteriori since the metric tensor is calculated and therefore the mesh is generated based on the computed solution of the BVP.

Numerical results are shown in Figs. 8 and 9. Once again, the advantage of using anisotropic meshes is obvious.

Example 5.3. The third example is to solve a BVP for the Laplace equation

$$v_{xx} + v_{yy} = 0 \tag{5.4}$$

defined on $\Omega \equiv \{0 \le r < 1,\, 0 < \theta < 7\pi/4\}$. The Dirichlet boundary condition is chosen

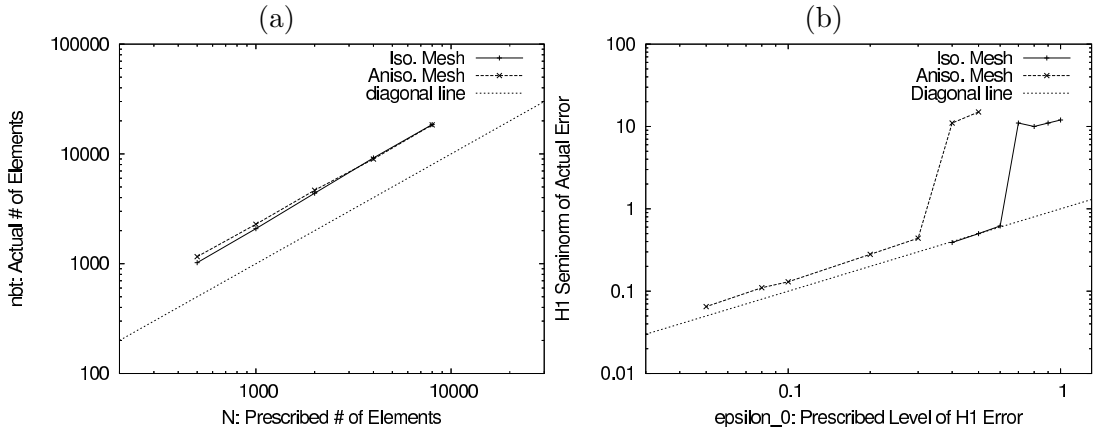

Figure 7: Example 5.1. (a) The actual number (nbt) of elements is depicted against the prescribed number (N) of elements. (b) The H^1 seminorm of interpolation error is shown against the prescribed level of the error.

such that the exact solution is given by

$$v = r^{\frac{2}{7}} \sin\left(\frac{2\pi}{7}\right). \tag{5.5}$$

The differential equation is discretized using linear elements on a triangular mesh.

The results are shown in Figs. 10 and 11. Note that the solution of this example does not have strong anisotropic features and the meshes generated using $\mathcal{M}_{ani,2}$ are isotropic and similar to those obtained with $\mathcal{M}_{iso}$. Fig. 11 shows that the solution error is almost the same on the meshes generated using metric tensors $\mathcal{M}_{iso}$ and $\mathcal{M}_{ani,2}$.

Example 5.4. The last example is to generate an adaptive mesh for a function with two components,

$$\begin{aligned} v_1 &= (10x^3 + y^3) + \text{atan2}(0.001, \sin(5y) - 2x), \\ v_2 &= (10y^3 + x^3) + \text{atan2}(0.01, \sin(5x) - 2y) \end{aligned} \tag{5.6}$$

defined on the unit disk. Here, $\text{atan2}(y, x) = \tan^{-1}(y/x)$ is the two-parameter arctangent function which returns an angle in the range from $-\pi$ to π. This example is taken from a document of BAMG [31] and used to verify the intersection of matrices discussed in Remark 3.3 of §3.1. Specifically, a metric tensor is constructed based on interpolation error for each of the two components. The final metric tensor is defined as the intersection of the metric tensors associated with the function components.

The numerical results are shown in Figs. 12 and 13. We can see that the linear interpolation error on the uniform and isotropic meshes barely converges in the H^1 seminorm or converges slowly in the L^2 norm in the considered range of the number of elements whereas that on the anisotropic mesh converges in the first and second order in the H^1 seminorm and the L^2 norm, respectively. The advantage of using an anisotropic mesh is clear, especially when the number of elements is relatively large.

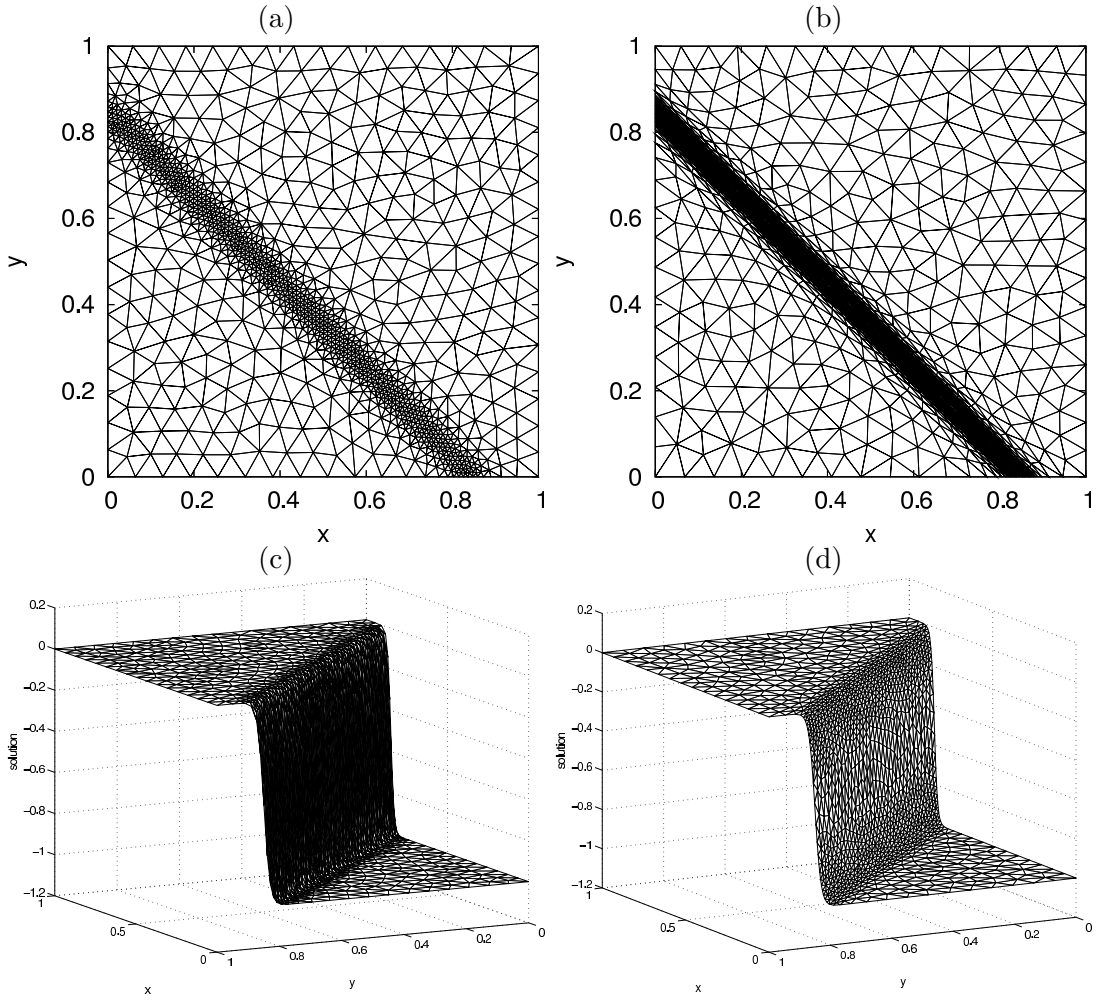

Figure 8: Example 5.2. (a) An isotropic mesh is obtained with metric tensor $\mathcal{M}_{iso}$ (4.8): $nbv = 1198$, $nbt = 2312$, $|e|_{H^1} = 1.1$, $\|e\|_{L^2} = 1.5e-2$, $\|Q_{geo}\|_\infty = 1.4$. (b) An anisotropic mesh is obtained with metric tensor $\mathcal{M}_{ani,2}$ (4.10): $nbv = 1229$, $nbt = 2347$, $|e|_{H^1} = 0.21$, $\|e\|_{L^2} = 2.1e-3$, $\|Q_{geo}\|_\infty = 16$, $\|Q_{ali}\|_\infty = 1.5$. (c) and (d) The computed solution is obtained on the meshes shown in (a) and (b), respectively.

6 Conclusions

Mesh adaptation has been studied in the previous sections in the mesh control point of view. It has been shown that the equidistribution and alignment principles, represented by the conditions (2.2) and (2.3), are necessary and sufficient for a complete control of the size, shape, and orientation of mesh elements throughout the physical domain. A key component in the conditions is the monitor function, a symmetric and positive definite matrix prescribed by the user for specifying the information on the mesh.

The definition of the monitor function $M = M(x)$ has been studied based on the

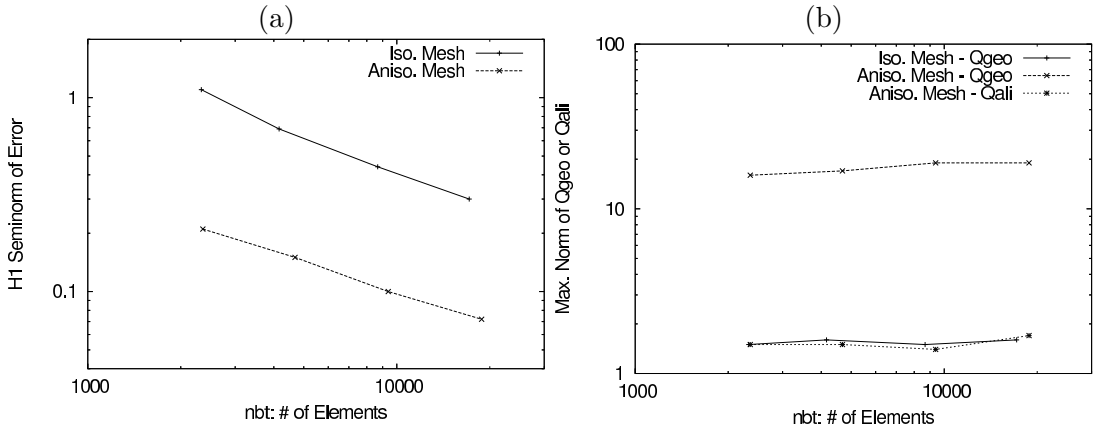

Figure 9: Example 5.2. (a) The H^1 seminorm of linear interpolation error is shown against the number of elements. (b) The geometric and alignment mesh quality measures are shown as function of the number of elements.

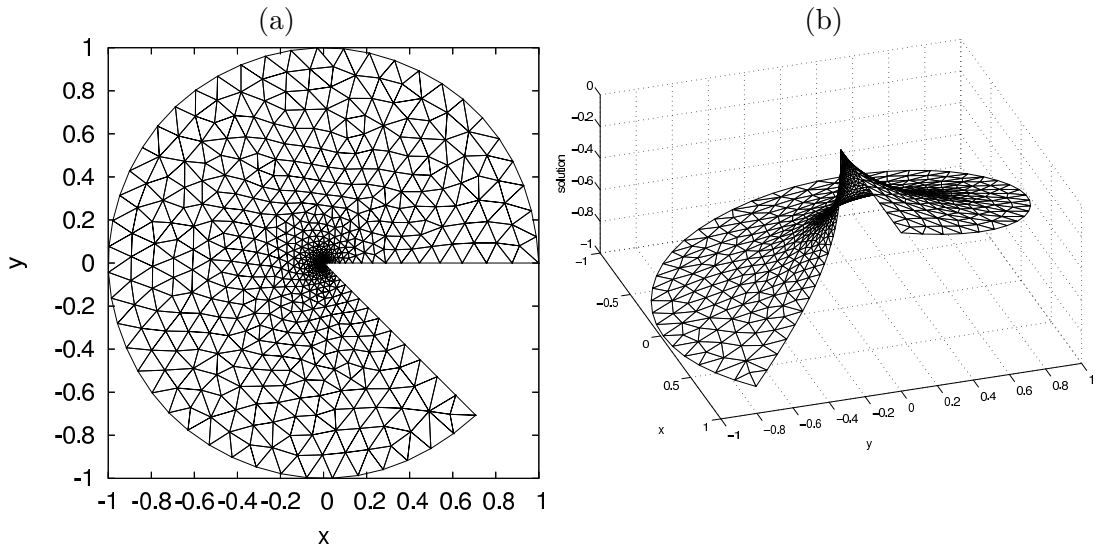

Figure 10: Example 5.3. (a) An anisotropic mesh is obtained with metric tensor $\mathcal{M}_{ani,2}$ (4.10): $nbv = 718$, $nbt = 1338$, $|e|_{H^1} = 3.9$, $\|e\|_{L^2} = 7.5e-4$, $\|Q_{geo}\|_\infty = 1.7$, $\|Q_{ali}\|_\infty = 1.6$. (b) The computed solution is obtained on the mesh shown in (a).

interpolation error on simplicial elements. The basic idea is to define M such that an error bound is minimized to some extent on a mesh satisfying the equidistribution and alignment conditions (2.2) and (2.3). The monitor function is given in (3.23) for the isotropic case, in (3.31) for the anisotropic case with $l = 1$ (such as piecewise constant interpolation), and in (3.42) for the anisotropic case with $l = 2$ (such as piecewise linear interpolation). These monitor functions are strictly positive definite and parameter free, and thus there is no need for regularization and parameter tuning in using them.

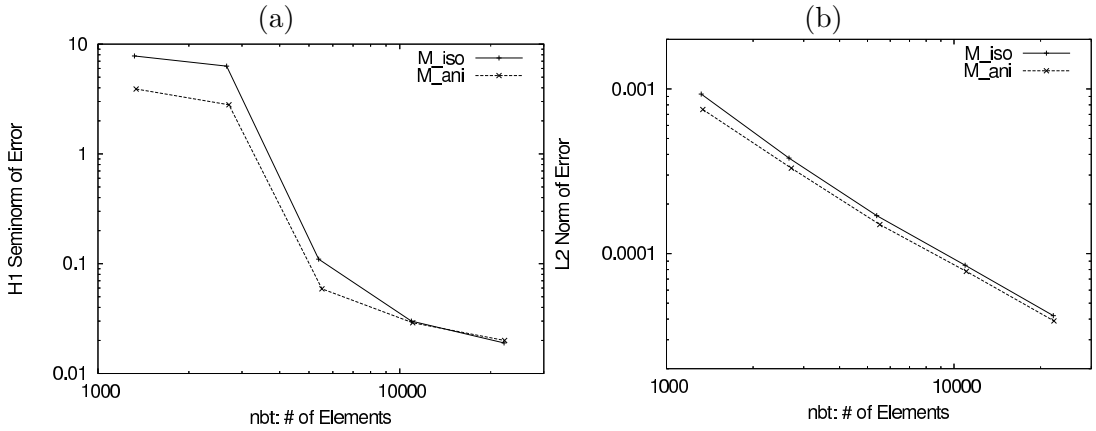

Figure 11: Example 5.3. The H^1 seminorm and L^2 norm of linear interpolation error are shown against the number of elements.

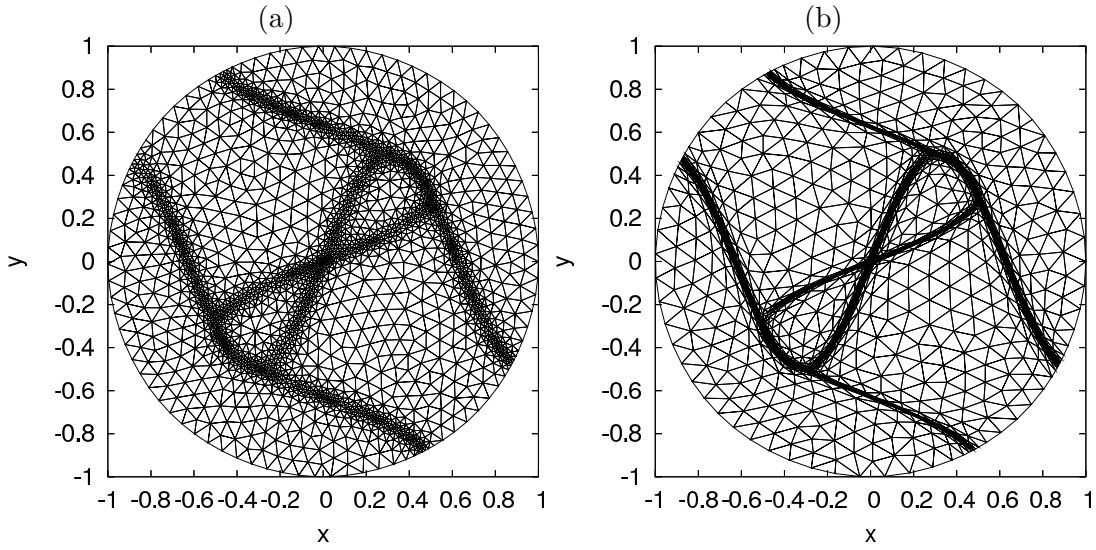

Figure 12: Example 5.4. (a) An isotropic mesh is obtained with metric tensor $\mathcal{M}_{iso}$ (4.8): $nbv = 2468$, $nbt = 4833$, $|e|_{H^1} = 150$, $\|e\|_{L^2} = 0.24$, $\|Q_{geo}\|_\infty = 1.7$, $Q_{mesh} = 1.1$. (b) An anisotropic mesh is obtained with metric tensor $\mathcal{M}_{ani,2}$ (4.10): $nbv = 2401$, $nbt = 4709$, $|e|_{H^1} = 83$, $\|e\|_{L^2} = 0.073$, $\|Q_{geo}\|_\infty = 120$, $\|Q_{ali}\|_\infty = 21$, $Q_{mesh} = 1.2$.

Algorithms for generating meshes satisfying conditions (2.2) and (2.3) for a given monitor function have been investigated. The focus has been on variational mesh adaptation and unstructured mesh refinement. A mesh adaptation functional, (4.4), first proposed in [33], was re-derived here using the equidistribution and alignment conditions. On the other hand, a common practice for generating unstructured, anisotropic meshes is to generate them as isotropic ones in the metric determined by a tensor (or a matrix-valued function). The metric tensor plays a crucial role in this approach. A metric tensor has

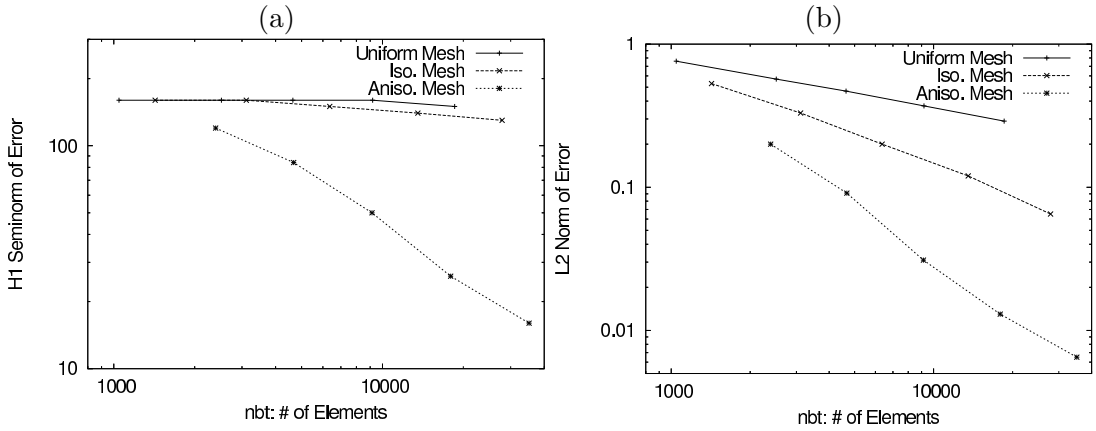

Figure 13: Example 5.4. The H^1 seminorm and L^2 norm of linear interpolation error are shown against the number of elements.

been developed based on interpolation error: (4.8) for the isotropic case, (4.9) for the anisotropic case with $l = 1$, and (4.10) for the anisotropic case with $l = 2$, corresponding to the monitor function for each case. These formulas are expressed in terms of a prescribed number of elements. They can also be written in terms of a prescribed level of interpolation error; see (4.11), (4.12), and (4.13). Numerical results were presented to demonstrate that the metric tensors can be used to generate proper anisotropic meshes for a given function or the solution of a given PDE. They also showed that a properly generated anisotropic mesh can produce more accurate results than an isotropic one, especially when the solution of the PDE exhibits a strong anisotropic feature that the solution changes more significantly in one direction than the others.

References

[1] R. A. Adams, Sobolev Spaces, Academic Press, New York, 1975.
[2] M. Ainsworth and J. T. Oden, A Posteriori Error Estimation in Finite Element Analysis, Pure and Applied Mathematics (New York), Wiley-Interscience, John Wiley & Sons, New York, 2000.
[3] D. Ait-Ali-Yahia, G. Baruzzi, W. G. Habashi, M. Fortin, J. Dompierre and M.-G. Vallet, Anisotropic mesh adaptation: towards user-indepedent, mesh-independent and solver-independent CFD. Part II: Structured grids, Int. J. Numer. Meth. Fluids, 39 (2002), 657-673.
[4] M. J. Baines, Least squares and approximate equidistribution in multidimensions, Numer. Meth. P. D. E., 15 (1999), 605 – 615.
[5] G. Beckett, J. A. Mackenzie and M. L. Robertson, A moving mesh finite element method for the solution of two-dimensional Stefan problems, J. Comput. Phys., 168 (2001), 500 – 518.
[6] H. Borouchaki, P. L. George, P. Hecht, P. Laug and E. Saletl, Delaunay mesh generation governed by metric specification: Part I. Algorithms, Finite Elem. Anal. Des., 25 (1997), 61 – 83.
[7] H. Borouchaki, P. L. George and B. Mohammadi, Delaunay mesh generation governed by metric specification: Part II. Applications, Finite Elem. Anal. Des., 25 (1997), 85 – 109.

[8] F. J. Bossen and P. S. Heckbert, A pliant method for anisotropic mesh generation, In: Proceedings, 5th International Meshing Roundtable, Sandia National Laboratories, Albuquerque, NM, 1996, pp. 63–74, Sandia Report 96-2301.
[9] J. U. Brackbill, An adaptive grid with directional control, J. Comput. Phys., 108 (1993), 38 – 50.
[10] J. U. Brackbill and J. S. Saltzman, Adaptive zoning for singular problems in two dimensions, J. Comput. Phys., 46 (1982), 342 – 368.
[11] H. G. Burchard, Splines (with optimal knots) are better, Appl. Anal., 3 (1974), 309 – 319.
[12] W. Cao, R. Carretero-González, W. Huang and R. D. Russell, Variational mesh adaptation methods for axisymmetrical problems, SIAM J. Numer. Anal., 41 (2004), 235 – 257 (electronic).
[13] W. Cao, W. Huang and R. D. Russell, An r-adaptive finite element method based upon moving mesh pdes, J. Comp. Phys., 149 (1999), 221 – 244.
[14] 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.
[15] G. Carey, Computational Grids: Generation, Adaptation, and Solution Strategies, Taylor & Francis, Washington, DC, 1997.
[16] M. J. Castro-Díaz, F. Hecht, B. Mohammadi and O. Pironneau, Anisotropic unstructured mesh adaption for flow simulations, Int. J. Numer. Meth. Fluids, 25 (1997), 475 – 491.
[17] P. G. Ciarlet, The Finite Element Method for Elliptic Problems, North-Holland, Amsterdam, 1978.
[18] E. F. D'Azevedo, Optimal triangular mesh generation by coordinate transformation, SIAM J. Sci. Stat. Comput., 12 (1991), 755 – 786.
[19] E. F. D'Azevedo and R. B. Simpson, On optimal triangular meshes for minimizing the gradient error, Numer. Math., 59 (1991), 321 – 348.
[20] C. de Boor, Good approximation by splines with variables knots II, in: G. A. Watson (Ed.), Lecture Notes in Mathematics 363, Springer-Verlag, Berlin, 1974, pp. 12-20.
[21] V. Dolejsi, Anisotropic mesh adaptation for finite volume and finite element methods on triangular meshes, Comput. Visual. Sci., 1 (1998), 165 – 178.
[22] J. Dompierre, M.-G. Vallet, Y. Bourgault, M. Fortin and W. G. Habashi, Anisotropic mesh adaptation: towards user-indepedent, mesh-independent and solver-independent CFD. Part III: Unstructured meshes, Int. J. Numer. Meth. Fluids, 39 (2002) 675 – 702.
[23] A. S. Dvinsky, Adaptive grid generation from harmonic maps on riemannian manifolds, J. Comput. Phys., 95 (1991), 450 – 476.
[24] J. E. Castillo (Ed.), Mathematical Aspects of Numerical Grid Generation, Frontiers in Applied Mathematics vol 8, SIAM, Philadelphia, 1991.
[25] P. Frey and P.-L. George, Mesh Generation: Application to Finite Elements, Hermes Science, 2000.
[26] R. V. Garimella and M. S. Shephard, Boundary layer meshing for viscous flows in complex domain, in: Proceedings, 7th International Meshing Roundtable, Sandia National Laboratories, Albuquerque, NM, 1998, pp. 107 – 118.
[27] P. L. George, Automatic mesh generation and finite element computation, in: P. G. Ciarlet and J. L. Lions (Eds.), Handbook of Numerical Analysis, Vol. IV, Elsevier Science B.V., 1996.
[28] P. L. George and F. Hecht, Nonisotropic grids, in: J. F. Thompson, B. K. Soni and N. P. Weatherill (Eds.), Handbook of Grid Generation, CRC Press, Boca Raton, 1999, pp. 20:1 – 20:29.
[29] W. G. Habashi, J. Dompierre, Y. Bourgault, D. Ait-Ali-Yahia, M. Fortin and M.-G. Val-

let, Anisotropic mesh adaptation: towards user-indepedent, mesh-independent and solver-independent CFD. Part I: General principles, Int. J. Numer. Meth. Fluids, 32 (2000), 725 – 744.
[30] G. H. Hardy, J. E. Littlewood and G. Pólya, Inequalities, Cambridge University Press, Cambridge, 1934.
[31] F. Hecht, Bidimensional anisotropic mesh generator, Technical Report, INRIA, Rocquencourt, 1997, Source code: http://www-rocq1.inria.fr/gamma/cdrom/www/bamg/eng.htm.
[32] W. Huang, Practical aspects of formulation and solution of moving mesh partial differential equations, J. Comput. Phys., 171 (2001), 753 – 775.
[33] W. Huang, Variational mesh adaptation: isotropy and equidistribution, J. Comput. Phys., 174 (2001), 903 – 924.
[34] W. Huang, Measuring mesh qualities and application to variational mesh adaptation, SIAM J. Sci. Comput., 26 (2005), 1643 – 1666.
[35] W. Huang, Metric tensors for anisotropic mesh generation, J. Comput. Phys., 204 (2005), 633 – 665.
[36] W. Huang and R. D. Russell, A high dimensional moving mesh strategy, Appl. Numer. Math., 26 (1997), 63 – 76.
[37] W. Huang and D. M. Sloan, A simple adaptive grid method in two dimensions, SIAM J. Sci. Comput., 15 (1994), 776 – 797.
[38] W. Huang and W. Sun, Variational mesh adaptation II: Error estimates and monitor functions, J. Comput. Phys., 184 (2003), 619 – 648.
[39] P. Knupp, L. Margolin and M. Shashkov, Reference jacobian optimization-based rezone strategies for arbitrary lagrangian eulerian methods, J. Comput. Phys., 176 (2002), 93 – 128.
[40] P. Knupp and S. Steinberg, Fundamentals of Grid Generation, CRC Press, Boca Raton, 1994.
[41] P. M. Knupp, Jacobian-weighted elliptic grid generation, SIAM J. Sci. Comput., 17 (1996), 1475 – 1490.
[42] P. M. Knupp and N. Robidoux, A framework for variational grid generation: conditioning the Jacobian matrix with matrix norms, SIAM J. Sci. Comput., 21 (2000), 2029 – 2047.
[43] R. Li, T. Tang and P. W. Zhang, Moving mesh methods in multiple dimensions based on harmonic maps, J. Comput. Phys., 170 (2001), 562 – 588.
[44] R. Li, T. Tang and P. W. Zhang, A moving mesh finite element algorithm for singular problems in two and three space dimensions, J. Comput. Phys., 177 (2002), 365 – 393.
[45] V. D. Liseikin, Grid Generation Methods, Springer, Berlin, 1999.
[46] S. Owen, Meshing software survey, 1998, http://www.andrew.cmu.edu/user/sowen/softsurv.html.
[47] J. Peraire, M. Vahdati, K. Morgan and O. C. Zienkiewicz, Adaptive remeshing for compressible flow computations, J. Comput. Phys., 72 (1997), 449 – 466.
[48] J. Remacle, X. Li, M. S. Shephard and J. E. Flaherty, Anisotropic adaptive simulation of transient flows using discontinuous Galerkin methods, Int. J. Numer. Meth. Engrg., 62 (2005), 899-923.
[49] R. B. Simpson, Anisotropic mesh transformations and optimal error control, Appl. Numer. Math., 14 (1994), 183 – 198.
[50] S. Steinberg and P. J. Roache, Variational grid generation, Numer. Meth. P. D. E., 2 (1986), 71 – 96.
[51] H. Tang and T. Tang, Adaptive mesh methods for one- and two-dimensional hyperbolic conservation laws, SIAM J. Numer. Anal., 41 (2003), 487 – 515, (electronic).
[52] J. F. Thompson, B. K. Soni, and N. P. Weatherill (Eds.), Handbook of Grid Generation, CRC Press, Boca Raton and London, 1999.

[53] J. F. Thompson, Z. A. Warsi and C. W. Mastin, Numerical Grid Generation: Foundations and Applications, North-Holland, New York, 1985.
[54] A. M. Winslow, Adaptive mesh zoning by the equipotential method, Technical Report UCID-19062, Lawrence Livemore Laboratory, 1981.
[55] S. Yamakawa and K. Shimada, High quality anisotropic tetrahedral mesh generation via ellipsoidal bubble packing, in: Proceedings, 9th International Meshing Roundtable, Sandia National Laboratories, Albuquerque, NM, 2000, pp. 263 – 273, Sandia Report 2000-2207.
[56] P. A. Zegeling and H. P. Kok, Adaptive moving mesh computations for reaction-diffusion systems, J. Comput. Appl. Math., 168 (2004), 519 – 528.
[57] Z. Zhang and A. Naga, A new finite element gradient recovery method: Superconvergence property, SIAM J. Sci. Comput., 26 (2005), 1192-1213.
[58] O. C. Zienkiewicz and J. Z. Zhu, The superconvergence patch recovery and a posteriori error estimates. Part 1: The recovery technique, Int. J. Numer. Meth. Engrg., 33 (1992), 1331 – 1364.
[59] O. C. Zienkiewicz and J. Z. Zhu, The superconvergence patch recovery and a posteriori error estimates. Part 2: Error esimates and adaptivity, Int. J. Numer. Meth. Engrg., 33 (1992), 1365 – 1382.