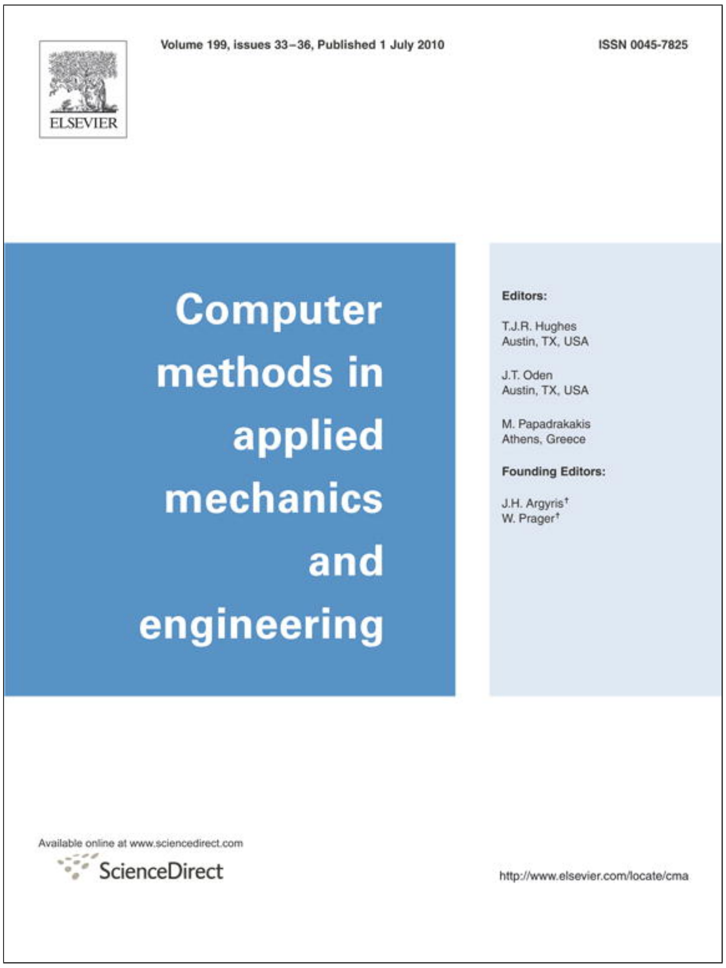
Volume 199, issues 33–36, Published 1 July 2010
ISSN 0045-7825
ELSEVIER
Computer methods in applied mechanics and engineering
Editors:
T.J.R. Hughes
Austin, TX, USA
J.T. Oden
Austin, TX, USA
M. Papadrakakis
Athens, Greece
Founding Editors:
J.H. Argyris†
W. Prager†
Available online at www.sciencedirect.com
ScienceDirect
http://www.elsevier.com/locate/cma

Computer Methods in Applied Mechanics and Engineering 199 (2010) 2195–2203

Contents lists available at ScienceDirect

Computer Methods in Applied Mechanics and Engineering

journal homepage: www.elsevier.com/locate/cma

Minimization of gradient errors of piecewise linear interpolation on simplicial meshes

Abdellatif Agouzal [a], Yuri V. Vassilevski [b,*]

[a] U.M.R. 5585 - Equipe d'Analyse Numerique Lyon Saint-Etienne Universite de Lyon 1, Laboratoire d'Analyse Numerique Bat. 101, 69 622 Villeurbanne Cedex, France
[b] Institute of Numerical Mathematics, Russian Academy of Sciences, Moscow 119333, Russia

A R T I C L E   I N F O

Article history:
Received 25 August 2009
Received in revised form 21 February 2010
Accepted 17 March 2010
Available online 25 March 2010

Keywords:
Optimal mesh
Gradient interpolation error
Metric based adaptation

A B S T R A C T

The paper is devoted to the analysis of optimal simplicial meshes which minimize the gradient error of the piecewise linear interpolation over all conformal simplicial meshes with a fixed number of cells N_T. We present theoretical results on asymptotic dependencies of L_p-norms of the gradient error on N_T for spaces of arbitrary dimension d. Our analysis is based on a geometric representation of the gradient error of linear interpolation on a simplex and a relaxed saturation assumption. We derive a metric field $\mathfrak{M}_p$ such that a $\mathfrak{M}_p$-quasi-uniform mesh is quasi-optimal, for arbitrary d and $p \in]0, +\infty]$. Quasi-optimal meshes provide the same asymptotics of the L_p-norm of the gradient error as the optimal meshes.

1. Introduction

The paper is devoted to the analysis of optimal meshes. These meshes minimize the gradient error of the piecewise linear interpolation over all conformal simplicial meshes with a fixed number of cells N_T. Possible anisotropy of optimal meshes hampers the interpolation error analysis. We present theoretical results on asymptotic dependencies of L_p-norms of the error on N_T for spaces of arbitrary dimension d. Asymptotic analysis of optimal meshes minimizing L_p-norm of the interpolation error was done in [1] for $p = +\infty$, $d = 2$, in [2] for $p = +\infty$, $d = 3$, and in [3] for $p \in]0, +\infty]$, $d = 2, 3$ (a similar result was obtained in [4] for convex functions only). The present work is the generalization of these results to the case of the gradient error and arbitrary d, $p \in]0, +\infty]$, provided that a relaxed saturation assumption [5] holds true.

In practice, the conventional adaptive procedures produce meshes close to optimal. Such meshes are called quasi-optimal. They give slightly higher errors but the same asymptotic rate of error reduction. Quasi-optimal meshes are shown to be uniform or quasi-uniform in an appropriate continuous tensor metric. In most papers, the metric is based on the continuous Hessian of the interpolated function: [1] for $p = +\infty$, $d = 2$, [2] for $p = +\infty$, $d = 3$, [3] ([4] for convex functions) for $p \in]0, +\infty]$, $d = 2, 3$, [6,7] for the gradient interpolation error, arbitrary d and $p \in [1, +\infty]$. The use of Hessian-based metrics requires a method of the discrete Hessian recovery. The accuracy of the Hessian recovery is very low producing relative errors as much as 50% and more, although the adaptive methods exhibit surprisingly good behavior in practice [8–10].

Error estimators [11–15] may provide a reliable alternative for metric recovery. We consider linear interpolation of quadratic functions and suggest a new method of computation of the gradient error. The method yields a reliable and efficient estimator of the interpolation error for general functions provided a relaxed saturation assumption is valid. We prove the relaxed saturation assumption up to the oscillation term which is small on a wide class of fine meshes. With the estimator we derive a metric field $\mathfrak{M}_p$ such that a $\mathfrak{M}_p$-quasi-uniform mesh is quasi-optimal one with respect to L_p-norm of the gradient error for arbitrary d and $p \in]0, +\infty]$. Neither the metric, nor the analysis of quasi-optimal meshes rely on the recovered Hessian of the interpolated function. The estimator can be extended to FEM discretizations of PDEs [15]. We mention also an alternative approach to the construction of quasi-optimal finite element discretizations based on the best tree approximation [16].

In addition to the error analysis, we discuss technical issues of the numerical implementation. In particular, we consider recovery of a continuous tensor metric field from a given piecewise constant tensor metric field. Also we present our technique for generation of $\mathfrak{M}_p$-quasi-uniform meshes with a prescribed number of elements.

The main results of this paper are as follows. First, we give an asymptotic error analysis for optimal meshes in d-dimensional spaces L_p, $p \in]0, +\infty]$. Second, we present and motivate a new reliable and efficient estimator for the gradient of interpolation error. Third, we define a particular metric yielding quasi-optimal meshes. Asymptotic error analysis for these meshes is given in d-dimensional spaces L_p, $p \in]0, +\infty]$.

The paper outline is as follows. In Section 2 we present the new method of metric recovery based on function values associated with mesh edges. In Section 3 we derive the new error estimator for linear interpolation of quadratic functions. In Section 4 we extend this estimator to general functions using the relaxed saturation assumption. In

* Corresponding author. Tel.: +7 495 9383911; fax: +7 495 9381821.
E-mail addresses: agouzal@univ-lyon1.fr (A. Agouzal), vasilevs@dodo.inm.ras.ru (Y.V. Vassilevski).

0045-7825/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.cma.2010.03.019

Section 5 we prove that the relaxed saturation assumption holds up to an oscillation term. A local gradient error analysis in general spaces L_p is presented in Section 6. Asymptotic error analysis of both optimal and quasi-optimal meshes is given in Section 7. In Section 8 we discuss algorithmic aspects of our methodology. Numerical experiments illustrating our analysis are presented in Section 9.

2. Metric recovery based on simplex edge data

Let Δ be a d-simplex (triangle for $d=2$, tetrahedron for $d=3$) with vertices $\mathbf{a}_i$, $i=1,\ldots,d_1$, $d_1=d+1$, and edges $\mathbf{e}_k$, $k=1,\ldots,n_d$, $n_d=d(d+1)/2$, such that $\mathbf{e}_k=\mathbf{a}_j-\mathbf{a}_i$, $k=j-i+i(i-1)/2$, $1\le i<j\le d_1$. Assume that a real number α_k is assigned to each edge $\mathbf{e}_k$, $k=1,\ldots,n_d$. In this section we shall construct a constant tensor metric $\mathfrak{M}$ on Δ such that

$$c_1|\Delta|_{\mathfrak{M}}^{2/d}\le\sum_{k=1}^{n_d}|\alpha_k|\le c_2|\partial\Delta|_{\mathfrak{M}}^2, \tag{1}$$

where constants c_1, c_2 depend only on d. Here, $|\Delta|_{\mathfrak{M}}$ and $|\partial\Delta|_{\mathfrak{M}}$ denote the volume and the perimeter (sum of edge lengths) of the simplex Δ in the metric $\mathfrak{M}$,

$$|\Delta|_{\mathfrak{M}} = (\det\mathfrak{M})^{1/2}|\Delta|,\quad |\partial\Delta|_{\mathfrak{M}} = \sum_{k=1}^{n_d}(\mathfrak{M}\mathbf{e}_k,\mathbf{e}_k)^{1/2}.$$

Let $u_2\in P_2(\Delta)$ be a quadratic function with the Hessian H_2 such that $u_2(\mathbf{a}_i)=0$, $i=1,\ldots,d_1$, and $u_2(\mathbf{c}_k)=-\frac{\alpha_k}{8}$, where $\mathbf{c}_k=(\mathbf{a}_i+\mathbf{a}_j)/2$ denotes the mid-point of $\mathbf{e}_k$, $k=1,\ldots,n_d$. The explicit form of u_2 will be given later. The trace of u_2 on $\mathbf{e}_k$ is a quadratic function w_2 vanishing at endpoints $\mathbf{a}_i$, $\mathbf{a}_j$ of $\mathbf{e}_k$ with an extremum at $\mathbf{c}_k$. Therefore, $w_2'(\mathbf{c}_k)=0$ and $\nabla u_2(\mathbf{c}_k)\cdot\mathbf{e}_k=0$. Applying the multi-point Taylor formula [17,18] for u_2 at endpoints $\mathbf{a}_i$, $\mathbf{a}_j$ of $\mathbf{e}_k$:

$$0=u_2(\mathbf{a}_i)=u_2(\mathbf{c}_k)-\frac{1}{2}\nabla u_2(\mathbf{c}_k)\cdot\mathbf{e}_k+\frac{1}{8}(H_2\mathbf{e}_k,\mathbf{e}_k) \tag{2}$$

$$0=u_2\left(\mathbf{a}_j\right)=u_2(\mathbf{c}_k)+\frac{1}{2}\nabla u_2(\mathbf{c}_k)\cdot\mathbf{e}_k+\frac{1}{8}(H_2\mathbf{e}_k,\mathbf{e}_k)$$

we obtain

$$\alpha_k=(H_2\mathbf{e}_k,\mathbf{e}_k).$$

The Hessian H_2 may be not positive definite and hence may not be used to define the metric $\mathfrak{M}$. In order to make it positive semidefinite, we take the spectral module of H_2:

$$|H_2| = W^T|\Lambda|W, \tag{3}$$

where $H_2=W^T\Lambda W$ is the spectral decomposition of the symmetric matrix H_2.

The degeneracy of the matrix $|H_2|$ is controlled by $detH_2$. If $detH_2\ne 0$, we set $\mathfrak{M}=|H_2|$.

Lemma 1. *Let α_k, $k=1,\ldots,n_d$ generate the quadratic function u_2, $u_2(\mathbf{a}_i)=0$, $i=1,\ldots,d_1$, with Hessian satisfying $(H_2\mathbf{e}_k,\mathbf{e}_k)=\alpha_k$ and $detH_2\ne 0$. Then for $\mathfrak{M}=|H_2|$ the estimate (1) holds with*

$$c_1=2\left(\frac{(d+1)(d+2)}{d!}\right)^{-\frac{1}{d}},\quad c_2=1.$$

Proof. We denote $H=H_2$. Since

$$|\partial\Delta|_{|H|}^2=\left(\sum_{k=1}^{n_d}(|H|\mathbf{e}_k,\mathbf{e}_k)^{1/2}\right)^2\ge\sum_{k=1}^{n_d}(|H|\mathbf{e}_k,\mathbf{e}_k)\ge\sum_{k=1}^{n_d}|(H\mathbf{e}_k,\mathbf{e}_k)| = \sum_{k=1}^{n_d}|\alpha_k|,$$

we have $c_2=1$.

To estimate c_1, we generalize the Cayley-Menger determinant to the case $H\ne I$:

$$detH|\Delta|^2=\frac{(-1)^{d-1}}{2^d(d!)^2}detK(H) \tag{4}$$

where

$$K(H)=\begin{pmatrix}(H\mathbf{a}_{11},\mathbf{a}_{11}) & \cdots & (H\mathbf{a}_{1d_1},\mathbf{a}_{1d_1}) & 1\\ \vdots & \ddots & \vdots & \vdots\\ (H\mathbf{a}_{d_11},\mathbf{a}_{d_11}) & \cdots & (H\mathbf{a}_{d_1d_1},\mathbf{a}_{d_1d_1}) & 1\\ 1 & \cdots & 1 & 0\end{pmatrix}, \tag{5}$$

$\mathbf{a}_{ij}\equiv\mathbf{a}_i-\mathbf{a}_j$. Setting $\mathbf{1}=(1,\ldots,1)^T\in\mathfrak{R}^{d_1}$, $k_{i,j}=(H\mathbf{a}_{ij},\mathbf{a}_{ij})$ and denoting by $[k_{i,j}]$ the matrix with entries $k_{i,j}$ we rewrite $K(H)$:

$$K(H)=\begin{pmatrix}[k_{i,j}] & \mathbf{1}\\ \mathbf{1}^T & 0\end{pmatrix}.$$

Since $H=H^T$, we have

$$k_{i,j}=(H\mathbf{a}_i,\mathbf{a}_i)+\left(H\mathbf{a}_j,\mathbf{a}_j\right)-2\left(H\mathbf{a}_i,\mathbf{a}_j\right).$$

Let the ith row of the matrix V be equal to $\mathbf{a}_i^T$, $i=1,\ldots,d_1$. Then

$$VHV^T=\left[\left(H\mathbf{a}_i,\mathbf{a}_j\right)\right].$$

In the following sequence of equalities we exploit the known dependence of a matrix determinant on linear operations with rows and columns:

$$\begin{aligned}\frac{(-1)^{d-1}}{2^d}detK(H) &= \frac{(-1)^{d-1}}{2^d}det\begin{pmatrix}\left[k_{i,j}-(H\mathbf{a}_i,\mathbf{a}_i)-\left(H\mathbf{a}_j,\mathbf{a}_j\right)\right] & \mathbf{1}\\ \mathbf{1}^T & 0\end{pmatrix}\\ &= \frac{(-1)^{d-1}}{2^d}det\begin{pmatrix}\left[-2\left(H\mathbf{a}_i,\mathbf{a}_j\right)\right] & \mathbf{1}\\ \mathbf{1}^T & 0\end{pmatrix}\\ &= -det\begin{pmatrix}\left[\left(H\mathbf{a}_i,\mathbf{a}_j\right)\right] & \mathbf{1}\\ \mathbf{1}^T & 0\end{pmatrix}\\ &= -det\left(\begin{pmatrix}V & \mathbf{1} & 0\\ \mathbf{0}^T & 0 & 1\end{pmatrix}\begin{pmatrix}HV^T & \mathbf{0}\\ \mathbf{0}^T & 1\\ \mathbf{1}^T & 0\end{pmatrix}\right)\\ &= det(V\quad \mathbf{1})det\begin{pmatrix}HV^T\\ \mathbf{1}^T\end{pmatrix}\\ &= det\begin{pmatrix}\mathbf{a}_1^T & 1\\ \mathbf{a}_2^T-\mathbf{a}_1^T & 0\\ \vdots & \vdots\\ \mathbf{a}_{d_1}^T-\mathbf{a}_1^T & 0\end{pmatrix}det\begin{pmatrix}H\mathbf{a}_1 & H(\mathbf{a}_2-\mathbf{a}_1) & \cdots & H\left(\mathbf{a}_{d_1}-\mathbf{a}_1\right)\\ 1 & 0 & \cdots & 0\end{pmatrix}\\ &= det\begin{pmatrix}\mathbf{a}_2^T-\mathbf{a}_1^T\\ \vdots\\ \mathbf{a}_{d_1}^T-\mathbf{a}_1^T\end{pmatrix}det\left(H(\mathbf{a}_2-\mathbf{a}_1)\quad\cdots\quad H\left(\mathbf{a}_{d_1}-\mathbf{a}_1\right)\right)\\ &= detH(det(\mathbf{a}_2-\mathbf{a}_1\quad\cdots\quad\mathbf{a}_{d_1}-\mathbf{a}_1))^2=(d!)^2detH|\Delta|^2\end{aligned}$$

which proves Eq. (4).

Therefore,

$$\begin{aligned}|\Delta|_{|H|}^2 &= det|H||\Delta|^2=\frac{1}{2^d(d!)^2}detK(|H|)\\ &\le\frac{1}{2^d(d!)^2}\sup_{\boldsymbol{\alpha}\in\mathfrak{R}^{n_d}}\frac{|detK(H)|}{\max_{k=1,\ldots,n_d}|\alpha_k|^d}\left(\sum_{k=1}^{n_d}|\alpha_k|\right)^d.\end{aligned} \tag{6}$$

For square matrices of order N with elements $b_{i,j}$ it holds

$$|det[b_{i,j}]| \le |\sum_{\sigma} \prod_{i=1}^{N} b_{i,\sigma}| \le N! \max_{\sigma} |\prod_{i=1}^{N} b_{i,\sigma}|,$$

where the summation is performed over all possible permutations σ of matrix rows and columns. Since $|k_{i,j}| = |(H\mathbf{e}_k, \mathbf{e}_k)| = |\alpha_k|$, $1 \le i < j \le d_1$, from definition (5) we derive that $detK(H)$ is a homogeneous polynomial of degree d of $\boldsymbol{\alpha}$ and

$$\sup_{\boldsymbol{\alpha}\in\mathfrak{R}^{n_d}} \frac{|detK(H)|}{\max\limits_{k=1,\dots,n_d} |\alpha_k|^d} \le (d+2)! \sup_{\boldsymbol{\alpha}\in\mathfrak{R}^{n_d}} \frac{\max\limits_{k=1,\dots,n_d} |\alpha_k|^d}{\max\limits_{k=1,\dots,n_d} |\alpha_k|^d} \le (d+2)!.$$

Therefore, we conclude from inequality (6) that

$$|\Delta|^2_{|H|} \le \frac{1}{2^d} \frac{(d+1)(d+2)}{d!} \left(\sum_{k=1}^{n_d} |\alpha_k| \right)^d,$$

$$|\Delta|^{\frac{2}{d}}_{|H|} \le \frac{1}{2} \left(\frac{(d+1)(d+2)}{d!} \right)^{\frac{1}{d}} \sum_{k=1}^{n_d} |\alpha_k|,$$

which implies

$$c_1 = 2\left(\frac{(d+1)(d+2)}{d!} \right)^{-\frac{1}{d}}.$$

If $detH_2 = 0$, the Hessian H_2 may not be the basis for the metric $\mathfrak{M}$. In this case we *modify* the edge data specifying the quadratic function so that its Hessian is positive definite and estimate (1) is satisfied. For the sake of simplicity we restrict ourselves to the case $0 \le \alpha_1 \le \alpha_2 \le \ldots \le \alpha_{n_d}$ and $\alpha_{n_d} \ne 0$ (in the method presented in the next section non-negativity of α_k is guaranteed). We introduce the modified edge data

$$\tilde{\alpha}_k = \alpha_k, \quad k = 1, \dots, n_d - 1, \quad \tilde{\alpha}_{n_d} = (1+\delta)\alpha_{n_d}, \tag{7}$$

where $\delta \in]0, 1]$. Let $\tilde{u}_2(\delta) \in P_2(\Delta)$ be a quadratic function such that $\tilde{u}_2(\mathbf{a}_i) = 0$, $i = 1, \ldots, d_1$, $\tilde{u}_2(\mathbf{c}_k) = -\frac{\tilde{\alpha}_k}{8}$, $k = 1, \ldots, n_d$, and $\tilde{H}_2(\delta)$ be its Hessian. Due to Eqs. (4) and (5) $p(\delta) = \det\tilde{H}_2(\delta)$ is a polynomial of degree two. Since $p(0) = \det\tilde{H}_2(0) = detH_2 = 0$, there exists $\delta_0 \in]0,1]$ such that $\det\tilde{H}_2(\delta_0) \ne 0$. We set $\mathfrak{M} = |\tilde{H}_2(\delta_0)|$ and check

$$\sum_{k=1}^{n_d} |\alpha_k| \le \sum_{k=1}^{n_d} |\tilde{\alpha}_k| \le \sum_{k=1}^{n_d} (|\tilde{H}_2(\delta_0)|\mathbf{e}_k, \mathbf{e}_k)$$
$$\le \left(\sum_{k=1}^{n_d} (|\tilde{H}_2(\delta_0)|\mathbf{e}_k, \mathbf{e}_k)^{1/2} \right)^2 = |\Delta|^2_{\mathfrak{M}},$$

$$\sum_{k=1}^{n_d} |\alpha_k| \ge \frac{1}{2} \sum_{k=1}^{n_d} |\tilde{\alpha}_k| \ge \left(\frac{(d+1)(d+2)}{d!} \right)^{-\frac{1}{d}} |\Delta|^{\frac{2}{d}}_{\mathfrak{M}}.$$

Thus we suggested such a modification of edge data that the recovered metric satisfies estimate (1) and using Lemma 1 we proved the following theorem.

Theorem 2. *Let a sequence $0 \le \alpha_1 \le \ldots \le \alpha_{n_d}$, $\alpha_{n_d} \ne 0$, generate the quadratic function u_2, $u_2(\mathbf{a}_i) = 0, i = 1, \dots, d_1$, with non-singular Hessian satisfying $(H_2\mathbf{e}_k, \mathbf{e}_k) = \alpha_k$, $k = 1, \ldots, n_d - 1$, and $(H_2\mathbf{e}_{n_d}\mathbf{e}_{n_d})$ equal to α_{n_d} or $\tilde{\alpha}_{n_d}$ from (7). Then for $\mathfrak{M} = |H_2|$ the estimate (1) holds with*

$$c_1 = \left(\frac{(d+1)(d+2)}{d!} \right)^{-\frac{1}{d}}, \quad c_2 = 1. \tag{8}$$

3. Energy of the interpolation error for quadratic functions

In this section we derive a new edge-based representation of the energy norm of the interpolation error. The application of Theorem 2 will give us the geometric representation of the error (11).

On a d-simplex Δ we define d_1 linear functions λ_i through their values at the vertices: $\lambda_i(\mathbf{a}_j) = \delta_i^j$ and n_d quadratic bubble functions $b_k = \lambda_i\lambda_j$ associated with edges $\mathbf{e}_k = [\mathbf{a}_i, \mathbf{a}_j]$, $k = 1, \ldots, n_d$. Note that $b_k(\mathbf{c}_k) = 1/4$, $b_k(\mathbf{c}_{k_1}) = 0$, $k_1 \ne k$, $b_k(\mathbf{a}_i) = 0$, $i = 1, \ldots, d$.

The linear interpolation operator is defined as

$$i_\Delta u = \sum_{i=1}^{d_1} u(\mathbf{a}_i)\lambda_i$$

and the error of linear interpolation of a quadratic function u_2 is

$$e_2 = u_2 - i_\Delta u_2.$$

Any quadratic function u_2 may be represented through its values at $\mathbf{a}_i$, $i = 1, \ldots, d_1$, and $\mathbf{c}_k$, $k = 1, \ldots, n_d$:

$$u_2 = i_\Delta u_2 + 4\sum_{k=1}^{n_d} (u_2(\mathbf{c}_k) - i_\Delta u_2(\mathbf{c}_k))b_k.$$

In particular, the function from Lemma 1 is $u_2 = -\frac{1}{2}\sum_{k=1}^{n_d} \alpha_k b_k$. The error of linear interpolation of any $u_2 \in P_2(\Delta)$ is

$$e_2 = u_2 - i_\Delta u_2 = 4\sum_{k=1}^{n_d} (u_2(\mathbf{c}_k) - i_\Delta u_2(\mathbf{c}_k))b_k = -\frac{1}{2}\sum_{k=1}^{n_d} \gamma_k b_k$$

where $\gamma_k = -8(u_2(\mathbf{c}_k) - i_\Delta u_2(\mathbf{c}_k))$. Since

$$\nabla e_2 = -\frac{1}{2}\sum_{k=1}^{n_d} \gamma_k \nabla b_k,$$

we get

$$\|\nabla e_2\|^2_{L_2} = |\Delta|(B\boldsymbol{\gamma}, \boldsymbol{\gamma}),$$

where $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_{n_d})^T$ and the $n_d \times n_d$ matrix B has elements

$$B_{i,j} = \frac{1}{4|\Delta|} \int_\Delta \nabla b_i \cdot \nabla b_j \mathrm{d}x. \tag{9}$$

Hereafter we omit the domain of integration in notations of integral norms unless this lead to an ambiguity.

The gradient error is only a number; it does not provide any directional information. To recover this information, we split this error into n_d edge-based error estimates $\alpha_k \ge 0$ such that

$$\sum_{k=1}^{n_d} \alpha_k = (B\boldsymbol{\gamma}, \boldsymbol{\gamma}) \quad \text{and} \quad \alpha_k \sim |\gamma_k|, k = 1, \dots, n_d.$$

The last requirement stems from the equidistribution of $||e_2||_{L_\infty(\mathbf{e}_k)}$ over all edges of the simplex Δ [14]. Setting

$$\alpha_k = |\gamma_k|(B\boldsymbol{\gamma}, \boldsymbol{\gamma})\left(\sum_{k=1}^{n_d} |\gamma_k| \right)^{-1} \tag{10}$$

and using Theorem 2 we define the metric $\mathfrak{M}$ such that

$$c_1|\Delta||\Delta|^{\frac{2}{d}}_{\mathfrak{M}} \le \|\nabla e_2\|^2_{L_2} \le c_2|\Delta||\partial\Delta|^2_{\mathfrak{M}}. \tag{11}$$

Remark 1. In general, other selections of n_d non-negative numbers α_k satisfying $\sum_{k=1}^{n_d} \alpha_k = (B\boldsymbol{\gamma}, \boldsymbol{\gamma})$ are possible (see Remark 1, [14]). According to the numerical evidence, the choice (10) provides recovery of anisotropic tensor metrics and generation of adaptive anisotropic meshes.

4. Energy of the interpolation error for general function

In this section, we derive a geometric representation of the energy norm of the interpolation error based on the relaxed saturation assumption.

For $p\in]0,+\infty]$ we introduce a normed (quasi-normed for $0<p<1$) space with the norm (quasi-norm)

$$\|u\|^p_{W^1_p} = \|u\|^p_{L_p} + \|\nabla u\|^p_{L_p}.$$

We recall that for quasi-normed spaces the triangular inequality is modified:

$$\|\nabla(v+w)\|_{L_p} \le C_p\left(\|\nabla v\|_{L_p} + \|\nabla w\|_{L_p}\right), \quad C_p = max\left(1, 2^{\frac{1-p}{p}}\right), \tag{12}$$

which follows from $(x+y)^p \le x^p + y^p \le 2^{1-p}(x+y)^p$.

For any function $u\in C(\overline{\Delta})\cap W^1_p(\Delta)$ with $p\in]0,+\infty]$ we define its quadratic interpolant

$$i_{2,\Delta}u = i_\Delta u + 4\sum_{k=1}^{n_d}(u(\mathbf{c}_k) - i_\Delta u(\mathbf{c}_k))b_k.$$

The generalization of estimate (11) is based on *the saturation assumption*: There exists $0<q_\Delta<1$ such that

$$\|\nabla(u-i_{2,\Delta}u)\|_{L_p} \le q_\Delta\|\nabla(u-i_\Delta u)\|_{L_p}. \tag{13}$$

The hypothesis (13) implies *the relaxed saturation assumption*:

$$c_s\|\nabla(i_{2,\Delta}u-i_\Delta u)\|_{L_p} \le \|\nabla(u-i_\Delta u)\|_{L_p} \le C_s\|\nabla(i_{2,\Delta}u-i_\Delta u)\|_{L_p} \tag{14}$$

with $c_s = \frac{1}{2C_p}, C_s = \frac{C_p}{1-q_\Delta C_p}$.

We define

$$\begin{aligned}\gamma_k &= -8(u(\mathbf{c}_k)-i_\Delta u(\mathbf{c}_k)) \\ &= -8\left(i_{2,\Delta}u(\mathbf{c}_k)-i_\Delta i_{2,\Delta}u(\mathbf{c}_k)\right) \\ &= -8\left(i_{2,\Delta}u(\mathbf{c}_k)-i_\Delta u(\mathbf{c}_k)\right),\end{aligned} \tag{15}$$

compute the metric $\mathfrak{M}$ for $\alpha_k = \gamma_k$ and combine the inequalities (11) and (14) in order to estimate the interpolation error $e\equiv u-i_\Delta u$ in the following theorem.

Theorem 3. *Let* inequality (14) *hold true and the metric* $\mathfrak{M}$ *be built using* (15), (9) *and* (10). *Then*

$$c_s^2 c_1|\Delta|\,|\Delta|^{\frac{2}{d}}_{\mathfrak{M}} \le \|\nabla e\|^2_{L_2} \le C_s^2 c_2|\Delta|\,|\partial\Delta|^2_{\mathfrak{M}} \tag{16}$$

where

$$c_1 = \left(\frac{(d+1)(d+2)}{d!}\right)^{-\frac{1}{d}}, \quad c_2 = 1.$$

The geometric representation of the energy norm of the error (16) is not final since it contains measures in different metrics. This will be corrected by a simple re-scaling of the metric $\mathfrak{M}$ (22) discussed in Section 6.

Although the saturation assumption is conventional in numerical analysis [19], its usage may be argued. We note that our analysis is based on the relaxed saturation assumption (14) rather than the saturation assumption (13).

5. Justification of the relaxed saturation assumption

In this section we justify the relaxed saturation assumption (14) for general (possibly anisotropic) simplexes. The motivation is based on the oscillation term studied in [5] in the context of an a posteriori error analysis.

We define a quadratic function $g = \frac{1}{2}(Gx, x)$ where $\mathcal{G}$ is a matrix from the space $\mathcal{G}$ of symmetric $d\times d$-matrices. Since $i_{2,\Delta}g = g$, we write

$$u-i_{2,\Delta}u = u-i_{2,\Delta}u-\left(g-i_{2,\Delta}g\right) = (u-g)-i_{2,\Delta}(u-g).$$

We denote the Hessians of u and g by H and G, respectively. By virtue of the multi-point Taylor formula [2,18]

$$\begin{aligned}\nabla\left(u-i_{2,\Delta}u\right)(\mathbf{x}) &= \nabla\left(u-g-i_{2,\Delta}(u-g)\right)(\mathbf{x}) \\ &= -\frac{1}{2}\sum_{j=1}^{d_1+n_d}\left(\left(H\left(\zeta\left(\mathbf{x},\mathbf{s}_j\right)\right)-G\right)\left(\mathbf{x}-\mathbf{s}_j\right),\mathbf{x}-\mathbf{s}_j\right)\nabla\left(p_j\right),\end{aligned}$$

where p_j, $j=1,\ldots,d_1+n_d$, are the basis functions for quadratic Lagrangian interpolation with nodes $\mathbf{s}_j=\mathbf{a}_j$, $j=1,\ldots,d_1$, $\mathbf{s}_{k+d_1}=\mathbf{c}_k$, $k=1,\ldots,n_d$, respectively, satisfying

$$\begin{aligned}\|\nabla(p_j)\|_{L_p} &\le C|\Delta|^{\frac{1}{p}}\max_{1\le i\le d_1}\|\nabla\lambda_i\|_{L_\infty} \le C|\Delta|^{\frac{1}{p}}\max_{1\le i\le d_1}\text{dist}^{-1}(\mathbf{a}_i,f_i) \\ &\le 2c(d)|\Delta|^{\frac{1}{p}}\frac{|\partial\Delta|^{d-1}}{|\Delta|}.\end{aligned}$$

Here $\text{dist}(\mathbf{a}_i, f_i)$ denotes the distance between $\mathbf{a}_i$ and the opposite face f_i.

Since $\mathbf{x}-\mathbf{s}_j = \sum_{j=1}^{d_1+n_d}\delta_j(\mathbf{x})\mathbf{e}_j$ with $|\delta_j|\le 1$, we derive

$$\begin{aligned}\sum_{j=1}^{d_1+n_d}\left(\left(H\left(\zeta\left(\mathbf{x},\mathbf{s}_j\right)\right)-G\right)\left(\mathbf{x}-\mathbf{s}_j\right),\mathbf{x}-\mathbf{s}_j\right) &\le c(d)\sum_{j=1}^{d_1+n_d}\left(H_G\mathbf{e}_j,\mathbf{e}_j\right) \\ &\le c(d)|\partial\Delta|^2_{H_G},\end{aligned}$$

where

$$H_G = |H(\zeta(\hat{\mathbf{x}},\hat{\mathbf{s}}))-G|, \quad (\hat{\mathbf{x}},\hat{\mathbf{s}}) = \arg\max_{\mathbf{x}\in\Delta,\mathbf{s}\in\Delta}(|H(\zeta(\mathbf{x},\mathbf{s}))-G|(\mathbf{x}-\mathbf{s}),\mathbf{x}-\mathbf{s}).$$

In order to emphasize the extremum features of H_G, we re-denote

$$|\partial\Delta|_{|H-G|_{\infty,\Delta}} := |\partial\Delta|_{H_G}.$$

Therefore,

$$\|\nabla(u-i_{2,\Delta}u)\|_{L_p} \le C(d)|\Delta|^{1/p}\frac{|\partial\Delta|^{d-1}}{|\Delta|}|\partial\Delta|^2_{|H(\mathbf{x})-G|_{\infty,\Delta}}.$$

We define the oscillation term

$$\text{osc}(H,\Delta)_p = C(d)|\Delta|^{1/p}\frac{|\partial\Delta|^{d-1}}{|\Delta|}\inf_{G\in\mathcal{G}}|\partial\Delta|^2_{|H(\mathbf{x})-G|_{\infty,\Delta}}. \tag{17}$$

Taking $v = i_{2,\Delta}u-i_\Delta u$, $w = u-i_{2,\Delta}u$ and using the triangular inequality (12), we obtain

$$\begin{aligned}\|\nabla(u-i_\Delta u)\|_{L_p} &\le C_p\left(\|\nabla(i_\Delta u-i_{2,\Delta}u)\|_{L_p} + \|\nabla(u-i_{2,\Delta}u)\|_{L_p}\right) \\ &\le C_p\left(\|\nabla(i_\Delta u-i_{2,\Delta}u)\|_{L_p} + \text{osc}(H,\Delta)_p\right).\end{aligned} \tag{18}$$

Similar use of the triangular inequality leads us to

$$\begin{aligned}\|\nabla(i_\Delta u-i_{2,\Delta}u)\|_{L_p} &\le C_p\left(\|\nabla(u-i_\Delta u\|_{L_p} + \|\nabla(u-i_{2,\Delta}u)\|_{L_p}\right) \\ &\le C_p\left(\|\nabla(u-i_\Delta u)\|_{L_p} + \text{osc}(H,\Delta)_p\right),\end{aligned}$$

which implies

$$C_p^{-1}\|\nabla(i_\Delta u - i_{2,\Delta}u)\|_{L_p} - \mathrm{osc}(H,\Delta)_p \le \|\nabla(u - i_\Delta u)\|_{L_p}. \quad (19)$$

Thus, we proved the following lemma.

Lemma 4. *Estimate* (14) *holds with* $c_s = C_p^{-1}$, $C_s = C_p$ *up to the oscillation term* (17).

The value of $\mathrm{osc}(H,\Delta)_p$ is small for $p \le 1$, $u \in C^2(\overline{\Delta})$ and small $|\partial\Delta|$. Moreover, for arbitrary $p \in]0, +\infty]$ and $u \in C^2(\overline{\Delta})$ we have

$$\inf_{G\in\mathcal{G}} |\partial\Delta|^2_{|H(\mathbf{x})-G|_{\infty,\Delta}} \le C \inf_{G\in\mathcal{G}} |H-G|_{\infty,\Delta}|\partial\Delta|^2,$$

$$\mathrm{osc}(H,\Delta)_p \le c(d)\frac{|\partial\Delta|^{d+1}}{|\Delta|^{\frac{p-1}{p}}} \inf_{G\in\mathcal{G}} |H-G|_{\infty,\Delta}$$

and the value of $\mathrm{osc}(H,\Delta)_p$ is small in simplices satisfying

$$\frac{|\partial\Delta|^{d+1}}{|\Delta|^{\frac{p-1}{p}}} \inf_{G\in\mathcal{G}} |H-G|_{\infty,\Delta} = o(1).$$

For instance, for shape regular simplices we have $|\partial\Delta|^d \le C|\Delta|$ and

$$\frac{|\partial\Delta|^{d+1}}{|\Delta|^{\frac{p-1}{p}}} \le C|\partial\Delta|\,|\Delta|^{\frac{1}{p}},$$

$$\mathrm{osc}(H,\Delta)_p \le C|\partial\Delta|\,|\Delta|^{\frac{1}{p}} \inf_{G\in\mathcal{G}} |H-G|_{\infty,\Delta}.$$

6. Gradient error of interpolation in general spaces L_p

In the previous sections, we considered the energy norm of the error corresponding to $p = 2$. The relaxed saturation assumption as well as its justification were discussed for general positive p. In this section we generalize Theorem 3 to the case of $p \in]0, +\infty]$ and derive the final geometric representation of the gradient of the interpolation error in Theorem 6.

Lemma 5. *For any* $p \in]0, +\infty]$ *and any non-negative* $v \in P_2(\Delta)$ *it holds*

$$C_{1/p}^{-\frac{1}{p}}|\Delta|^{\frac{1}{p}-1}\|v\|_{L_1} \le \|v\|_{L_p} \le C_p|\Delta|^{\frac{1}{p}-1}\|v\|_{L_1} \quad (20)$$

with

$$\begin{cases} C_p = 1 & \text{if } 0<p\le 1,\\ C_p = (d+1)(d+2)(d!)^{\frac{1}{p}}\left(\prod_{j=1}^{d}(p+j)\right)^{-\frac{1}{p}} & \text{if } 1<p<+\infty,\\ C_\infty = \lim_{p\to+\infty} C_p = (d+1)(d+2), & \\ C_{1/\infty} = \lim_{p\to+\infty} C_{1/p} = 1. & \end{cases}$$

Proof. First we prove the right inequality (20).

Let $p \in]0,1[$. We estimate $\|v\|^p_{L_p}$ using Hölder's inequality with $s = p^{-1} > 1$ and $r = (1-p)^{-1}$ for which $s^{-1} + r^{-1} = 1$:

$$\|v\|^p_{L_p} = \int_\Delta v^p dx \le |\Delta|^{1-p}\|v\|^p_{L_1}$$

that is

$$\|v\|_{L_p} \le C_p|\Delta|^{\frac{1}{p}-1}\|v\|_{L_1}.$$

For $p = 1$ the last estimate is trivial.

Let $p \in]1, +\infty]$. We present $0 \le v = \sum_{i=1}^{d+1} a_i\lambda_i + \sum_{k=1}^{n_d} c_k b_k$ with some $a_i \ge 0$, $c_k \ge 0$. Then

$$\|v\|_{L_1} = \int_\Delta v dx = \frac{|\Delta|}{d+1}\sum_{i=1}^{d+1} a_i + \frac{|\Delta|}{(d+1)(d+2)}\sum_{k=1}^{n_d} c_k$$

and

$$\sum_{i=1}^{d+1} a_i + \sum_{k=1}^{n_d} c_k \le \frac{(d+1)(d+2)}{|\Delta|}\|v\|_{L_1}.$$

Since $\forall p>1$, $1 \le i \le d+1$, $1 \le k \le n_d$ it holds

$$\|b_k\|^p_{L_p} \le \|\lambda_i\|^p_{L_p} = d!|\Delta| / \prod_{j=1}^{d}(p+j),$$

we derive

$$\|v\|_{L_p} \le \left(d!|\Delta| / \prod_{j=1}^{d}(p+j)\right)^{\frac{1}{p}}\left(\sum_{i=1}^{d+1} a_i + \sum_{k=1}^{n_d} c_k\right) \le C_p|\Delta|^{\frac{1}{p}-1}\|v\|_{L_1}.$$

In order to show the left inequality (20) for $p \in]0, +\infty[$, we set $w = v^{1/q}$ and write

$$\|v\|^{1/q}_{L_1} = \|w\|_{L_q} \le C_q|\Delta|^{\frac{1}{q}-1}\|w\|_{L_1} = C_q|\Delta|^{\frac{1}{q}-1}\|v\|^{1/q}_{L_{1/q}}$$

which implies $\forall q>0$

$$\|v\|_{L_1} \le C_q^q|\Delta|^{1-q}\|v\|_{L_{1/q}}$$

and for $p = 1/q$

$$C_{1/p}^{-\frac{1}{p}}|\Delta|^{\frac{1}{p}-1}\|v\|_{L_1} \le \|v\|_{L_p}.$$

For $p = +\infty$ we define $C_{1/\infty} = \lim_{p\to+\infty} C_{1/p} = 1$ and derive

$$C_{1/\infty}|\Delta|^{-1}\|v\|_{L_1} = \lim_{p\to+\infty} C_{1/p}^{-\frac{1}{p}}|\Delta|^{-1}\|v\|_{L_1} \le \lim_{p\to+\infty}|\Delta|^{-\frac{1}{p}}\|v\|_{L_p} = \|v\|_{L_\infty}.$$

Now we consider the error of linear interpolation of a quadratic function $e_2 = u_2 - i_\Delta u_2$. Since the function

$$v(x) = \sum_{j=1}^{d}\left(\frac{\partial e_2}{\partial x_j}\right)^2$$

is quadratic, we can apply Lemma 5

$$\|\nabla e_2\|_{L_p} = \|v\|^{1/2}_{L_{p/2}} \le C^{1/2}_{p/2}|\Delta|^{\frac{1}{p}-\frac{1}{2}}\|v\|^{1/2}_{L_1} = C^{1/2}_{p/2}|\Delta|^{\frac{1}{p}-\frac{1}{2}}\|\nabla e_2\|_{L_2}. \quad (21)$$

Let $\mathfrak{M}$ be the metric generated by Theorem 2 and let the scaled metric be

$$\mathfrak{M}_p = (\det\mathfrak{M})^{-\frac{1}{d+p}}\mathfrak{M} \quad (22)$$

for which it holds

$$|\Delta|^{\frac{1}{p}}|\partial\Delta|_{\mathfrak{M}} = |\Delta|^{\frac{1}{p}}_{\mathfrak{M}_p}|\partial\Delta|_{\mathfrak{M}_p}, \qquad |\Delta|^{\frac{1}{p}}|\Delta|^{\frac{1}{d}}_{\mathfrak{M}} = |\Delta|^{\frac{1}{p}+\frac{1}{d}}_{\mathfrak{M}_p}.$$

By virtue of inequalities (11) and (21)

$$\|\nabla e_2\|_{L_p} \le C_{p/2}^{1/2}|\Delta|^{\frac{1}{p}-\frac{1}{2}}\|\nabla e_2\|_{L_2} \le \left(C_{p/2}c_2\right)^{1/2}|\Delta|^{\frac{1}{p}}|\partial\Delta|_{\mathfrak{M}} = \left(C_{p/2}c_2\right)^{1/2}|\Delta|_{\mathfrak{M}_p}^{\frac{1}{p}}|\partial\Delta|_{\mathfrak{M}_p}. \quad (23)$$

Using the same arguments, we can apply the second half of Lemma 5:

$$\|\nabla e_2\|_{L_p} = \|v\|_{L_{p/2}}^{1/2} \ge C_{2/p}^{-1/p}|\Delta|^{\frac{1}{p}-\frac{1}{2}}\|v\|_{L_1}^{1/2} = C_{2/p}^{-1/p}|\Delta|^{\frac{1}{p}-\frac{1}{2}}\|\nabla e_2\|_{L_2}. \quad (24)$$

In view of estimates (11) and (24)

$$\|\nabla e_2\|_{L_p} \ge C_{2/p}^{-1/p}c_1^{1/2}|\Delta|^{\frac{1}{p}}|\Delta|_{\mathfrak{M}}^{\frac{1}{d}} = C_{2/p}^{-1/p}c_1^{1/2}|\Delta|_{\mathfrak{M}_p}^{\frac{1}{p}+\frac{1}{d}}. \quad (25)$$

For $p=+\infty$ we have $\mathfrak{M}_\infty=\mathfrak{M}$ and use

$$\|w\|_{L_\infty} = \lim_{p\to+\infty}\left(\frac{1}{|\Delta|}\int_\Delta w^p dx\right)^{\frac{1}{p}} = \lim_{p\to+\infty}|\Delta|^{-\frac{1}{p}}\|w\|_{L_p}$$

which yields

$$\|\nabla e_2\|_{L_\infty} = \lim_{p\to+\infty}|\Delta|^{-\frac{1}{p}}\|\nabla e_2\|_{L_p} \le \lim_{p\to+\infty}|\Delta|^{-\frac{1}{p}}(C_{p/2}c_2)^{1/2}|\Delta|_{\mathfrak{M}_p}^{\frac{1}{p}}|\partial\Delta|_{\mathfrak{M}_p} = \lim_{p\to+\infty}(C_{p/2}c_2)^{1/2}|\partial\Delta|_{\mathfrak{M}_p} = (C_\infty c_2)^{1/2}|\partial\Delta|_{\mathfrak{M}_\infty}, \quad (26)$$

$$\|\nabla e_2\|_{L_\infty} = \lim_{p\to+\infty}|\Delta|^{-\frac{1}{p}}\|\nabla e_2\|_{L_p} \ge \lim_{p\to+\infty}|\Delta|^{\frac{1}{p}}C_{2/p}^{-1/p}c_1^{1/2}|\Delta|_{\mathfrak{M}_p}^{\frac{1}{p}+\frac{1}{d}} = c_1^{1/2}|\Delta|_{\mathfrak{M}_\infty}^{\frac{1}{d}}. \quad (27)$$

Applying estimates (14), (23), (25), (26) and (27) we prove the following theorem.

Theorem 6. *Let the relaxed saturation assumption* (14) *hold true and the metric $\mathfrak{M}_p$ be built using.* (15), (9), (10) *and* (22). *Then for any $u\in C(\overline{\Delta})\cap W_p^1(\Delta)$, $p\in]0,+\infty]$*

$$c_sC_{2/p}^{-1/p}c_1^{1/2}|\Delta|_{\mathfrak{M}_p}^{\frac{1}{p}+\frac{1}{d}} \le \|\nabla(u-i_\Delta u)\|_{L_p} \le C_s(C_{p/2}c_2)^{1/2}|\Delta|_{\mathfrak{M}_p}^{\frac{1}{p}}|\partial\Delta|_{\mathfrak{M}_p}. \quad (28)$$

7. Gradient error on optimal and quasi-optimal meshes

In this section we take advantage of the local analysis summarized in Theorem 6 and present the asymptotic analysis of interpolation errors on optimal and quasi-optimal meshes.

Let $\Omega\in\Re^d$ be a polyhedral domain and Ω^h be its conformal d-simplicial partitioning into $\mathcal{N}(\Omega^h)$ cells (elements). Let $C(\overline{\Omega})$ denote the space of continuous functions over $\overline{\Omega}$, and $P_1(\Omega^h)$ denote the space of continuous piecewise linear functions, and let $\mathcal{P}_{\Omega^h}: C(\overline{\Omega})\to P_1(\Omega^h)$ be the linear interpolation operator.

Definition 1. Let $p\in]0,+\infty]$ and $u\in C(\overline{\Omega})\cap W_p^1(\Omega)$ be given. A mesh $\Omega_{opt}^h(N_T,u)$ consisting of at most N_T elements is called optimal if it is a solution of the optimization problem

$$\Omega_{opt}^h(N_T,u) = \arg\min_{\Omega^h:\mathcal{N}(\Omega^h)\le N_T}\|\nabla(u-\mathcal{P}_{\Omega^h}u)\|_{L_p(\Omega)}. \quad (29)$$

Although the existence of the optimal mesh is not known, there exist meshes providing error norms which are arbitrary close to the minimum in formula (29).

Theorem 7. *Let the optimal mesh $\Omega_{opt}^h(N_T,u)$ exist and $u\in C(\overline{\Omega})\cap W_p^1(\Omega)$ and $p\in]0,+\infty]$ are such that the relaxed saturation assumption* (14) *holds $\forall\Delta\in\Omega_{opt}^h(N_T,u)$ with $c_s<1$. Then there exists a tensor metric $\mathfrak{M}_p$, piecewise constant on Ω^h, such that*

$$c_sC_{2/p}^{-1/p}c_1^{1/2}|\Omega|_{\mathfrak{M}_p}^{\frac{1}{p}+\frac{1}{d}}N_T^{-\frac{1}{d}} \le \|\nabla(u-\mathcal{P}_{\Omega_{opt}^h}u)\|_{L_p(\Omega)}.$$

Proof. For $p\in]0,+\infty[$ *we use Hölder's inequality with* $s=1+\frac{p}{d}$ *and* $r=1+\frac{d}{p}(s^{-1}+r^{-1}=1)$ *to derive:*

$$|\Omega|_{\mathfrak{M}_p} = \sum_{\Delta\in\Omega_{opt}^h(N_T,u)}|\Delta|_{\mathfrak{M}_{p,\Delta}} \le \left(\sum_{\Delta\in\Omega_{opt}^h(N_T,u)}|\Delta|_{\mathfrak{M}_{p,\Delta}}^s\right)^{\frac{1}{s}}\left(\sum_{\Delta\in\Omega_{opt}^h(N_T,u)}1^r\right)^{\frac{1}{r}} = \left(\sum_{\Delta\in\Omega_{opt}^h(N_T,u)}|\Delta|_{\mathfrak{M}_{p,\Delta}}^{1+\frac{p}{d}}\right)^{\frac{d}{d+p}}\mathcal{N}(\Omega_{opt}^h)^{\frac{p}{d+p}}. \quad (30)$$

By virtue of Theorem 6 for any $\Delta\in\Omega_{opt}^h(N_T,u)$ there exists a tensor metric $\mathfrak{M}_{p,\Delta}$:

$$|\Delta|_{\mathfrak{M}_{p,\Delta}}^{1+\frac{p}{d}} \le c_s^{-p}C_{2/p}c_1^{-p/2}\|\nabla(u-\mathcal{P}_{\Omega_{opt}^h}u)\|_{L_p(\Delta)}^p. \quad (31)$$

Since $\mathcal{N}(\Omega_{opt}^h)\le N_T$, we get from inequalities (30) and (31)

$$|\Omega|_{\mathfrak{M}_p}N_T^{-\frac{p}{d+p}} \le \left(c_s^{-p}C_{2/p}c_1^{-p/2}\sum_{\Delta\in\Omega_{opt}^h(N_T,u)}\|\nabla(u-\mathcal{P}_{\Omega_{opt}^h}u)\|_{L_p(\Delta)}^p\right)^{\frac{d}{d+p}}$$

or

$$|\Omega|_{\mathfrak{M}_p}^{\frac{1}{p}+\frac{1}{d}}N_T^{-\frac{1}{d}} \le c_s^{-1}C_{2/p}^{1/p}c_1^{-1/2}\|\nabla\left(u-\mathcal{P}_{\Omega_{opt}^h}u\right)\|_{L_p(\Omega)}.$$

For $p=+\infty$ we use $\mathfrak{M}_\infty=\mathfrak{M}$ and estimate (28):

$$c_sc_1^{1/2}|\Omega|_{\mathfrak{M}}^{\frac{1}{d}} \le c_sc_1^{1/2}N_T^{\frac{1}{d}}\max_{\Delta\in\Omega_{opt}^h(N_T,u)}|\Delta|^{\frac{1}{d}} \le N_T^{\frac{1}{d}}\max_{\Delta\in\Omega_{opt}^h(N_T,u)}\|\nabla(u-i_\Delta u)\|_{L_\infty(\Delta)} = N_T^{\frac{1}{d}}\|\nabla\left(u-\mathcal{P}_{\Omega_{opt}^h}u\right)\|_{L_\infty(\Omega)}.$$

The optimal mesh is an ideal object which is not available. From a practical standpoint, it is sufficient to deal with meshes providing similar to optimal (albeit larger) gradient errors of interpolation. In particular, such meshes should demonstrate the optimal asymptotic rate of error reduction. We define such meshes as *quasi-optimal*. The next theorem shows that $\mathfrak{M}_p$-quasi-uniform meshes are quasi-optimal. In what follows we assume that the tensor metric $\mathfrak{M}_p$ is composed of elemental metrics $\mathfrak{M}_{p,\Delta}$ defined on each simplex by Theorem 6.

Definition 2. A conformal mesh Ω^h is called $\mathfrak{M}_p$-quasi-uniform, if there exist positive constants C_{sh}, C_{vl}:

$$|\partial\Delta|_{\mathfrak{M}_{p,\Delta}}^d \le C_{sh}|\Delta|_{\mathfrak{M}_{p,\Delta}}, \quad \forall\Delta\in\Omega^h, \quad (32)$$

$$\mathcal{N}(\Omega^h)\max_{\Delta\in\Omega^h}|\Delta|_{\mathfrak{M}_{p,\Delta}} \le C_{vl}|\Omega|_{\mathfrak{M}_p}, \quad \forall\Delta\in\Omega^h. \quad (33)$$

Theorem 8. *Let $u\in C(\overline{\Omega})\cap W_p^1(\Omega)$, $p\in]0,+\infty]$ and let Ω^h, $\mathcal{N}(\Omega^h)=N_T$, be a conformal $\mathfrak{M}_p$-quasi-uniform mesh such that the relaxed saturation assumption* (14) *holds $\forall\Delta\in\Omega^h$ with certain constants c_s, C_s. Then*

$$\|\nabla(u-\mathcal{P}_{\Omega^h}u)\|_{L_p(\Omega)} \le C_s\left(C_{p/2}c_2\right)^{1/2} C_{sh}^{\frac{1}{d}} C_{vl}^{\frac{1}{p}+\frac{1}{d}} N_T^{-\frac{1}{d}} |\Omega|_{\mathfrak{M}_p}^{\frac{1}{p}+\frac{1}{d}}. \quad (34)$$

Proof. By virtue of Theorem 6 and Definition 2 we have

$$\begin{aligned}\|\nabla(u-P_{\Omega^h}u)\|_{L_p(\Omega)} &= \left(\sum_{\Delta\in\Omega^h}\|\nabla(u-i_\Delta u)\|_{L_p(\Delta)}^p\right)^{\frac{1}{p}}\\ &\le C_s\left(C_{p/2}c_2\right)^{1/2}\left(\sum_{\Delta\in\Omega^h}|\Delta|_{\mathfrak{M}_{p,\Delta}}|\partial\Delta|_{\mathfrak{M}_{p,\Delta}}^{p}\right)^{\frac{1}{p}}\\ &\le C_s\left(C_{p/2}c_2\right)^{1/2}C_{sh}^{\frac{1}{d}}\left(\sum_{\Delta\in\Omega^h}|\Delta|_{\mathfrak{M}_{p,\Delta}}^{1+\frac{p}{d}}\right)^{\frac{1}{p}}\\ &\le C_s\left(C_{p/2}c_2\right)^{1/2}C_{sh}^{\frac{1}{d}}N_T^{\frac{1}{p}}\left(\max_{\Delta\in\Omega^h}|\Delta|_{\mathfrak{M}_{p,\Delta}}\right)^{\frac{1}{p}+\frac{1}{d}}\\ &\le C_s\left(C_{p/2}c_2\right)^{1/2}C_{sh}^{\frac{1}{d}}C_{vl}^{\frac{1}{p}+\frac{1}{d}}N_T^{-\frac{1}{d}}|\Omega|_{\mathfrak{M}_p}^{\frac{1}{p}+\frac{1}{d}}.\end{aligned}$$

8. From theory to algorithms

Theorem 8 gives the constructive description of quasi-optimal meshes: an optimal asymptotics of the gradient error of interpolation is achieved on $\mathfrak{M}_p$-quasi-uniform meshes. The piecewise constant metric $\mathfrak{M}_p$ may be recovered by Theorem 3 and scaling (22) on the basis of mid-edge interpolation data. This leads us to the adaptive iterative algorithm: a) given a current mesh, compute the metric $\mathfrak{M}_p$; b) given the metric $\mathfrak{M}_p$, generate an $\mathfrak{M}_p$-quasi-uniform mesh. After several loops of the algorithm we shall produce a mesh which will be quasi-uniform in the metric recovered on the same mesh. The algorithm is known to be practical for continuous tensor metric fields.

The above algorithm rises three technical issues:

1. What is the measure of $\mathfrak{M}_p$-quasi-uniformity?
2. How to produce a continuous metric $\mathfrak{M}_p$?
3. How to produce an $\mathfrak{M}_p$-quasi-uniform mesh?

Below we discuss these issues.

8.1. Measure of mesh quasi-uniformity

Given a metric $\mathfrak{M}_{p,\Delta}$ on a d-simplex Δ and a desirable simplex size h, we define the quality of Δ with respect to h as

$$Q_{\Delta,h} = \frac{d!(d(d+1))^d}{\sqrt{2^d(d+1)}}\frac{|\Delta|_{\mathfrak{M}_{p,\Delta}}}{|\partial\Delta|_{\mathfrak{M}_{p,\Delta}}^d}F\left(\frac{d(d+1)h}{2|\partial\Delta|_{\mathfrak{M}_{p,\Delta}}}\right),$$

where $F:\mathfrak{R}_+\to]0,1]$ is a smooth function such that $F(0)=0$, $F(1)=1$, $F'(x)>0$, $x\in]0,1[$ and $F(x)=F(x^{-1})$, $\forall x>0$. An example of such a function is

$$F=\sqrt{\min(x,x^{-1})(2-\min(x,x^{-1}))}.$$

The function F defines the size factor of the simplex quality $Q_{\Delta,h}$ since its maximum is attained on simplexes whose perimeter is equal to $n_d h$. The remaining factors of $Q_{\Delta,h}$ define the shape factor attaining its maximum (equal to one) on $\mathfrak{M}_p$-equilateral simplexes. Therefore, the maximal quality $Q_{\Delta,h}=1$ is achieved on $\mathfrak{M}_p$-equilateral simplexes Δ with $\mathfrak{M}_p$-length of edges h.

Now we define *the mesh quality*. Given a conformal mesh Ω^h, a metric $\mathfrak{M}_p$ on $\overline{\Omega}$ and a desirable number of mesh elements N_T, we define the mesh quality $Q(\Omega^h, N_T)$ as

$$Q(\Omega^h,N_T)=\min_{\Delta\in\Omega^h}Q_{\Delta,h},\qquad h=\left(\frac{2^{\frac{d}{2}}d!\,|\Omega|_{\mathfrak{M}_p}}{N_T\sqrt{d+1}}\right)^{\frac{1}{d}}.$$

The parameter h is chosen to be the $\mathfrak{M}_p$-length of edge of an $\mathfrak{M}_p$-equilateral d-simplex whose $\mathfrak{M}_p$-volume is $|\Omega|_{\mathfrak{M}_p}/N_T$.

Since the shape factor and the size factor of $Q_{\Delta,h}$ do not exceed 1, we conclude

$$0\le Q\left(\Omega^h,N_T\right)\le 1.$$

Now we show that if $Q(\Omega^h, N_T)\ge\underline{Q}>0$ then inequalities (32) and (33) hold with constants C_{sh}, C_{vl} dependent on F, $\underline{Q}$ and d only. Indeed, since $F(x)\le 1$, for any $\Delta\in\Omega^h$ we have

$$\underline{Q}\le\frac{d!(d(d+1))^d}{\sqrt{2^d(d+1)}}\frac{|\Delta|_{\mathfrak{M}_{p,\Delta}}}{|\partial\Delta|_{\mathfrak{M}_{p,\Delta}}^d},$$

$$|\partial\Delta|_{\mathfrak{M}_{p,\Delta}}^d\le\frac{d!(d(d+1))^d}{\underline{Q}\sqrt{2^d(d+1)}}|\Delta|_{\mathfrak{M}_{p,\Delta}},$$

i.e., $C_{sh}=\frac{d!(d(d+1))^d}{\underline{Q}\sqrt{2^d(d+1)}}$. On the other hand, since the shape factor does not exceed 1, for any $\Delta\in\Omega^h$ we have

$$\underline{Q}\le F\left(\frac{d(d+1)h}{2|\partial\Delta|_{\mathfrak{M}_{p,\Delta}}}\right)$$

which implies

$$z\le\frac{d(d+1)h}{2|\partial\Delta|_{\mathfrak{M}_{p,\Delta}}}\le z^{-1},\qquad z=F^{-1}(\underline{Q})\le 1.$$

From this and the boundedness of the shape factor we derive

$$|\Delta|_{\mathfrak{M}_{p,\Delta}}\le\frac{\sqrt{2^d(d+1)}}{d!(d(d+1))^d}|\partial\Delta|_{\mathfrak{M}_{p,\Delta}}^d\le\frac{|\Omega|_{\mathfrak{M}_p}}{z^dN_T},$$

i.e., $C_{vl}=(z(F,\underline{Q}))^{-d}$.

Therefore, the mesh quality $Q(\Omega^h, N_T)$ is a good measure of $\mathfrak{M}_p$-quasi-uniformity.

8.2. Generation of $\mathfrak{M}_p$-quasi-uniform meshes

As it was mentioned earlier, numerical evidence suggests the use of a continuous metric within the adaptation loop. Continuous metrics provide faster convergence and the resulted meshes are more smooth. However, in Section 2 we considered the metric recovery element-by-element which implies a *discontinuous* tensor metric field. The simplest way to define a continuous metric is to assume that metric entries are continuous piecewise linear functions specified at mesh nodes. However, the nodal metric entries may not be defined via independent recovery of respective piecewise constant functions: such an approach may produce degenerate or non-definite matrices. We suggest a simple method of nodal metric recovery. For each node $\mathbf{a}_i$ of Ω^h we define the superelement σ_i as the union of all d-simplices sharing $\mathbf{a}_i$ and assign the metric with the largest determinant from all metrics available in superelement σ_i. The method takes always the worst metric in the vicinity of the node.

We generate $\mathfrak{M}_p$-quasi-uniform meshes as follows. Assume that an initial mesh and a continuous piecewise linear tensor metric $\mathfrak{M}_p$ are given and that the quality of that mesh is small. The basic strategy for the generation of a $\mathfrak{M}_p$-quasi-uniform mesh with a desirable number of elements is to modify the mesh using local operations which increase the mesh quality. The list of local operations includes moving, adding and deleting mesh nodes, and swapping of mesh edges and faces [1,2,8]. Since the mesh quality is equal to the quality of the worst simplex, the local mesh modifications are applied to this simplex. The local nature of topological operations makes the algorithm robust at least for $d=2,3$ [1,2]. Two- and three-dimensional implementations of this method are available in packages Ani2D and Ani3D [20,21], respectively, developed by K. Lipnikov and Yu. Vassilevski.

9. Numerical experiments

In this section, we examine numerically asymptotic properties of quasi-optimal meshes. We study the interpolation problem for two two-dimensional functions: weakly anisotropic and strongly anisotropic. Twenty steps of the adaptation loop were performed for each run. The minimal computed error was chosen to be the numerical result. The mesh quality was maintained at 0.5 within the adaptation cycle.

In the first example, we consider the problem of minimizing the gradient error of interpolation for the function [22]

$$u(x_1,x_2) = \frac{(x_1-0.5)^2-\left(\sqrt{10}x_2+0.2\right)^2}{\left((x_1-0.5)^2+\left(\sqrt{10}x_2+0.2\right)^2\right)^2}$$

defined over the unit square $[0,1]^2$. The function has a weak anisotropic singularity at the point $\left(0.5,-0.2/\sqrt{10}\right)$ which is outside the computational domain but close to its boundary. Isolines of u are shown in Fig. 1, (left). Fig. 1 also shows the quasi-optimal meshes for $N_T=2500$ for two values of p: the middle picture corresponds to the case $p=1$, the right picture corresponds to the case $p=+\infty$. Table 1 presents the L_p norms of the gradient error of interpolation $\|\nabla(u-\mathcal{P}_{\Omega^h}u)\|_{L_p(\Omega)}$ for $p=1,2,4,+\infty$. Large values of the error are attributable to the large gradient of u, $\|\nabla u\|_{L_\infty(\Omega)}=790.6$.

We observe the correct asymptotics of the error reduction:

$$\|\nabla(u-P_{\Omega^h}u)\|_{L_p(\Omega)}\sim N_T^{-1/2} \quad (35)$$

for $p=1,2,4,+\infty$.

In the second experiment, we build the quasi-optimal meshes for the function proposed in [23]:

$$u(x_1,x_2) = x_2x_1^2+x_2^3+tanh(6(sin(5x_2)-2x_1)).$$

Table 1
Experiment 1: L_p-norms of the gradient error of interpolation for different p.

$N_T\backslash p$	$+\infty$	4	2	1
600	79	10.1	3.9	1.29
2500	40	4.9	1.9	0.65
10,000	23	2.5	1.0	0.34
40,000	14	1.23	0.51	0.17

The computational domain is the square $[-1,1]^2$. The solution is anisotropic along the zigzag curve (see left picture in Fig. 2) and changes sharply in the direction normal to this curve. Table 2 shows the L_p-norms of the gradient error of interpolation. The dependence (35) is clearly observed.

In the middle picture in Fig. 2 we present the quasi-optimal mesh minimizing $\|\nabla(u-\mathcal{P}_{\Omega^h}u)\|_{L_\infty(\Omega)}$ with $N_T=2500$. For the sake of comparison, in the right picture in Fig. 2 we show a quasi-optimal mesh minimizing $\|u-\mathcal{P}_{\Omega^h}u\|_{L_\infty(\Omega)}$.

10. Conclusion

We analyzed the optimal meshes minimizing $L_p(\Omega)$-norms of the gradient interpolation error over all simplicial meshes with a fixed number of cells N_T. The analysis is given for functions satisfying the relaxed saturation assumption and is performed for arbitrary $p\in]0,+\infty]$ in spaces of arbitrary dimension d. The error norms are estimated by the product of factors depending on p, d, Ω, N_T. The explicit forms of the factors are derived.

We also presented and analyzed the new method of recovery of the tensor metric field $\mathfrak{M}_p$ used in the adaptation procedure. We proved that $\mathfrak{M}_p$-quasi-uniform meshes are quasi-optimal. This implies that they give slightly higher errors but the same asymptotic rate of error reduction as the optimal mesh. The error norms on quasi-optimal meshes are estimated with an explicit dependence on p, d, Ω, N_T, and the quality of the mesh $Q(\Omega^h,N_T)$. Our metric recovery is based on the new edge-based estimator of the interpolation error. This estimator is shown to be reliable and efficient for general functions provided that the relaxed saturation assumption holds true. The latter may be derived from the classical saturation assumption. Nevertheless we prove the relaxed saturation assumption up to the oscillation term which is small on a wide class of fine meshes.

We discussed practical implications of the developed theory. In particular, we presented the method of construction of a continuous tensor metric field from piecewise constant metric recovered elementwise. Also we explained our approach to the generation of $\mathfrak{M}_p$-quasi-uniform meshes with a prescribed number of elements.

Two-dimensional numerical experiments confirmed the predicted asymptotic rates of the error reduction for different p. Quasi-optimal

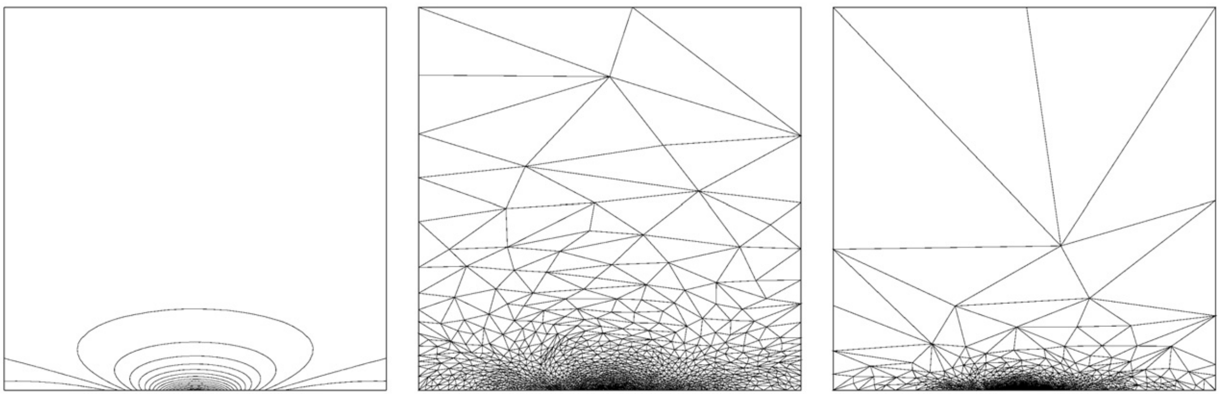

Fig. 1. Isolines of function u from Experiment 1 (left), quasi-optimal mesh for $p=1$ (center), quasi-optimal mesh for $p=+\infty$ (right); $N_T=2500$.

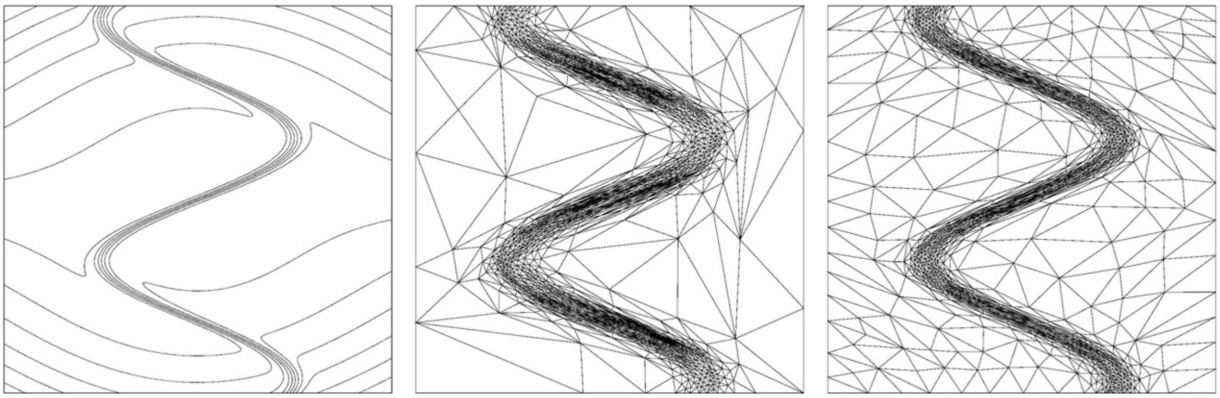

Fig. 2. Isolines of function u from Experiment 2 (left), quasi-optimal meshes minimizing $\|\nabla(u-P_{\Omega^h}u)\|_{L_\infty(\Omega)}$ (center) and $\|u-P_{\Omega^h}u\|_{L_\infty(\Omega)}$ (right); $N_T=2500$.

Table 2
Experiment 2: L_p-norms of the gradient error of interpolation.

$N_T \setminus p$	$+\infty$	4	2	1
600	8.9	2.5	2.0	2.1
2500	4.4	1.1	0.88	0.95
10,000	2.4	0.54	0.44	0.47
40,000	1.3	0.27	0.22	0.23

meshes were generated for both weakly and strongly anisotropic functions.

Our method of quasi-optimal mesh generation may be extended to the solution of boundary value problems [15].

Acknowledgments

This research was partly supported by the Russian Foundation for Basic Research through grant 08-01-00159-a and by RAS program “Optimal methods for problems of mathematical physics”.

The authors thank K. Lipnikov for fruitful discussions and useful remarks.

References

[1] Y. Vassilevski, K. Lipnikov, Adaptive algorithm for generation of quasi-optimal meshes, Comp. Math. Math. Phys. 39 (1999) 1532–1551.
[2] A. Agouzal, K. Lipnikov, Y. Vassilevski, Adaptive generation of quasi-optimal tetrahedral meshes, East-West J. Numer. Math. 7 (1999) 223–244.
[3] Y. Vassilevski, A. Agouzal, Unified asymptotic analysis of interpolation errors for optimal meshes, Doklady Mathematics 72 (2005) 879–882.
[4] L. Chen, P. Sun, J. Xu, Optimal anisotropic meshes for minimizing interpolation errors in L_p-norm, Math.Comp. 76 (2007) 179–204.
[5] W. Dörfler, R. Nochetto, Small data oscillation implies the saturation assumption, Numer. Math. 91 (2002) 1–12.
[6] W. Huang, Metric tensors for anisotropic mesh generation, J. Comp.Phys. 204 (2005) 633–665.
[7] W. Huang, Mathematical principles of anisotrpic mesh adaptation, Comm. Comp. Phys. 1 (2006) 276–310.
[8] G. Buscaglia, E. Dari, Anisotropic mesh optimization and its application in adaptivity, Inter. J. Numer. Meth. Engrg. 40 (1997) 4119–4136.
[9] M. Castro-Diaz, F. Hecht, B. Mohammadi, O. Pironneau, Anisotropic unstructured mesh adaptation for flow simulations, Int. J. Numer. Meth. Fluids 25 (1997) 475–491.
[10] K. Lipnikov, Y. Vassilevski, Parallel adaptive solution of 3D boundary value problems by Hessian recovery, Comp.Methods Appl. Mech. Eng. 192 (2003) 1495–1513.
[11] L. Formaggia, S. Perotto, Anisotropic error estimates for elliptic problems, Numer. Math. 94 (2003) 67–92.
[12] G. Kunert, An a posteriori residual error estimator for the finite element method anisotropic tetrahedral meshes, Numer. Math. 86 (2000) 471–490.
[13] G. Kunert, *A posteriori error estimation for anisotropic tetrahedral and triangular finite element meshes*. Ph.D. Thesis, TU Chemnitz, 1999.
[14] A. Agouzal, K. Lipnikov, Vassilevski, Hessian-free metric-based mesh adaptation via geometry of interpolation error, Comp. Math. Math. Phys. 50 (1) (2010) 124–138.
[15] A. Agouzal, K. Lipnikov, Y. Vassilevski. Anisotropic mesh adaptation for solution of finite element problems using hierarchical edge-based error estimates, *Proceedings of 18th International Meshing Roundtable*, Springer (B.W. Clark Editor), 2009, 595–610).
[16] P. Binev, R. DeVore, Fast computation in adaptive tree approximation, Numer. Math. 97 (2004) 193–217.
[17] P. Ciarlet, C. Wagschal, Multipoint Taylor formulas and applications to the finite element method, Numer. Math. 17 (1971) 84–100.
[18] P. Ciarlet, The finite element method for elliptic problems, North-Holland, 1978.
[19] R. Verfurth, A review of a posteriori error estimation and adaptive mesh-refinement techniques, Wiley-Teubner, Stuttgart, Germany, 1996.
[20] http://sourceforge.net/projects/ani2d.
[21] http://sourceforge.net/projects/ani3d.
[22] E. D'Azevedo, Optimal triangular mesh generation by coordinate transformation, SIAM J. Sci. Stat. Comput. 12 (1991) 755–786.
[23] F. Hecht, A few snags in mesh adaptation loops, Proceedings of the 14th International Meshing Roundtable, Springer-Verlag, Berlin, 2005, 301–313.