Anisotropic hp-mesh optimization technique based on the continuous mesh and error models*

Vít Dolejší[1], Georg May[2], Filip Roskovec[1] and Pavel Solin[3]

[1] Charles University, Prague, Faculty of Mathematics and Physics, Czech Republic

[2] Aachen Institute for Advanced Study in Computational Engineering Science, RWTH Aachen, Germany

[3] Department of Mathematics and Statistics, University of Nevada in Reno, USA

Abstract

We develop a new mesh adaptive technique for the numerical solution of partial differential equations (PDEs) using the *hp*-version of the finite element method (*hp*-FEM). The technique uses a combination of approximation and interpolation error estimates to generate anisotropic triangular elements as well as appropriate polynomial approximation degrees. We present a *hp*-version of the continuous mesh model as well as the continuous error model which are used for the formulation of a mesh optimization problem. Solving the optimization problem leads to *hp*-mesh with the smallest number of degrees of freedom, under the constraint that the approximate solution has an error estimate below a given tolerance. Further, we propose an iterative algorithm to find a suitable anisotropic *hp*-mesh in the sense of the mesh optimization problem. Several numerical examples demostrating the efficiency and applicability of the new method are presented.

Keyword: *hp*-methods; anisotropic mesh adaptation; continuous mesh model; continuous error model; mesh optimization
MSC: 65N50; 65N15; 65D05

1 Introduction

This paper deals with the development of adaptive numerical methods for the solution of various partial differential equations (PDEs) describing practical problems. Adaptive methods which automatically enhance the functional space where the approximate solution is sought, can significantly reduce the computational cost. We employ triangular *anisotropic hp-adaptive* grids where the corresponding mesh elements are generally anisotropic (thin and long) and for each triangle, a different polynomial approximation degree is used generally. By $\mathscr{T}_{hp}$, we denote a *hp*-mesh, which consists of the a triangular grid itself (the set of non-overlapping triangles covering the computational domain) and the set of the integers corresponding to each triangle denoting the local polynomial approximation degree. Although the presented technique was designed for the *discontinuous Galerkin method*, we suppose, that after a small modification, it can be used for any other finite element method allowing varying polynomial degrees.

Anisotropic *hp*-grids offer substantial flexibility in mesh design. They can efficiently discretize the computational domain near edge singularities, interior and boundary layers, etc. Roughly speaking, the presented techniques combine the *(isotropic) hp-adaptive methods*, which allow the adaptation in the element size h as well as in the polynomial degree of approximation p, (cf. [1, 2, 3, 4]), and the *anisotropic mesh adaptation* techniques, which generate triangles having the shape extended in one dominant direction, cf. [5, 6, 7, 8]. The anisotropy of elements can be described by a Riemann metric represented by a 2×2 symmetric positive definite matrix. The anisotropy of elements is usually set in order to minimize the interpolation error of the exact solution. The works mentioned above deal mostly with piecewise linear approximation, thus the Hessian matrix (=matrix of the second order derivatives) is naturally employed for the definition of the Riemann metric generating the anisotropy of grids. Furthermore, in [9, 10, 11, 12], optimal grids for fixed higher polynomial degree of approximation (> 1) were derived.

For the given *hp*-grid $\mathscr{T}_{hp}$, we define the space of discontinuous piecewise polynomial functions S_{hp},where the approximate solution is sought. Our general aim is to develop an algorithm which generates a suitable *hp*-grid $\mathscr{T}_{hp}$ (and therefore the corresponding space S_{hp}) such that

(i) the computational error of the numerical solution from S_{hp} computed on $\mathscr{T}_{hp}$ is under the given tolerance,

*This work was supported by grant No. 13-00522S of the Czech Science Foundation and by the Fulbright Visiting Program.

(ii) the number of *degrees of freedom* (= dimension of S_{hp}) is small, in the ideal case as small as possible.

In [13] we developed a somewhat heuristic adaptive technique seeking an "optimal" mesh in the sense mentioned above. There, the *interpolation error* was considered in the condition (i). Moreover, the optimality of the resulting *hp*-mesh was not guaranteed. In [14], we presented a technique which seeks a suitable anisotropic *hp*-grid with respect to the adjoint-based error estimates. This approach uses several heuristic thresholds for the adaptation and the mesh optimality is again quite open.

In this paper, we introduce the continuous mesh and error models which allow us to employ a variational calculus for the seeking of the optimal *hp*-mesh in the sense mentioned above. The framework of continuous mesh model based on an interpolation error estimate for piecewise linear approximation was introduced in [15, 16]. In [17], the continuous mesh model was extended to higher order polynomial approximation taking into account the interpolation error estimates from [13]. Let us also mention [18], deriving the *hp*-adaptation framework for quadrilateral grids where elements can be split onto two daughter elements by a line in either the vertical or the horizontal directions.

The novel ideas of the present paper may be summarized as follows. Firstly, we define the *hp*-version of the *continuous mesh model*, which is formed by a quartet of functions defined on the computational domain. Secondly, in a generalization of the method presented in [17], several more or less standard assumptions on the chosen error estimator allow us to introduce the *continuous error model* described by two additional functions. Using both continuous models, we formulate the *hp-mesh optimization problem*. Unfortunately, due to its complexity, we have no guarantee of the existence of its solution for a general error estimator and we are unable to solve it at this moment. However, using a technique similar to [17], we are able to solve the *h*-version of the optimization problem, whose consequence is the *equi-distribution principle* of the error estimate. Using this property, we propose an algorithm, which iteratively seeks a possible solution of the *hp*-version of the mesh optimization problem. Numerical experiments show that this algorithm converges, but we are not able to guarantee how close the limit *hp*-grid is to the hypothetical optimal one.

The content of the rest of the paper is the following. In Section 2, we introduce the basic notations and properties of (discrete) anisotropic *hp*-meshes. In Section 3, we introduce the *hp*-version of the continuous mesh model. In Section 4, we recall some results concerning the anisotropic interpolation error estimates from [13, 19]. In Section 5, we deal with the formulation and the solution of the *h*-version of the mesh optimization problem, while the extension to the *hp*-version is treated in Section 6, where we also introduce the final algorithm. Several technical aspects for practical computations are mentioned in Section 7. The numerical experiments demonstrating the efficiency and applicability of the presented technique are given in Section 8.

2 Discrete setting: *hp*-mesh and error estimate

In this section we introduce the usual discrete setting for the *hp*-method as a domain discretization, the corresponding finite element space of functions and an abstract error estimate. Finally, we introduce the main goal of our effort. By $\mathbb{R}$ and $\mathbb{N}$ we denote the set of real and positive integer numbers, respectively. Further, we put $\mathbb{R}^+ := (0,\infty)$, $\mathbb{R}_0^+ := [0,\infty)$ and $\mathbb{R}_1^+ := [1,\infty)$. Let $\Omega \subset \mathbb{R}^2$ be a bounded polygonal computational domain with a polygonal boundary $\partial\Omega$. For simplicity, we assume that Ω is convex, however this assumption can be relaxed.

2.1 *hp*-mesh

By $\mathscr{T}_h = \{K\}$ ($h > 0$) we denote a conforming *triangulation* of Ω with standard finite element properties, see, e.g., [20], and $|K|$ is the area (= 2D Lebesgue measure) of $K \in \mathscr{T}_h$.

Definition 2.1. *Let $\mathscr{T}_h = \{K\}$ be a triangulation of Ω. To each $K \in \mathscr{T}_h$, we assign a positive integer p_K (=local polynomial approximation degree on K). We put $\mathbf{p} := \{p_K;\ K \in \mathscr{T}_h\}$. The pair $\mathscr{T}_{hp} := \{\mathscr{T}_h, \mathbf{p}\}$ is called the hp-*mesh.

For the given *hp*-mesh $\mathscr{T}_{hp}$, there exists the space of piecewise polynomial discontinuous functions given by

$$S_{hp} := \{v \in L^2(\Omega);\ v|_K \in P^{p_K}(K)\ \forall K \in \mathscr{T}_h\}, \tag{1}$$

where $P^{p_K}(K)$ is the space of polynomials of degree $\le p_K$ on $K \in \mathscr{T}_h$. The dimension of S_{hp} is

$$D_{hp} := \sum_{K\in\mathscr{T}_h} (p_K+1)(p_K+2)/2. \tag{2}$$

We call D_{hp} the *number of degrees of freedom* (DoF) of the *hp*-mesh $\mathscr{T}_{hp}$.

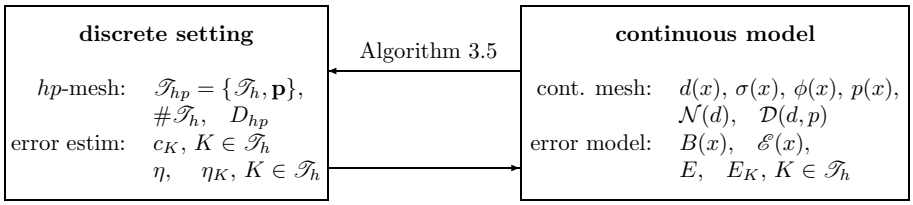

Fig. 1: Relation between the discrete settings and the continuous mesh model.

2.2 Error estimator

Let us consider a *boundary value problem* (BVP) with the unknown exact solution $u : \Omega \to \mathbb{R}$, $u \in V$, where V is a suitable space of functions. Let $u_h \in S_{hp}$ be the approximate solution of BVP computed by a suitable numerical method, e.g., the discontinuous Galerkin method. We are interested either in the solution u_h itself or in the *quantity of interest* represented by a linear functional $J : V \cup S_{hp} \to \mathbb{R}$. The functional J can be, e.g., a (regularized) value at a given point in Ω or a boundary/volume integral over a given subset of Ω. In the former case, we need to estimate the error $\|u - u_h\|_X$, where $\|\cdot\|_X$ is a suitable norm. In the latter case, we have to estimate the error $J(u) - J(u_h)$. The former case can be also considered in the framework of the latter case [21, Section 4]. For simplicity, we denote the error by e_h for both cases.

Our aim is to generate an *hp*-mesh having the minimal possible number of DoF such that the error e_h computed on this mesh is under the given tolerance. Let us assume that we have an error estimator η estimating e_h which satisfies the following assumptions.

Assumption 2.2. *Let $\mathscr{T}_{hp}$ be a hp-mesh, e_h the corresponding error of the quantity of interest and $\eta \in \mathbb{R}$ its estimator satisfying the following conditions:*

(A1) η is computable from the approximate solution u_h,

(A2) The error estimate may be written $\eta^2 = \sum_{K \in \mathscr{T}_h} \eta_K^2$, where η_K is the error estimator on K computable from u_h on a neighbourhood of K,

(A3) $\eta \approx e_h$.

The assumptions (A1)–(A3) are roughly standard and they are valid for any reasonable error estimate.

2.3 Main goal of the mesh adaptation

Our aim is to achieve the prescribed error tolerance using the smallest possible number of degree of freedom. Hence we formulate the following task.

Problem 2.3. *Let $\omega > 0$ be the given tolerance. We seek a hp-mesh $\mathscr{T}_{hp} = \{\mathscr{T}_h, \mathbf{p}\}$ such that*

(p1) $\eta = (\sum_{K \in \mathscr{T}_h} \eta_K^2)^{1/2} \le \omega$,

(p2) D_{hp} (=number of DoF of $\mathscr{T}_{hp}$) is minimal,

where η, η_K, $K \in \mathscr{T}_h$ are the error estimate of $u_h \in S_{hp}$ computed on the hp-mesh $\mathscr{T}_{hp}$ (cf. Assumption 2.2).

This problem is easy to define but very complex to solve. One way, how to treat it, is to employ the continuous mesh and error models which are introduced in the next sections.

3 Continuous anisotropic *hp*-mesh model

In the previous section, we introduced the standard *discrete setting*, namely the *hp*-mesh $\mathscr{T}_{hp}$ and error estimates η and η_K. Now we present the corresponding *continuous model* where the discrete quantities are represented by continuous functions defined on Ω, cf. Figure 1. Let us note that having a continuous mesh model, we can easily construct the corresponding (discrete) *hp*-mesh using Algorithm 3.5 presented below. On the other hand, the setting of the continuous model from the discrete setting is rather heuristic.

The continuous model has two parts: the *continuous hp-mesh model* (Section 3.2) and the *continuous error model* (Section 5 and Section 6). Before the definition of the continuous *hp*-mesh model, we are going to describe the anisotropy of a triangular element. The anisotropy is usually given by 3 real parameters. We use a different (but equivalent) description in comparison with [13] or [17].

3.1 Anisotropy of element

We denote the set of 2×2 symmetric and positive definite matrices by

$$\text{Sym} := \left\{ \mathbb{M} = \{m_{ij}\}_{i,j=1}^{2} \in \mathbb{R}^{2\times 2};\ m_{12} = m_{21},\ x^T \mathbb{M} x > 0\ \forall x \in \mathbb{R}^2,\ x \neq 0 \right\}, \tag{3}$$

where x^{T} is transpose of the column vector $x = (x_1, x_2)$ and $x^{\mathrm{T}}\mathbb{M}x := m_{11}x_1^2 + 2m_{12}x_1x_2 + m_{22}x_2^2$.

Let K be a triangle, not necessary isosceles. We describe its anisotropy by the circumscribed ellipse Σ_K having the minimal possible area, see Figure 2 for an illustration. Such ellipse is called the *Steiner ellipse* (see, e.g., [22]), its center coincides with the barycentre of K and the areas of K and Σ_K satisfy the relations

$$|K| = c_a r_{K,1} r_{K,2} = c_a \frac{|\Sigma_K|}{\pi}, \tag{4}$$

where $r_{K,1}$ and $r_{K,2}$ are the lengths of the semi-axes of the ellipse Σ_K and $c_a = 3\sqrt{3}/4$. The Steiner ellipse Σ_K is given by the relation

$$\Sigma_K := \left\{ x \in \mathbb{R}^2;\ x^{\mathrm{T}} \mathbb{M}_K x \leq 1 \right\}, \tag{5}$$

where $\mathbb{M}_K = \{m_{ij}\}_{i,j=1}^{2} \in \text{Sym}$ and its coefficients of $\mathbb{M}$ are the solution of the following linear algebraic system

$$\begin{pmatrix} (x_1^A - x_1^B)^2 & 2(x_1^A - x_1^B)(x_2^A - x_2^B) & (x_2^A - x_2^B)^2 \\ (x_1^B - x_1^C)^2 & 2(x_1^B - x_1^C)(x_2^B - x_2^C) & (x_2^B - x_2^C)^2 \\ (x_1^C - x_1^A)^2 & 2(x_1^C - x_1^A)(x_2^C - x_2^A) & (x_2^C - x_2^A)^2 \end{pmatrix} \begin{pmatrix} m_{11} \\ m_{12} \\ m_{22} \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix}, \tag{6}$$

where (x_1^A, x_2^A), (x_1^B, x_2^B) and (x_1^C, x_2^C) are the Cartesian coordinates of the vertices of K. Moreover, we recall that the eigenvalue decomposition of $\mathbb{M}_K \in \text{Sym}$

$$\mathbb{M}_K = \mathbb{Q}_{\phi_K}^{\mathrm{T}} \begin{pmatrix} \lambda_{K,1} & 0 \\ 0 & \lambda_{K,2} \end{pmatrix} \mathbb{Q}_{\phi_K}, \qquad \text{where} \quad \mathbb{Q}_\phi := \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}, \quad \phi \in [0, 2\pi), \tag{7}$$

gives $r_{K,i} = 1/\sqrt{\lambda_{K,i}}$, $i = 1, 2$ and ϕ_K is the angle between the major semi-axis of Σ_K and the axis x_1, see Figure 2, left. Now, we are able to present the following results. Let us note that the similar result was presented in [13] for isosceles triangles only.

Lemma 3.1. *Let K be a triangle, not necessary isosceles, Σ_K be the correspoding Steiner ellipse and $\mathbb{M}_K \in \text{Sym}$ with entries m_{ij}, $i, j = 1, 2$ satisfying* (6)*. Then*

$$\|\boldsymbol{e}_{K,i}\|_{\mathbb{M}_K} = \left(\boldsymbol{e}_{K,i}^{\mathrm{T}} \mathbb{M}_K \boldsymbol{e}_{K,i}\right)^{1/2} = \sqrt{3}, \quad i = 1, 2, 3, \tag{8}$$

where $\boldsymbol{e}_{K,i}$, $i = 1, 2, 3$ denote the edges of $K_{\mathbb{M}}$, which are considered as vectors from $\mathbb{R}^2$ given by their endpoints.

Proof. Equalities (8) follow directly from (6). □

Furthermore, we define the anisotropy of element K.

Definition 3.2. *Let K be a triangle and Σ_K its Steiner ellipse. Let $r_{K,1} > 0$, $r_{K,2} > 0$ ($r_{K,1} \geq r_{K,2}$) be the lengths of the semi-axis of Σ_K and let ϕ_K denote the angle between the major semi-axes of Σ_K and the axes x_1. We say that $\sigma_K := r_{K,2}/r_{K,1} \leq 1$ is the* aspect ratio *of K and ϕ_K is the* orientation *of K. The pair $\{\sigma_K,\ \phi_K\}$ is called the anisotropy of K. The area $|K|$ is further called the* size of element *K.*

Obviously, if K is an equilateral triangle then $\sigma_K = 1$ and $\phi_K \in [0, \pi)$ can be arbitrary since then Σ_K is a circle.

Remark 3.3. Let us note that for the given triangle, the corresponding Steiner ellipse is unique. On the other hand, a given ellipse can be the Steiner one for more triangles, see Figure 2, right, where an acute-angle and an obtuse triangles have the same Steiner ellipse and therefore the same anisotropy.

Finally, a simple consequence of the previous equalities is the relation between the pairs $\{\lambda_{K,1}, \lambda_{K,2}\}$ and $\{|K|, \sigma_K\}$ given by

$$\lambda_{K,1}\,\lambda_{K,2} = (c_a|K|)^{-2} \qquad \& \qquad \lambda_{K,1}/\lambda_{K,2} = \sqrt{\sigma_K}, \tag{9}$$

where the constant c_a was introduced in (4).

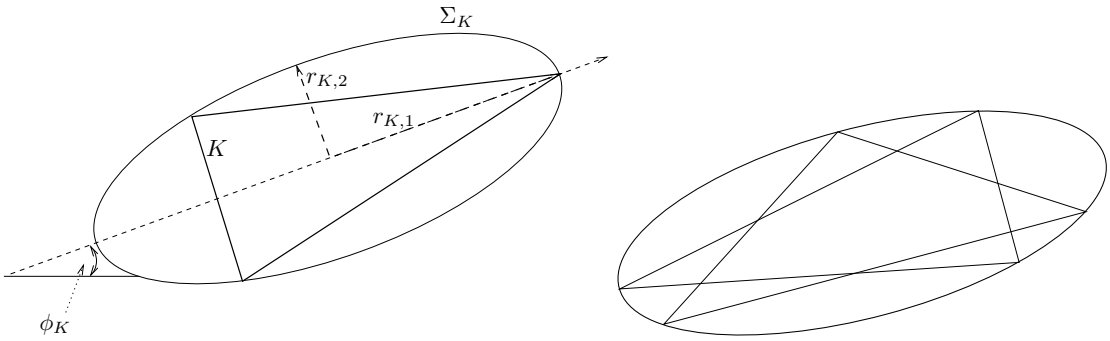

Fig. 2: Anisotropy of the triangle K given by the minimal possible circumscribed (Steiner) ellipse Σ_K with the length of semi-axes $r_{K,1}$, $r_{K,2}$ and the orientation ϕ_K (left). The acute-angle and the obtuse-angle anisotropic triangles having the same anisotropy (right).

3.2 Continuous anisotropic *hp*-mesh model

In this section we introduce the *hp*-version of the *continuous mesh framework* developed in [15, 16] for the piecewise linear approximation, which was extended in [17] to higher order polynomial approximation taking into account the interpolation error estimates from [13].

An anisotropic element of an *hp*-mesh (cf. Definitions 2.1 and 3.2), can be characterized by the parameters

$$|K| > 0, \quad \sigma_K \in (0,1], \quad \phi_K \in [0,\pi), \quad p_K \in \mathbb{N}, \qquad K \in \mathscr{T}_h. \tag{10}$$

In the *continuous mesh model* we replace these parameters with appropriate smooth functions.

Definition 3.4. *Let us consider the following continuous bounded functions:*

- $d : \Omega \to \mathbb{R}_0^+$ *called the* element-size distribution function*,*
- $\sigma : \Omega \to (0,1]$ *called the* aspect-ratio distribution function*,*
- $\phi : \Omega \to [0,\pi)$ *called the* orientation distribution function*,*
- $p : \Omega \to \mathbb{R}_1^+$ *called the* polynomial degree distribution function*.*

The quartet of functions $\{d,\ \sigma,\ \phi,\ p\}$ *is called the* continuous anisotropic *hp*-mesh model.

The meaning of the continuous anisotropic *hp*-mesh model is the following. Let $\mathscr{T}_{hp}$ be an *hp*-mesh then the corresponding continuous mesh model satisfies

$$|K| \approx d(x_K), \quad \sigma_K \approx \sigma(x_K), \quad \phi_K \approx \phi(x_K), \quad p_K \approx p(x_K), \quad x_K \in K,\ K \in \mathscr{T}_h. \tag{11}$$

Taking into account the usual techniques (e.g., [15, 17, 13]), we define the following three-steps algorithm which, for the given continuous anisotropic *hp*-mesh model, constructs the corresponding (discrete) *hp*-mesh $\mathscr{T}_{hp} = \{\mathscr{T}_h, \mathbf{p}\}$.

Algorithm 3.5. Let $\{d,\ \sigma,\ \phi,\ p\}$ be the given continuous anisotropic *hp*-mesh model.

(S1) In virtue of (7), we define the mapping $\mathcal{M} : \Omega \to \mathrm{Sym}$ by

$$\mathcal{M}(x) := \mathbb{Q}_{\phi(x)}^{\mathrm{T}} \mathrm{diag}(\lambda_1(x), \lambda_2(x)) \mathbb{Q}_{\phi(x)}, \qquad x \in \Omega, \tag{12}$$

where $\mathbb{Q}_{\phi(x)}$ is the rotation matrix (7) of its argument $\phi(x)$ and $\lambda_1(x), \lambda_2(x)$ are nonnegative functions satisfying relations (cf. (9))

$$\lambda_1(x)\,\lambda_2(x) = (c_a d(x))^{-2} \qquad \& \qquad \lambda_1(x)/\lambda_{2(x)} = \sqrt{\sigma(x)}, \qquad x \in \Omega. \tag{13}$$

This mapping $\mathcal{M}$ defines the Riemann metric in Ω, where the distance between $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ is given by

$$\|\boldsymbol{v}_2 - \boldsymbol{v}_1\|_{\mathcal{M}} := \int_0^1 \big((\boldsymbol{v}_2 - \boldsymbol{v}_1)^{\mathrm{T}} \mathcal{M}(\boldsymbol{v}_1 + t(\boldsymbol{v}_2 - \boldsymbol{v}_1))(\boldsymbol{v}_2 - \boldsymbol{v}_1)\big)^{1/2}\, \mathrm{d}t, \quad \boldsymbol{v}_1, \boldsymbol{v}_2 \in \Omega. \tag{14}$$

(S2) In virtue of (8), we generate a $\mathcal{M}$-uniform triangular mesh in the least square sense, namely we construct the triangulation $\mathscr{T}_h$

$$\mathscr{T}_h = \arg\min_{\mathscr{T}_h'} \sum_{\boldsymbol{e}\in\mathscr{F}_h'} \left(\|\boldsymbol{e}\|_{\mathcal{M}} - \sqrt{3}\right)^2, \tag{15}$$

where $\|\boldsymbol{e}\|_{\mathcal{M}}$ is the distance of the end points of $\boldsymbol{e}$ given by (14), the minimum is taken over all possible triangulations $\mathscr{T}_h'$ of Ω and $\mathscr{F}_h'$ is the set of all faces of $\mathscr{T}_h'$. Of course, such a triangulation is not unique. Let us note that there exist algorithms and codes, e.g., [23], [24], which construct mesh $\mathscr{T}_h$ for the given metric $\mathcal{M}$ in the sense of (15).

(S3) We define the polynomial approximation degrees p_K for each $K \in \mathscr{T}_h$ by

$$p_K := \operatorname{int}\left[\frac{1}{|K|}\int_K p(x)\,\mathrm{d}x\right], \quad K \in \mathscr{T}_h, \tag{16}$$

where $\operatorname{int}[a] := \lfloor a + 1/2 \rfloor$ denotes the integer part of the number $a + 1/2,\ a \geq 0$.

Remark 3.6. Algorithm 3.5 describes the construction of the discrete mesh from the continuous mesh model. We need also an opposite procedure, i.e., set the continuous model from the discrete mesh. The simplest way is to evaluate functions d, σ, ϕ and p at barycentres of all $K \in \mathscr{T}_h$ putting $d(x_K) := |K|$, $p(x_K) := p_K$ and the construction of the Steiner ellipse at x_K for the setting of $\sigma(x_K)$ and $\phi(x_K)$. Then these functions can be interpolated at the mesh vertexes (using some weighted averages from barycentres of elements sharing the given mesh vertex) and thus we have a continuous piecewise linear functions defining the continuous mesh model.

Furthermore, we introduce the *number of elements* and the *number of degrees of freedom* of the continuous *hp*-mesh model $\{d,\ \sigma,\ \phi,\ p\}$ by

$$\mathcal{N}(d) := \int_\Omega \frac{1}{d(x)}\,\mathrm{d}x \qquad \text{and} \qquad \mathcal{D}(d,p) := \int_\Omega \frac{(p(x)+1)(p(x)+2)}{2d(x)}\,\mathrm{d}x, \tag{17}$$

respectively. Taking into account (11), we find the correlation between the number of triangles and the number of degrees of freedom of the discrete and continuous mesh models. Namely,

$$\mathcal{N}(d) = \int_\Omega \frac{1}{d(x)}\,\mathrm{d}x = \sum_{K\in\mathscr{T}_h}\int_K \frac{1}{d(x)}\,\mathrm{d}x \approx \sum_{K\in\mathscr{T}_h}\int_K \frac{1}{|K|}\,\mathrm{d}x = \sum_{K\in\mathscr{T}_h} 1 = \#\mathscr{T}_h, \tag{18}$$

and, similarly,

$$\mathcal{D}(d,p) = \sum_{K\in\mathscr{T}_h}\int_K \frac{(p(x)+1)(p(x)+2)}{2d(x)}\,\mathrm{d}x \approx \sum_{K\in\mathscr{T}_h}\int_K \frac{(p_K+1)(p_K+2)}{2|K|}\,\mathrm{d}x = \sum_{K\in\mathscr{T}_h}\frac{(p_K+1)(p_K+2)}{2} = D_{hp}. \tag{19}$$

Let us note that the quantities $\mathcal{N}$ and $\mathcal{D}$ are independent of $\sigma(x)$ and $\phi(x)$ for the given continuous mesh model. We interpret the function $(p(x)+1)(p(x)+2)/(2d(x))$ as the *density of the number of degrees of freedom* of the continuous *hp*-mesh model.

Finally, let us discuss 3 special cases of the continuous anisotropic *hp*-mesh model:

- Continuous *isotropic hp-mesh model*: putting $\sigma(x) = 1,\ x \in \Omega$ and $\phi(x)$ arbitrary (e.g., $\phi(x) = 0,\ x \in \Omega$), we obtain the isotropic case. Thus Algorithm 3.5 generates triangulation consisting from (almost) equilateral triangles.
- Continuous *anisotropic h-mesh model*: putting $p(x) = \bar{p} = \text{const},\ x \in \Omega$ we obtain the *h*-version. Then Algorithm 3.5 generates anisotropic triangles and the same polynomial approximation degree for all $K \in \mathscr{T}_h$ of course.
- Continuous *isotropic h-mesh model*: putting $\sigma(x) = 1,\ x \in \Omega$, $\phi(x)$ arbitrary and $p(x) = \bar{p} = \text{const},\ x \in \Omega$, we obtain the isotropic *h*-version which is the combination of two previous special cases.

4 Anisotropic interpolation error estimates

In this section we briefly recall the basic results from [13, 19] concerning the optimal setting of the anisotropy of individual triangles with respect to the interpolation error. Let u be a sufficiently regular function (for simplicity let $u \in C^\infty(\overline{\Omega})$),

$\bar{x} = (\bar{x}_1, \bar{x}_2) \in \Omega$ and $\bar{p} \in \mathbb{N}$ be given. Let $\pi_{\bar{x},\bar{p}} u$ be the $\bar{p}$-degree polynomial approximation of u given by the Taylor expansion at $\bar{x}$. Then the error of this interpolation can be expressed as

$$u(x) - \pi_{\bar{x},\bar{p}} u(x) = e^{\mathrm{int}}_{\bar{x},\bar{p}}(x) + O(|x-\bar{x}|^{\bar{p}+2}), \tag{20}$$

where

$$e^{\mathrm{int}}_{\bar{x},\bar{p}}(x) := \frac{1}{(\bar{p}+1)!} \sum_{l=0}^{\bar{p}+1} \binom{\bar{p}+1}{l} \frac{\partial^{\bar{p}+1} u(\bar{x})}{\partial x_1^l \partial x_2^{\bar{p}+1-l}} (x_1 - \bar{x}_1)^l (x_2 - \bar{x}_2)^{\bar{p}+1-l}, \quad x = (x_1, x_2) \in \Omega \tag{21}$$

is the *interpolation error function* of degree p located at $\bar{x}$. Let us note that the right-hand side of (21) is the $(\bar{p}+1)^{\mathrm{th}}$-order scaled directional derivative of u at $\bar{x}$ along the direction $x - \bar{x}$. Obviously, $e^{\mathrm{int}}_{\bar{x},\bar{p}}(\bar{x}) = 0$ and $e^{\mathrm{int}}_{\bar{x},\bar{p}}(x) \approx u(x) - \pi_{\bar{x},\bar{p}} u(x)$ up to the higher order terms.

4.1 Shape-optimality with respect to the L^q-norm, $q \in [1, \infty]$

In [13, Lemma 3.12], we derived the estimate

$$\left|e^{\mathrm{int}}_{\bar{x},\bar{p}}(x)\right| \le A_{\bar{x},\bar{p}} \left((x-\bar{x})^{\mathrm{T}} \mathbb{Q}_{\varphi_{\bar{x},\bar{p}}} \mathbb{D}_{\rho_{\bar{x},\bar{p}}} \mathbb{Q}^{\mathrm{T}}_{\varphi_{\bar{x},\bar{p}}} (x-\bar{x}) \right)^{\frac{p+1}{2}} \ \forall\, x \in \Omega, \tag{22}$$

where $A_{\bar{x},\bar{p}} > 0$, $\mathbb{Q}_{\varphi_{\bar{x},\bar{p}}}$ is the rotation through angle $\varphi_{\bar{x},\bar{p}}$ given by (7) and $\mathbb{D}_{\rho_{\bar{x},\bar{p}}}$ is the matrix given by

$$\mathbb{D}_{\rho_{\bar{x},\bar{p}}} := \begin{pmatrix} 1 & 0 \\ 0 & \rho_{\bar{x},\bar{p}}{}^{-2/(\bar{p}+1)} \end{pmatrix}, \quad \rho_{\bar{x},\bar{p}} \ge 1. \tag{23}$$

The values $A_{\bar{x},\bar{p}} \ge 0$, $\rho_{\bar{x},\bar{p}} \ge 1$ and $\varphi_{\bar{x},\bar{p}} \in [0, 2\pi)$ represent the *size*, the *aspect ratio* and the *orientation* of the interpolation error function $e^{\mathrm{int}}_{\bar{x},\bar{p}}$, which are defined in such a way that the estimate (22) is as sharp as possible. For more details see [13, Section 3]. Let us note that the value $A_{\bar{x},\bar{p}}$ represents the maximal value of the $(\bar{p}+1)^{\mathrm{th}}$-order scaled directional derivative of u at $\bar{x}$ (cf. the right-hand side of (21)), $\varphi_{\bar{x},\bar{p}}$ is the angle of the direction of the maximal derivative and $\rho_{\bar{x},\bar{p}}$ corresponds to the ratio between $A_{\bar{x},\bar{p}}$ and the $(\bar{p}+1)^{\mathrm{th}}$-order scaled directional derivative along the perpendicular direction.

The estimate (22) is the base of the following Lemma, for the proof see [13, Lemma 3.17].

Lemma 4.1. *Let $u \in C^\infty(\overline{\Omega})$, $\bar{x} \in \Omega$, $\bar{p} \in \mathbb{N}$, $q \in [1, \infty]$ and $\bar{\omega} > 0$ be given. Let $\{A_{\bar{x},\bar{p}}, \varphi_{\bar{x},\bar{p}}, \rho_{\bar{x},\bar{p}}\}$ be the anisotropy of the corresponding interpolation error function $e^{\mathrm{int}}_{\bar{x},\bar{p}}$ from (22). Then the triangle $K_{\bar{x},\bar{p}}$ with barycenter $\bar{x}$, the size*

$$|K_{\bar{x},\bar{p}}| := c_a \left(\frac{\bar{\omega}\, (\rho_{\bar{x},\bar{p}})^{1/2}}{c_{\bar{p},q}\, A_{\bar{x},\bar{p}}} \right)^{2q/(q(\bar{p}+1)+2)} \quad \text{for } q \in [1,\infty), \qquad |K_{\bar{x},\bar{p}}| := c_a \left(\frac{\bar{\omega}\, (\rho_{\bar{x},\bar{p}})^{1/2}}{A_{\bar{x},\bar{p}}} \right)^{2/(\bar{p}+1)} \quad \text{for } q = \infty \tag{24}$$

where $c_a = 3\sqrt{3}/4$, $c_{\bar{p},q} := \big(\big(2(\pi)^{-q(\bar{p}+1)/2}\big) / \big(q(\bar{p}+1)+2\big) \big)^{1/q}$, $\pi = 3.1415\ldots$, and the anisotropy

$$\sigma_{\bar{x},\bar{p}} = \rho_{\bar{x},\bar{p}}{}^{1/(\bar{p}+1)}, \quad \phi_{\bar{x},\bar{p}} = \varphi_{\bar{x},\bar{p}} - \pi/2 \tag{25}$$

satisfies $\left\| e^{\mathrm{int}}_{\bar{x},\bar{p}} \right\|_{L^q(K_{\bar{x},\bar{p}})} \le \bar{\omega}$ and has the maximal possible area.

This lemma allows us to define the size and the anisotropy of the optimal triangle for the given polynomial degree and the given tolerance in the L^q-norm.

4.2 Shape-optimality with respect to the H^1-seminorm

In [19, relation (3.17)], we derived the estimate

$$\left|\nabla e^{\mathrm{int}}_{\bar{x},\bar{p}}(x)\right|^2 \le A_{\bar{x},\bar{p}} \left((x-\bar{x})^{\mathrm{T}} \mathbb{Q}_{\varphi_{\bar{x},\bar{p}}} \bar{\mathbb{D}}_{\rho_{\bar{x},\bar{p}}} \mathbb{Q}^{\mathrm{T}}_{\varphi_{\bar{x},\bar{p}}} (x-\bar{x}) \right)^{p} \ \forall\, x \in \Omega, \tag{26}$$

where $A_{\bar{x},\bar{p}} > 0$, $\mathbb{Q}_{\varphi_{\bar{x},\bar{p}}}$ is the rotation through angle $\varphi_{\bar{x},\bar{p}}$ given by (7) and $\bar{\mathbb{D}}_{\rho_{\bar{x},\bar{p}}}$ is the matrix given by

$$\bar{\mathbb{D}}_{\rho_{\bar{x},\bar{p}}} := \begin{pmatrix} 1 & 0 \\ 0 & \rho_{\bar{x},\bar{p}}{}^{-1/p} \end{pmatrix}, \quad \rho_{\bar{x},\bar{p}} \ge 1. \tag{27}$$

The values $A_{\bar{x},\bar{p}} \ge 0$, $\rho_{\bar{x},\bar{p}} \ge 1$ and $\varphi_{\bar{x},\bar{p}} \in [0, 2\pi)$ represent the *size*, the *aspect ratio* and the *orientation* of the *square of the magnitude of the gradient* of the interpolation error function $|\nabla e^{\mathrm{int}}_{\bar{x},\bar{p}}|^2$, which are defined in such a way that the estimate (26) is as sharp as possible. For more details see Ref. [19, Section 3]. Let us note that the values $A_{\bar{x},\bar{p}}$, $\rho_{\bar{x},\bar{p}}$ and $\varphi_{\bar{x},\bar{p}}$ differ form those in Section 4.1 but can be evaluated using a similar technique.

The estimate (26) is the base of the following Lemma, for the proof see [19, Theorem 3.1].

Lemma 4.2. *Let $u \in C^\infty(\overline{\Omega})$, $\bar{x} \in \Omega$, $\bar{p} \in \mathbb{N}$ and $\bar{\omega} > 0$ be given. Let $\{A_{\bar{x},\bar{p}}, \varphi_{\bar{x},\bar{p}}, \rho_{\bar{x},\bar{p}}\}$ be the anisotropy of the corresponding $|\nabla e^{\mathrm{int}}_{\bar{x},\bar{p}}|^2$ from (26). Then the triangle $K_{\bar{x},\bar{p}}$ with barycenter $\bar{x}$, the size*

$$|K_{\bar{x},\bar{p}}| := \left(\frac{\bar{\omega}^2 \, (\rho_{\bar{x},\bar{p}})^{1/2}}{c_{\bar{p}} A_{\bar{x},\bar{p}}} \right)^{1/(\bar{p}+1)}, \tag{28}$$

where $c_{\bar{p}} := \pi^{-\bar{p}}/(\bar{p}+1)$, and the anisotropy

$$\sigma_{\bar{x},\bar{p}} = {\rho_{\bar{x},\bar{p}}}^{1/2\bar{p}}, \quad \phi_{\bar{x},\bar{p}} = \varphi_{\bar{x},\bar{p}} - \pi/2 \tag{29}$$

satisfies $\left|e^{\mathrm{int}}_{\bar{x},\bar{p}}\right|_{H^1(K_{\bar{x},\bar{p}})} \le \bar{\omega}$ and has the maximal possible area.

This lemma allows us to define the size and the anisotropy of the optimal triangle for the given polynomial degree and the given tolerance in the H^1-seminorm.

5 Continuous model of an error estimate: h-version

In Section 3.2, we introduced the continuous analogue of the (discrete) hp-mesh from Section 2.1. The next step is to define the continuous analogue to the (discrete) error estimator given in Section 2.2. In this section, we deal with the h-mesh model version only since the theoretical considerations are clear for this case. A partly heuristic extension to the hp-case is given in Section 6. We start with the isotropic case, an extension to the anisotropic case will be given in Section 6.3.

5.1 Continuous analogues of the discrete quantities: h-version

Let $p \in \mathbb{N}$ be the given polynomial approximation degree, i.e., $p_K = p$ for any $K \in \mathscr{T}_h$. Moreover, let us consider an error estimator $\eta^2 = \sum_{K \in \mathscr{T}_h} \eta_K^2$ satisfying (A1)–(A3) from Assumption 2.2. Moreover, we need to consider an additional assumption concerning the relation between the error estimates η_K and the size of the element.

Assumption 5.1. *Let $\mathscr{T}_{hp}$ be a hp-mesh, e_h be the corresponding error of the quantity of interest, $\eta^2 = \sum_{K \in \mathscr{T}_h} \eta_K^2$ be its estimates satisfying (A1)–(A3) such that $\eta = O(h^s) = e_h$, i.e., we have the s^{th}-order method. We assume that the local error estimates η_K, $K \in \mathscr{T}_h$ satisfy the following condition:*

(A4) $\eta_K^2 = C_K |K|^{s+1}$, $K \in \mathscr{T}_h$, where C_K is (almost) independent of the size of the element K but depends on the anisotropy of K and the order of the method s which correlates with the fixed polynomial degree p.

The order of the method s depends on the chosen error measure and the regularity of the exact solution. For example, if we choose $e_h = \|\nabla(u - u_h)\|_{L^2(\Omega)}$ and the exact solution is sufficiently regular then we have $s = p$ where p is polynomial approximation degree. On the other hand, for the choice $e_h = \|u - u_h\|_{L^2(\Omega)}$, we have $s = p + 1$. The case when the exact solution is not regular will be discussed later.

The equality in assumption (A4) is more or less technical, since we can put $C_K := \eta_K^2/|K|^{s+1}$, $K \in \mathscr{T}_h$. Let $n = 2$ be the spatial dimension. The power $s+1$ follows from the expectation that if the given element K is split (isotropically) onto n^2 sub-elements then the square of the error estimator decreases n^{2s}-times (we have s-order method) but the number of elements is n^2 times larger. Hence the square of the error estimator for one refined element has to be $n^{2s+2} = (n^2)^{s+1}$-times smaller. The factor n^2 corresponds to the decrease of the area of the refined elements. The independence of C_K on $|K|$ is more questionable, its experimental justification is given in Section 8.1. See also [25, Section 5.2] where similar property was studied for compressible flow problems.

By assumptions (A1)–(A4), we introduced the discrete quantities $\{|K|\}_{K \in \mathscr{T}_h}$, η, $\{\eta_K\}_{K \in \mathscr{T}_h}$ and $\{C_K\}_{K \in \mathscr{T}_h}$. Now, we describe their continuous analogue. In Section 3.2, we have already acquainted with

- the function $d : \Omega \to \mathbb{R}^+$ called *the element-size distribution function*, which satisfies

$$d(x_K) \approx |K| \qquad \text{for } x_K \in K,\ K \in \mathscr{T}_h. \tag{30}$$

Moreover, we introduce

- the function $B : \Omega \to \mathbb{R}_0^+$, called *density of the error estimate*, which satisfies

$$B(x_K) \approx C_K = \eta_K^2/|K|^{s+1} \qquad \text{for } x_K \in K,\ K \in \mathscr{T}_h, \tag{31}$$

  where C_K, $K \in \mathscr{T}_h$ are the constants from Assumption 5.1.

Furthermore, using functions d and B, we define the *error-estimate distribution function* $E : \Omega \to \mathbb{R}_0^+$ by

$$E(x) := B(x)d(x)^s, \qquad x \in \Omega \tag{32}$$

and the *local* and the *global total error estimates* by

$$\mathscr{E}_K(B,d) := \int_K E(x)\,\mathrm{d}x, \quad K \in \mathscr{T}_h \qquad \text{and} \qquad \mathscr{E}(B,d) := \int_\Omega E(x)\,\mathrm{d}x, \tag{33}$$

respectively.

The "discrete" global and local error estimates η and η_K, $K \in \mathscr{T}_h$ are connected with their continuous analogues $\mathscr{E}$ and $\mathscr{E}_K$, $K \in \mathscr{T}_h$, respectively. Namely, using (30) – (33), we have for the local error estimate η_K the relation

$$\mathscr{E}_K(B,d) = \int_K E(x)\,\mathrm{d}x = \int_K B(x)d(x)^s\,\mathrm{d}x \approx \int_K C_K|K|^s\,\mathrm{d}x = C_K|K|^{s+1} = \eta_K^2. \tag{34}$$

Furthermore, summing (34) over $K \in \mathscr{T}_h$ and using (33), we obtain the relation for the global error estimate η

$$\mathscr{E}(B,d) = \int_\Omega E(x)\,\mathrm{d}x = \sum_{K\in\mathscr{T}_h} \int_K E(x)\,\mathrm{d}x \approx \sum_{K\in\mathscr{T}_h} \eta_K^2 = \eta^2. \tag{35}$$

Finally, let us note that for the h-version of the continuous mesh model, the number of DoF reduces to (cf. (18)–(19))

$$\mathcal{D}(d,p) = \tilde{c}_p \int_\Omega d(x)^{-1}\,\mathrm{d}x = \tilde{c}_p \mathcal{N}(d), \qquad \tilde{c}_p = (p+1)/(p+2)/2. \tag{36}$$

Obviously, the minimization of DoF is equivalent to the minimization of the number of elements of the mesh $\mathscr{T}_h$.

5.2 Mesh optimization problem

Using the relations from Section 5.1, we formulate the continuous analogue of the h-version of Problem 2.3.

Problem 5.2. *Let $\omega > 0$ be the given tolerance and $B(x)$ be the given density of the error estimate. We seek the element-size distribution function $d : \Omega \to \mathbb{R}^+$ such that*

(p1) $\mathscr{E}(B,d) = \int_\Omega B(x)d(x)^s\,\mathrm{d}x \le \omega^2$,

(p2) $\mathcal{N}(d) = \int_\Omega d(x)^{-1}\,\mathrm{d}x$ *is minimal.*

In order to seek the minimum of the functional $\mathcal{N}(d)$ under the constrain $\mathscr{E}(B,d) = \omega^2$, we define the Lagrangian

$$\mathscr{L}(d,\lambda) := \int_\Omega d(x)^{-1}\,\mathrm{d}x + \lambda\left(\int_\Omega B(x)d(x)^s\,\mathrm{d}x - \omega^2\right), \quad \lambda \neq 0. \tag{37}$$

The minimum satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathscr{L}(d + t\tilde{d},\lambda)\Big|_{t=0} = 0 \tag{38}$$

for any perturbation $\tilde{d}$. Performing the derivation in (38), we have

$$0 = -\int_\Omega d(x)^{-2}\,\tilde{d}(x)\,\mathrm{d}x + \lambda\int_\Omega sB(x)d(x)^{s-1}\,\tilde{d}(x)\,\mathrm{d}x. \tag{39}$$

Since (39) should be valid for any $\tilde{d}$, we obtain

$$d(x)^{-2} = \lambda s B(x) d(x)^{s-1} \quad \Rightarrow \quad d(x) = \left(\lambda s B(x)\right)^{-1/(s+1)}. \tag{40}$$

Inserting the previous relation for d into the constrain, we derive

$$\omega^2 = \int_\Omega B(x)\left(\lambda s B(x)\right)^{-s/(s+1)}\,\mathrm{d}x = (\lambda s)^{-s/(s+1)}\int_\Omega B(x)^{1/(s+1)}\,\mathrm{d}x. \tag{41}$$

Eliminating λs from (41), we obtain

$$\lambda s = \left(\frac{\omega^2}{\int_\Omega B(x)^{1/(s+1)} \,\mathrm{d}x} \right)^{-(s+1)/s}. \tag{42}$$

Finally, inserting (42) into (40), we have

$$d(x) = MB(x)^{-1/(s+1)}, \quad \text{where } M := \left(\frac{\omega^2}{\int_\Omega B(x)^{1/(s+1)} \,\mathrm{d}x} \right)^{1/s} = \text{ const.} \tag{43}$$

This relation can be summarized in the following results.

Theorem 5.3. *Let $\omega > 0$ be the prescribed tolerance and $B(x)$ be the given density of the error estimate. Then the element-size distribution function d given by* (43) *is the solution of Problem 5.2.*

An important consequnce of Theorem 5.3 is the following *equi-distribution principle.*

Corollary 5.4. *Let the element-size distribution function d, given by* (43)*, be the solution of Problem 5.2. Then the appropriate local error estimates $\mathscr{E}_K$ are equal for all $K \in \mathscr{T}_h$ and satisfy*

$$\eta_K \approx \frac{\omega}{\sqrt{\#\mathscr{T}_h}}, \quad K \in \mathscr{T}_h. \tag{44}$$

Proof. From (43), we have $B(x) = d(x)^{-(s+1)} M^{s+1}$. Using this relation, (30) and (34), we derive

$$\mathscr{E}_K(B,d) = \int_K B(x) d(x)^s \,\mathrm{d}x = M^{s+1} \int_K d(x)^{-(s+1)} d(x)^s \,\mathrm{d}x = M^{s+1} \int_K d(x)^{-1} \,\mathrm{d}x \approx M^{s+1} \int_K |K|^{-1} \,\mathrm{d}x = M^{s+1}, \ K \in \mathscr{T}_h. \tag{45}$$

Hence, due to (34) and (45), we have $\eta_K^2 \approx M^{s+1}$ for all $K \in \mathscr{T}_h$, which means that η_K have to be the same for all $K \in \mathscr{T}_h$. Moreover, in virtue of (A2) from Assumption 2.2 and condition (p1) in Problem 2.3, we have $\eta^2 = \sum_{K \in \mathscr{T}_h} \eta_K^2 = \#\mathscr{T}_h \eta_K^2 = \omega^2$ which implies (44). □

Remark 5.5. In [17] a continuous-mesh error estimator for the L^q norm of the solution was defined, and the equilibration of the error contributions were experimentally verified (cf. [17], Section III.A).

6 Continuous model of an error estimate: *hp*-version

6.1 Formulation of the mesh optimization problem

Now, our aim is to extend the *h*-version of the continuous error model from Section 6 to the *hp*-version. For simplicity we assume that the order of the method is equal to the polynomial approximation degree, i.e., $s = p$ in Assumption 5.1. It is the case of the used error estimator in Section 7.2.1. The main complication is the fact that the constants $C_K,\ K \in \mathscr{T}_h$ in Assumption 5.1 depends on varying polynomial approximation degree p_K. Then we have to replace Assumption 5.1 by

Assumption 6.1. *Let $\mathscr{T}_{hp}$ be a hp-mesh, e_h the corresponding error of the quantity of interest, $\eta^2 = \sum_{K \in \mathscr{T}_h} \eta_K^2$ be its estimate, where the local error estimates satisfy the following condition:*

(A4′) $\eta_K^2 = C_K |K|^{p_K+1}$, $K \in \mathscr{T}_h$, where C_K is (almost) independent of the size of the element K but it depends on the anisotropy of K and the local polynomial degree p_K.

We can formally introduce the *density of the error estimate* $B : \Omega \to \mathbb{R}_0^+$ (cf. (31)) as a function satisfying

$$B(x_K, p_K) \approx C_K = \eta_K^2 / |K|^{p_K+1} \qquad \text{for } x_K \in K,\ K \in \mathscr{T}_h. \tag{46}$$

Obviously, due to Assumption (A4′), the function B depends on p_K. Then, in contrary to (31), we have $B = B(x, p(x))$, where $p(x)$ is the polynomial degree distribution function from Definition 3.4. Using (46), we can set $B(x, p(x))$ at any $x \in \Omega$, but we do not know the explicit dependence of B on p.

Furthermore, as in Section 5.1, we can formally define the *error-estimate distribution function*

$$E(x) := B(x, p(x))\, d(x)^{p(x)}, \qquad x \in \Omega, \tag{47}$$

and consequently the *local* and the *global total error estimates*

$$\mathscr{E}_K(B,d,p) := \int_K E(x)\,\mathrm{d}x, \quad K \in \mathscr{T}_h \qquad \text{and} \qquad \mathscr{E}(B,d,p) := \int_\Omega E(x)\,\mathrm{d}x, \tag{48}$$

respectively, cf. (32)–(33). Finally, we can formulate the continuous analogue of the *hp*-version of Problem 2.3.

Problem 6.2. *Let $\omega > 0$ be the given tolerance and $B(x,p(x))$ be the density of the error estimate. We seek the element-size distribution function $d : \Omega \to \mathbb{R}^+$ and the polynomial degree distribution function $p : \Omega \to \mathbb{R}_1^+$ such that*

(p1) $\mathscr{E}(B,d,p) = \int_\Omega B(x,p(x))\, d(x)^{p(x)}\,\mathrm{d}x \le \omega^2$,

(p2) $\mathcal{D}(d,p) = \int_\Omega \frac{(p(x)+1)(p(x)+2)}{2d(x)}\,\mathrm{d}x$ *is minimal.*

At this moment, we have no idea how to solve Problem 6.2 due to the unknown dependence of B on p. A more perspective possibility is to consider the interpolation error estimates where the dependence of the estimate of the interpolation error on the polynomial approximation degree is known (cf. [26]).

Nevertheless, we present a heuristic iterative algorithm which solves Problem 6.2 problem partly. Particularly, we obtain the functions $d(x)$ and $p(x)$ which satisfy condition (p1) of Problem 6.2, the corresponding $\mathcal{D}(d,p)$ is small but we can not guarantee that it is the minimal possible.

6.2 Heuristic solution of the mesh optimization problem

In order to develop an algorithm partly solving Problem 6.2, we use two ingredients:

- the *equi-distribution principle* from Corollary 5.4, which allows us to set the size of the elements,
- the *interpolation error estimate* from Lemmas 4.1 or 4.2, which allows us to set a suitable polynomial degree.

Before the presentation of the algorithm, we explain these items.

6.2.1 Use of the equi-distribution principle

We introduce this idea for the discrete setting. Let $K \in \mathscr{T}_h$, p_K be the given polynomial degree and η_K be the computed error estimate. From (A4′) we have

$$\eta_K^2 = C_K |K|^{p_K+1}. \tag{49}$$

Our aim is to set the new area of K denoted as d_K^{new} such that the corresponding error estimate is equal to the right-hand side in (44), i.e.,

$$\bar\omega = C_K (d_K^{\mathrm{new}})^{p_K+1}, \qquad \text{where } \bar\omega := \frac{\omega^2}{\#\mathscr{T}_h}. \tag{50}$$

Let us note that due to assumption (A4′) the constants C_K in (49) and (50) are equal. Eliminating this constant from these relations we obtain

$$d_K^{\mathrm{new}} = |K| \left(\frac{\bar\omega}{\eta_K^2}\right)^{1/(p_K+1)}. \tag{51}$$

6.2.2 Use of interpolation error estimates

We introduce this idea again for the discrete setting and for the case of the L^q norm. The case of the H^1-seminorm can be given analogously using Lemma 4.2 instead of Lemma 4.1. Let $q \in [1,\infty]$ be the Lebesgue index considered. Let $K \in \mathscr{T}_h$, x_K be its barycentre and p_K be the given polynomial degree. Using (21), we define the corresponding interpolation error function denoted by $e^{\mathrm{int}}_{x_K,p_K}$. With the aid of the technique from [13, Section 3], we evaluate the anisotropy of $e^{\mathrm{int}}_{x_K,p_K}$, set the right-hand side of the estimate (22), which is a polynomial function. Integrating the appropriate power of the right-hand side of (22), we evaluate the estimate of $e^{\mathrm{int}}_{x_K,p_K}$ in the $L^q(k)$ norm and we put $\bar\omega_K := \left\|e^{\mathrm{int}}_{x_K,p_K}\right\|_{L^q(K)}$.

Now, we apply Lemma 4.1 for $\bar x := x_K$, $\bar\omega := \bar\omega_K$ and $\bar p \in \{p_K - 1,\ p_K,\ p_K + 1\}$. Then we obtain three anisotropic triangles denoted by K_{-1}, K_0 and K_1 satisfying $\left\|e^{\mathrm{int}}_{x_K,p_K}\right\|_{L^q(K_j)} \le \bar\omega_K$, $j = -1, 0, 1$. From these three candidates, we choose this one, which has the smallest density of the number of degrees of freedom (cf. (19)). Hence, we define a new (more suitable) polynomial approximation degree

$$p_K^{\mathrm{new}} := p_K + \bar\jmath_K, \qquad \text{where} \quad \bar\jmath_K = \arg\min_{j\in\{-1,0,1\}} \frac{(p_K+j+1)(p_K+j+2)}{2|K_j|}. \tag{52}$$

6.2.3 Mesh optimization algorithm

Now, we are ready to describe the algorithm which seeks an approximate solution of Problem 6.2.

Algorithm 6.3. Let $\tilde{d}(x)$ and $\tilde{p}(x)$ be an initial guess of element-size distribution function and the polynomial degree distribution function, respectively. For simplicity, we can start with $\tilde{d}(x) = \text{const}$ as well as $\tilde{p}(x) = \text{const}$.

(S1) let $\eta_K,\ K \in \mathscr{T}_h$ be the error estimate of the approximate solution u_h computed on the *hp*-mesh corresponding to $\tilde{d}(x)$ and $\tilde{p}(x)$,

(S2) we define the function

$$\tilde{\eta}(x) :\approx \eta_K, \quad x \in K,\ K \in \mathscr{T}_h, \tag{53}$$

(S3) in virtue of (18) and the second relation in (50), we set $\bar{\omega} := \omega^2 \left(\int_\Omega \left(\tilde{d}(x) \right)^{-1} \mathrm{d}x \right)^{-1}$,

(S4) using (51), we put

$$d(x) = \tilde{d}(x) \left(\frac{\bar{\omega}}{\tilde{\eta}^2(x)} \right)^{1/(\tilde{p}(x)+1)}, \quad x \in \Omega, \tag{54}$$

(S5) for $x \in \Omega$, we apply the procedure from Section 6.2.2 with $\bar{p} := \mathrm{int}[\tilde{p}(x)]$, where $\mathrm{int}[\cdot]$ denotes the rounding to the nearest integer, cf. (16). Then we put

$$p(x) = \tilde{p}(x) + \bar{j}(x), \tag{55}$$

where $\bar{j}(x)$ is the argument of minimum from (52) applied for $\bar{x} = x$. The range of the function $\bar{j}(x)$ is the discrete set $\{-1, 0, 1\}$. The resulting function p is not continuous in general, hence some smoothing has to be applied if necessary.

(S6) If $\bar{j}(x) = 0$ (i.e., $\tilde{p}(x)$ is already optimal degree) then we stop the algorithm, otherwise we put $\tilde{d}(x) := d(x)$, $\tilde{p}(x) := p(x)$ and go to step (S1).

Remark 6.4. We discuss the particular steps in Algorithm (6.3).

- This algorithm defines the element-size distribution function $d(x)$ and the polynomial degree distribution function $p(x)$ but not the anisotropy characterized by the aspect ratio and orientation distribution functions $\sigma(x)$ and $\phi(x)$, respectively, cf. Definition 3.4. Putting $\sigma = 1$ and $\phi = 0$, the algorithm generates the isotropic *hp*-meshes. The anisotropic case is treated in Section 6.3.

- The steps (S1)–(S4) set the element size distribution $d(x)$ for the given polynomial degree distribution function $\tilde{p}(x)$. If $\tilde{p}(x) = \text{const}$ then Theorem 5.3 implies that this element size distribution is optimal. For non-constant $\tilde{p}$, we believe, that due to the *equi-distribution principle* given in Corollary 5.4 (which is p-independent), the resulting $d(x)$ is close to the (hypothetical) optimal one.

- The step (S5) allows us to change the polynomial degree distribution locally in Ω with possible decreasing the number of DoF while the interpolation error estimate is kept. Of course, we have no guarantee that the error estimates η_K, $K \in \mathscr{T}_h$ are also kept. However, we have no idea how to replace the estimate of the interpolation error functions by the estimates η_K, $K \in \mathscr{T}_h$ in step (S5).

- The meaning of the step (S6) is to stop the algorithm when the suitable polynomial degree distribution was found. In practice, we stop the computation when the prescribed error tolerance was achieved.

- In practice, we cannot perform Algorithm 6.3 for all $x \in \Omega$. Usually, we apply it only for the barycenters x_K of all $K \in \mathscr{T}_h$ and then the continuous functions are formally reconstructed.

- The evaluation of the high-order derivatives required in step (S5) is described in Section 7.1.

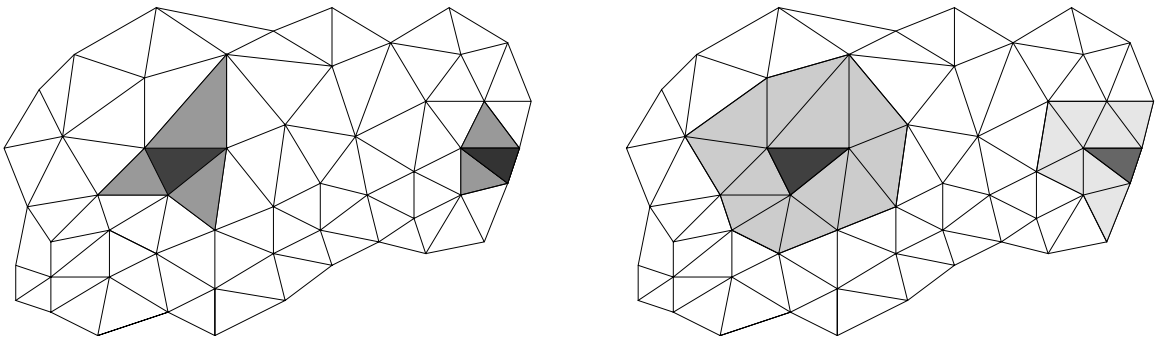

Fig. 3: Examples of patches $\mathscr{D}_K$ corresponding to interior and boundary elements, the small (left) and large (right) patches.

6.3 Anisotropic meshes

Algorithm 6.3 presented above gives isotropic grids only. In order to produce anisotropic grids we have to define the aspect-ratio distribution function $\sigma(x)$ and the orientation distribution function $\phi(x)$. Similarly as in step (S5) of Algorithm 6.3, we set the anisotropy of elements with respect to the interpolation error using Lemmas 4.1 or 4.2. Again, we have no idea how to set the optimal anisotropy with respect to the error estimate η.

In order to define the anisotropic grids, step (S5) has to be accompanied by the additional step

(S5′) for $x \in \Omega$ and the chosen $p(x)$ from (S5), we apply Lemmas 4.1 or 4.2 with $\bar{x} := x$, $\bar{p} := p(x)$ and $\bar{\omega}$ from step (S3). Then we set $\sigma(x) := \sigma_{\bar{x},\bar{p}}$ and $\phi(x) := \phi_{\bar{x},\bar{p}}$.

7 Technical aspects of practical computations

In this section we briefly mention some technical aspects of the practical computations, which are used in Section 8. First, we describe the evaluation of the high-order derivatives required by Algorithm 6.3. Further, we describe the used error estimators η, η_K, $K \in \mathscr{T}_h$. Let us note that we believe that the presented algorithm works with any reasonable estimator. Finally, we present a note about problems with a non-regular exact solution.

7.1 Evaluation of the high-order derivatives

In order to choose the optimal polynomial approximation degrees from the candidates $p_K - 1$, p_K and $p_K + 1$, $K \in \mathscr{T}_h$, in steps (S5) or (S5′) of Algorithm 6.3, we use the upper bounds in (22) or (26). Hence, we have to approximate the partial derivatives of degree p_K, $p_K + 1$ and $p_K + 2$ on each $K \in \mathscr{T}_h$. We employ the technique from [13] and [14], where, for each $K \in \mathscr{T}_h$, we define a patch $\mathscr{D}_K$ around K and the polynomial function $\tilde{u}_{K,r_K} \in P^{r_K}(\mathscr{D}_K)$ by

$$(\tilde{u}_{K,r_K}, \phi)_{1,\mathscr{D}_K} = (u_{hp}, \phi)_{1,\mathscr{D}_K} \quad \forall \phi \in P^{r_K}(\mathscr{D}_K), \quad r_K \in \{p_K - 1,\ p_K,\ p_k + 1\} \tag{56}$$

where $P^{r_K}(\mathscr{D}_K)$ is the space of polynomial functions of degree r_K on the patch $\mathscr{D}_K$ and $(\cdot,\cdot)_{1,\mathscr{D}_K}$ is the H^1-scalar product on $\mathscr{D}_K$. We consider here two possibilities: a small and a large patch. In the former case, the patch $\mathscr{D}_K$ consists of K and its neighbours and in the latter case, the patch $\mathscr{D}_K$ consists of all elements from $\mathscr{T}_h$ sharing at least a vertex with K, see Figure 3.

7.2 Error estimators

The majority of the numerical experiments presented in Section 8 is obtained by the *residual error estimator* (RES) developed in [27]. Although this estimator is not sufficiently theoretically supported, it works very well. Moreover, in Section 8.4, we compare it with the *dual weighted residual* (DWR) error estimators (see [21]), which is in a great interest of many authors. For the completeness, we briefly describe these error estimates.

7.2.1 Residual error estimator – RES

We introduce the main idea of this approach. Let V and V_h be the functional spaces, $\dim(V_h) < \infty$ and $V_h \in V$. Moreover, let

$$u \in V: \quad a_h(u,v) = \ell_h(v) \quad \forall v \in V, \qquad \text{and} \qquad u_h \in V_h: \quad a_h(u_h, v_h) = \ell_h(v_h) \quad \forall v_h \in V_h, \tag{57}$$

be the consistency property of the exact solution $u \in V$ and the definition of the approximate solution $u_h \in V_h$, respectively. The form a_h is linear with respect to its second argument and ℓ_h is a linear form. Then we define the error measure by

$$\sup_{v\in V} \frac{a_h(u_h,v)-a_h(u,v)}{\|v\|_X} = \sup_{v\in V} \frac{a_h(u_h,v)-\ell_h(v)}{\|v\|_X}, \tag{58}$$

where the equality follows from the consistency relation in (57). the norm $\|\cdot\|_X$ will be specified later. However, it is impossible to evaluate the supreme in (58) over the infinite dimensional space V. Hence, we introduce a finite dimensional space V_h^+ such that $V_h \subsetneq V_h^+$ and define

$$\eta := \sup_{v\in V_h^+} \frac{a_h(u_h,v)-\ell_h(v)}{\|v\|_X}, \qquad \eta_K := \sup_{v\in V_h^+,\mathrm{supp}(v)\subset K} \frac{a_h(u_h,v)-\ell_h(v)}{\|v\|_X}, \quad K\in\mathscr{T}_h, \tag{59}$$

where we put $\|\cdot\|_X := |\cdot|_{H^1(\Omega)}$ for the purely diffusive problems and $\|\cdot\|_X := \left(\|\cdot\|^2_{L^2(\Omega)} + \varepsilon|\cdot|^2_{H^1(\Omega)}\right)^{1/2}$ for the convection-diffusion problems with ε equal to the ratio between the diffusion and the convection. Such error estimator satisfies assumption (A1)–(A3) from Assumption 2.2 with $e_h := \|u-u_h\|_X$, cf. [27].

7.2.2 Dual weighted residual error estimator – DWR

The aim of the this approach is to estimate the error of the *quantity of interest*, see, e.g., [21]. For simplicity, we restrict to the linear form a_h in (57). Let $J : V \cup S_{hp} \to \mathbb{R}$ be a linear functional defining the quantity of interest. We consider the *dual problem*

$$z\in V: \quad a_h(y,z) = J(y) \quad \forall y\in V. \tag{60}$$

Using the Galerkin orthogonality, we derive

$$J(u_h)-J(u) = J(u_h-u) = a_h(u_h-u,z) = a_h(u_h,z)-\ell_h(z). \tag{61}$$

Using a numerical approximation of z we can approximate the right-hand side of (61) and thus estimate the error $J(u_h)-J(u)$. Let us note the link between (59) and (61). In the former case, we approximate the supremum of the residual form $a_h(u_h,\cdot)-\ell_h(\cdot)$ whereas in the latter case we insert there the dual solution.

7.3 Non-regular solutions

Assumptions (A4) and (A4′) as well as results from Lemmas 4.1 and 4.2 require a sufficiently regular exact solution. Here we give an argumentation that Algorithm 6.3 successfully works also for non-regular solution. If the exact solution is locally singular (e.g., due to the presence of interior corners or singular source terms) then the relation assumption (A4′) changes to $\eta_K^2 = C_K|K|^{z_K}$, where $z_K \le p_K+1$ is the local regularity index on $K\in\mathscr{T}_h$, i.e., $u|_K \in H^{z_K}(K)$, compare with (66). Then step (S4) in Algorithm 6.3 underestimates the optimal area. However, the next loop of the algorithm will improve this value towards the optimal one. Therefore, we can expect a slowdown of the convergence of the algorithm but there is no principal obstacle to achieve a reasonable final *hp*-mesh.

Moreover, concerning the interpolation error estimates, when the solution is not sufficiently regular, the higher order derivatives do not exist, they diverge. Nevertheless, their approximation (cf. Section 7.1) leads to finite but very large values of $A_{\bar{x},\bar{p}}$ at the right-hand sides of (22) or (26) and then the technique described in Section 6.2.2 avoids larger polynomial degrees since the fraction in (52) is too high. On the other hand, we can argue that the higher-order reconstruction smooths the solution and hence the algorithm "overestimates" the regularity of the solution and the resulting *hp*-grid will be far form the optimal one. However, numerical experiments in Section 8.3 demonstrate the efficiency of the algorithm in comparison with a technique which takes into account a prior knowledge of the regularity solution.

8 Numerical experiments

In this section, we present the computational performance of Algorithm 6.3. In Section 8.1, we give an example with sufficiently regular solution which supports the assumption (A4) (cf. Assumption 5.1). In Section 8.2, we consider an example, where the exact solution has strong anisotropic features - two thin boundary layers. Here, we compare several possible settings in Algorithm 6.3 and also demonstrate the efficiency of the *hp*-anisotropic adaptation in comparison to the *h*- and isotropic versions. In Section 8.3, we deal with an example having the exact solution with a mild singularity. We show that the presented algorithm works also for non-sufficiently regular exact solution and moreover it is superior to techniques

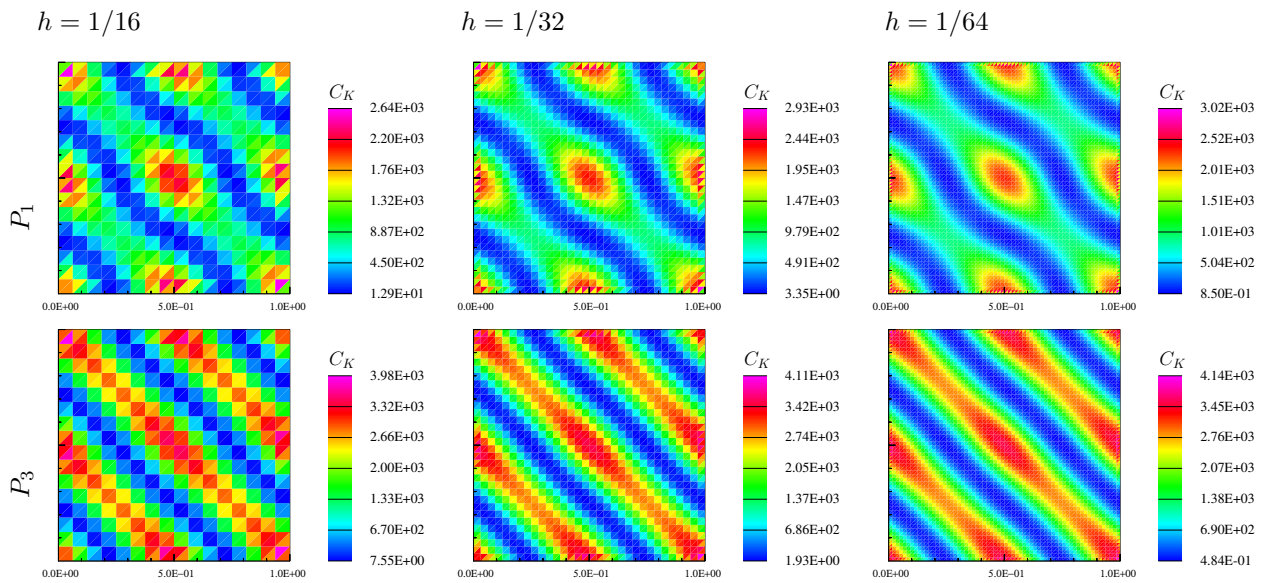

Fig. 4: Numerical justification of (A4) from Assumption 5.1.

taking into account a prior knowledge of the regularity. Finally, in Section 8.4, we demonstrate that Algorithm 6.3 can be used for different error estimators. Namely, we compare the performance of the residual (RES) and the dual weighted residual (DWR) error estimators. All numerical experiments in Sections 8.1 – 8.3 were performed with the residual error estimator from Section 7.2.1.

8.1 Experimental justification of (A4) from Assumption 5.1

In order to verify Assumption 5.1, we consider the Laplace problem $-\Delta u = f$ in $\Omega := (0,1) \times (0,1)$, $u = u_D$ on $\partial\Omega$, where f and u_D are chosen such that the exact solution has the form $u = \sin(2\pi x_1)\sin(2\pi x_2)$. We carried the computations using P_1 and P_3 polynomial approximation degrees on the sequence of three uniform triangular grids with spacing $h = 1/16$, $h = 1/32$ and $h = 1/64$. Figure 4 shows the distribution of the quantity $C_K := \eta_K^2/|K|^{p+1}$. The scale for each separate plot in Figure 4 is from the minimal value to the maximal one. We simply observe a very small dependence of the maximal value of C_K, $K \in \mathscr{T}_h$ on the mesh size. On the other hand, the dependence of the minimal value is not negligible. However, the elements with the smallest error estimator usually do not contribute too much to the total error.

8.2 Linear convection-diffusion equation with boundary layers

We consider the scalar linear convection-diffusion equation (similarly as in [28])

$$-\varepsilon\triangle u - \frac{\partial u}{\partial x_1} - \frac{\partial u}{\partial x_2} = g \quad \text{in } \Omega := (0,1)^2, \tag{62}$$

where $\varepsilon = 10^{-3}$ is a constant diffusion coefficient. We prescribe the Dirichlet boundary condition on $\partial\Omega$ and the source term g such that the exact solution has the form $u(x_1,x_2) = (c_1 + c_2(1-x_1) + e^{-x_1/\varepsilon})(c_1 + c_2(1-x_2) + e^{-x_2/\varepsilon})$ with $c_1 = -e^{-1/\varepsilon}$, $c_2 = -1 - c_1$. The solution contains two boundary layers along $x_1 = 0$ and $x_2 = 0$, whose width is proportional to ε. For the problem (62), we carried out two sets of experiments.

8.2.1 Comparison of technical aspects

We compare two aspects of Algorithm (6.3), namely, the type of the higher-order reconstruction (small and large patches, Figure 3) and the shape optimality with respect to the L^2-norm and the H^1-seminorm (Lemmas 4.1 and 4.2). These results are shown in Figure 5, left, which shows the dependence of the error estimator on the third root of the number of degrees of freedom in log-linear scale since the theoretical results for *hp*-methods give an exponential rate of the convergence in the form $e_h \approx C\exp(-b\,DoF^{1/3})$ where C and b are positive constants. We observe that the H^1-seminorm optimality is slightly superior to the L^2-norm one and that the large patches are a little more efficient than the small ones.

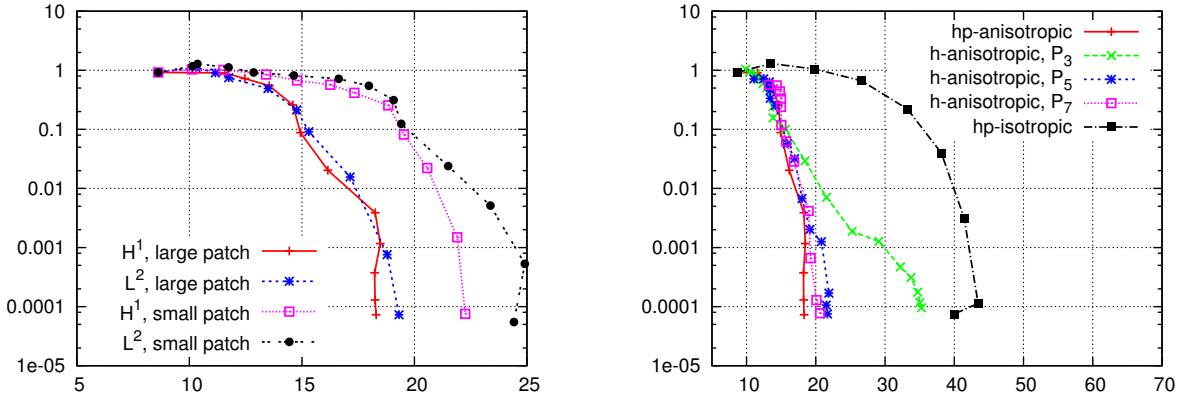

Fig. 5: Problem (62), convergence of the error estimator w. r. t. $\mathrm{DoF}^{1/3}$: the small/large patches in the higher-order reconstruction and the L^2 / H^1 optimal element shapes (left); the *hp*-anisotropic, *h*-anisotropic (P_3, P_5 and P_7 discretizations) and *hp*-isotropic adaptations (right).

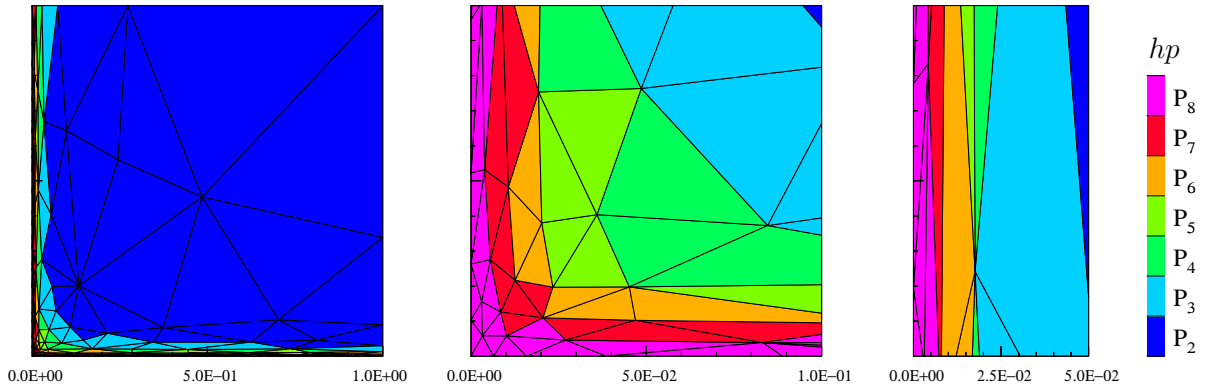

Fig. 6: Problem (62), final *hp*-mesh obtained by the anisotropic *hp*-mesh adaptation: the total view (left), the zoom of the left-bottom corner (center) and the zoom of the left boundary layer (right).

8.2.2 Comparison of the *hp*-anisotropic method with the *h*-anisotropic and the *hp*-isotropic ones

Further, we compare the performance of the *hp*-anisotropic adaptation with the *h*-anisotropic adaptation using P_3, P_5 and P_7 approximations and with the *hp*-isotropic adaptation. These versions can be easily obtained by the modification of Algorithm 6.3, where we either fix polynomial approximation degree or the anisotropy of elements, see the end of Section 3.2. We use the higher-order reconstruction on the large patches and the shape-optimality with respect to the H^1-seminorm. Figure 5, right, shows the dependence of the error estimator on the third root of the number of degrees of freedom. Although the *hp*-isotropic adaptation gives exponential order of convergence, it is incomparable with the anisotropic techniques due to the anisotropic features of the problem. Finally, Figure 6 shows the final *hp*-grid obtained by the the *hp*-anisotropic adaptation.

8.3 Interior line singularity

We consider the scalar nonlinear diffusion equation

$$-\nabla\cdot(\boldsymbol{K}(u)\nabla u) = g \qquad \text{in } \Omega=(-1,1)^2, \qquad u = u_D \quad \text{on } \partial\Omega, \tag{63}$$

where $\boldsymbol{K}(u)$ is given by

$$\boldsymbol{K}(u) = \begin{pmatrix} 2+\arctan(u) & (2-\arctan(u))/4 \\ 0 & (4+\arctan(u))/2 \end{pmatrix}. \tag{64}$$

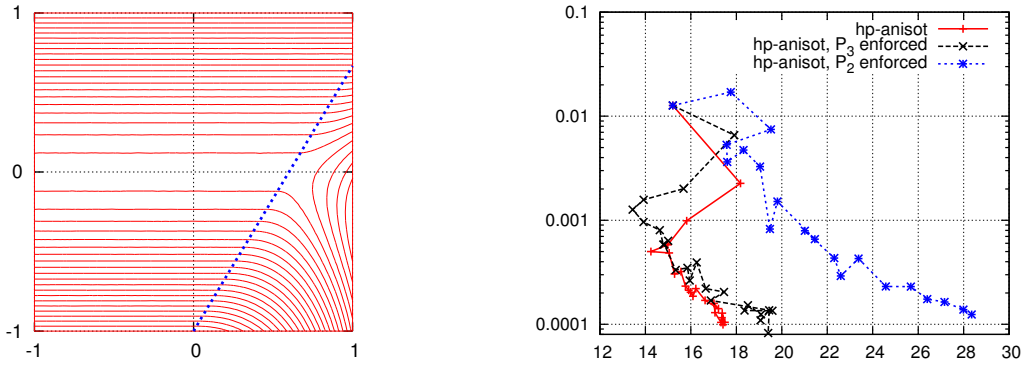

Fig. 7: Isolines of the exact solution (65), the singularity lines is dashed (left). The convergence of the error estimator w. r. t. $\mathrm{DoF}^{1/3}$ (right)

We prescribe a Dirichlet boundary condition u_D on $\partial\Omega$ and set the source term g such that the exact solution is

$$u(x_1,x_2)=\begin{cases}\cos(\pi x_2/2) & \text{for } x_1\le\beta(x_2-1),\\ \cos(\pi x_2/2)+(x_1-\beta(x_2-1))^\alpha & \text{for } x_1>\beta(x_2-1),\end{cases} \tag{65}$$

where we put $\alpha=2$ and $\beta=0.6$. The solution satisfies $u\in H^{\alpha+1/2-\epsilon}(\Omega)$ and posses a mild singularity along the line $x_1-\beta(x_2-1)=0$. This line singularity is difficult to capture since it is very weak, see Figure 7, left, showing the isolines of the exact solution. However, without a sufficient h-refinement along this line, it is not possible to decrease the computational error under the given tolerance.

Some of the hp-adaptive grids from [29] are based on an estimation of the local regularity of the exact solution and then they set the appropriate polynomial approximation degree $p_K:=z_K-1$, where z_K is local regularity index on $K\in\mathscr{T}_h$, i.e., $u|_K\in H^{z_K}(K)$. This relation follows from the approximation properties of the L^2-projection $\Pi^{L^2}:H^{z_K}(K)\to P^{p_K}(K)$ written as

$$|\Pi^{L^2}v-v|_{H^1(K)}\le C_A\frac{h_K^{\mu_K-1}}{p^{z_K-1}}\|v\|_{H^{z_K}(K)},\quad v\in H^{z_K}(K),\ K\in\mathscr{T}_h, \tag{66}$$

where $C_A>0$, $\mu_K=\min(p_K+1,z_K)$ (cf. [30, Lemma 4.5]), and assuming the optimal setting is $\mu_K=p_K+1=z_K$. In our case, $u\in H^{2.5-\epsilon}(\Omega)$, $\epsilon>0$ and then the optimal polynomial degree should be 1 or 2.

We compare our anisotropic hp-adaptation algorithm with its modification, where for the elements, which are intersected by the interior line of the singularity, we enforce the P_2 and P_3 polynomial approximation. For all three computations, we use higher-order reconstruction on the large patches and the shape-optimality with respect to the H^1-seminorm and we prescribed the tolerance $\omega=10^{-4}$. Figure 7, right, shows dependence of the error estimator on the third root of the number of degrees of freedom. We observe that the convergence is not monotone as it is usual in standard technique since Algorithm 6.3 successively improves the quality of the mesh. The final hp-grids with the zoom is viewed in Figure 8. The triangles in the vicinity of the interior line singularity use mostly polynomial approximation degree 4 for the original Algorithm and degrees 2 and 3 for its modifications, where these degrees are enforced. We simply observe that the algorithm without this modification is more efficient, it achieves the same error tolerance using lower number of degrees of freedom. The modification of the algorithm with enforcing polynomial degrees in triangles intersecting the line singularity requires stronger h-refinement. From this experiments we deduce that the setting of the polynomial approximation degree from (66) is not optimal. Similar experiences were observed, e.g., in [31, Example 5] and [27, Section 6.1].

8.4 Double interior layers

We consider a linear convection-dominated problem [32, Example 6.2]

$$-\varepsilon\triangle u+b_1\frac{\partial u}{\partial x_1}+b_2\frac{\partial u}{\partial x_2}=0\quad\text{in }\Omega:=(0,1)^2, \tag{67}$$

where $\varepsilon=10^{-6}$ and $(b_1,b_2)=(-x_2,x_1)$, is the velocity field with curved characteristics. We prescribe the homogeneous Neumann data at the outflow part $\partial\Omega_N=\{0\}\times(0,1)$ and the discontinuous Dirichlet data $u=1$ at $(x_1,x_2)\in(\frac{1}{3},\frac{2}{3})\times\{0\}$

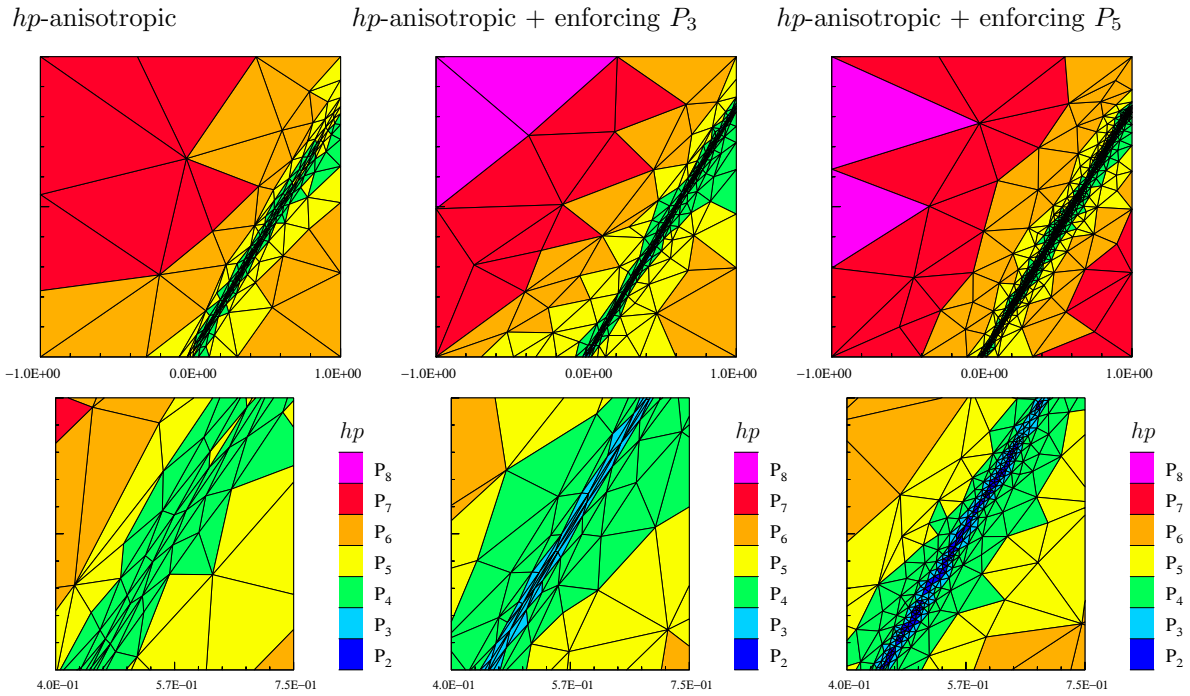

Fig. 8: Case (63): final *hp*-mesh obtained by the anisotropic *hp*-mesh adaptation, the total view (top) and the zoom of middle of the interior line singularity (right).

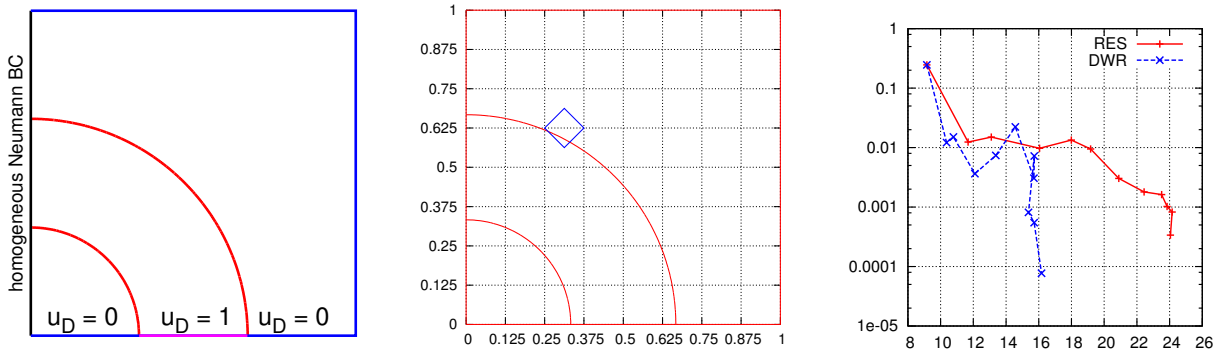

Fig. 9: Example (67), double interior layer: the setting of the boundary condition with interior layers (left), the sub-domain of the quantity of interest Ω_A (center) and the convergence of $|J(u) - J(u_h)|$ w.r.t. $\mathrm{DoF}^{1/3}$ for the RES and DWR error estimates (right).

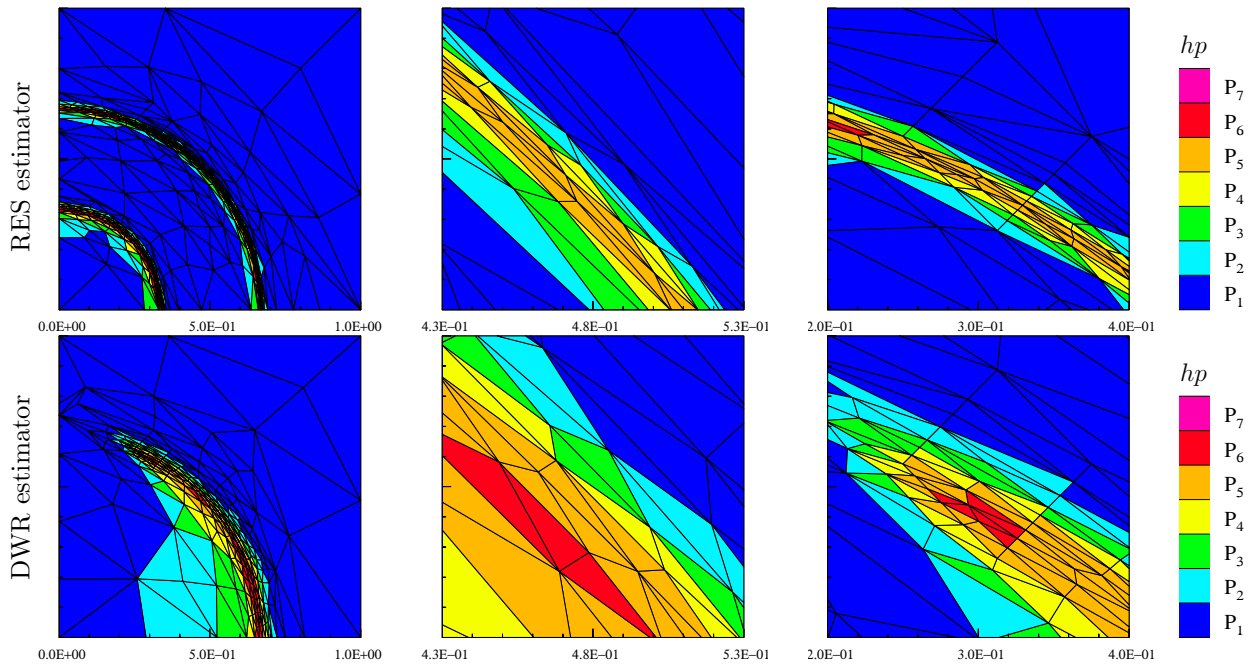

Fig. 10: Case (67): final *hp*-mesh obtained by the anisotropic *hp*-mesh adaptation obtained by the RES (top) and (DWR) bottom error estimators: the total view (left), the zoom of the middle of the outer arc (center) and the neighbourhood of Ω_A (right).

and $u = 0$ elsewhere on $\partial\Omega_D := \partial\Omega \setminus \partial\Omega_N$, see Figure 9, left. Then this discontinuous profile is basically transported along the characteristic curves leading to sharp characteristic interior layers.

We applied the *hp*-anisotropic mesh adaptation for the residual error estimator (RES) and the dual weighted residual error estimator (DWR). For the latter case, we defined the function $J(u)$ as the mean value over the small square sub-domain Ω_A defined by points $[\frac{1}{4}, \frac{5}{8}]$, $[\frac{5}{16}, \frac{9}{16}]$, $[\frac{3}{8}, \frac{5}{8}]$ and $[\frac{5}{16}, \frac{11}{16}]$, see Figure 9, center.

Figure 9, right, shows the convergence of the error $|J(u) - J(u_h)|$ with respect to the third root of DoF for both error estimators. Obviously, using DWR we obtain the same error level using smaller number of DoF. It easily follows from Figure 10, which shows the final *hp*-grids obtained by the algorithm for both error estimators. As we can expect, RES leads to a strong refinement along both interior layers whereas DWR leads to the refinement only in the vicinity of the outer arc and only in front of the domain if interest.

Conclusion

We presented an algorithm which generates anisotropic *hp*-grids such that an arbitrary error estimate of the corresponding approximate solution is under the given tolerance and the number of DoF is small. We developed and employed the *hp*-version of the continuous mesh model and the error model. The anisotropy of elements and the resulting polynomial approximation degrees are set using interpolation error estimates. Further goal is the development of an approach allowing to set the anisotropy and the polynomial approximation degrees using the error estimate itself.

Acknowledgements The first author is thankful to prof. Endre Suli (Oxford University) for a fruitful discussion during his stay at the Charles University in Prague.

References

[1] C. Schwab, *p*- and *hp*-Finite Element Methods, Clarendon Press, Oxford, 1998.

[2] J. M. Melenk, *hp*-finite element methods for singular perturbations, Vol. 1796 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, 2002.

[3] I. Babuška, M. Suri, The *p*- and *hp*-FEM a survey, SIAM Review 36 (1994) 578–632.

[4] L. Demkowicz, Computing with *hp*-adaptive finite elements. Vol. 1, Chapman & Hall/CRC Applied Mathematics and Nonlinear Science Series, Chapman & Hall/CRC, Boca Raton, FL, 2007, one and two dimensional elliptic and Maxwell problems, With 1 CD-ROM (UNIX).

[5] W. G. Habashi, J. Dompierre, Y. Bourgault, D. Ait-Ali-Yahia, M. Fortin, M.-G. Vallet, Anisotropic mesh adaptation: towards user-independent, mesh-independent and solver-independent CFD. Part I: general principles, Int. J. Numer. Methods Fluids 32 (6) (2000) 725–744.

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

[7] J. Dompierre, M.-G. Vallet, Y. Bourgault, M. Fortin, W. G. Habashi, Anisotropic mesh adaptation: towards user-independent, mesh-independent and solver-independent CFD. Part III. unstructured meshes, Int. J. Numer. Methods Fluids 39 (8) (2002) 675–702.

[8] P. J. Frey, F. Alauzet, Anisotropic mesh adaptation for CFD computations, Comput. Methods Appl. Mech. Engrg. 194 (2005) 5068–5082.

[9] W. Cao, Anisotropic measures of third order derivatives and the quadratic interpolation error on triangular elements, SIAM J. Sci. Comput. 29 (2) (2007) 756–781.

[10] W. Cao, An interpolation error estimate in R^2 based on the anisotropic measures of higher order derivatives, Math. Comp. 77 (261) (2008) 265–286.

[11] J.-M. Mirebeau, Optimal meshes for finite elements of arbitrary order, Constr. Approx. 32 (2) (2010) 339–383.

[12] J.-M. Mirebeau, Optimally adapted meshes for finite elements of arbitrary order and $W^{1,p}$ norms, Numer. Math. 120 (2) (2012) 271–305.

[13] V. Dolejší, Anisotropic *hp*-adaptive method based on interpolation error estimates in the L^q-norm, Appl. Numer. Math. 82 (2014) 80–114.

[14] A. Balan, M. Woopen, G. May, Adjoint-based *hp*-adaptivity on anisotropic meshes for high-order compressible flow simulations, Comput. Fluids (accepted). doi:http://dx.doi.org/10.1016/j.compfluid.2016.03.029.

[15] A. Loseille, F. Alauzet, Continuous mesh framework part I: well-posed continuous interpolation error, SIAM J. Numer. Anal. 49 (1) (2011) 38–60.

[16] A. Loseille, F. Alauzet, Continuous mesh framework part II: validations and applications, SIAM J. Numer. Anal. 49 (1) (2011) 61–86.

[17] A. Rangarajan, A. Balan, G. May, Mesh adaptation and optimization for discontinuous Galerkin methods using a continuous mesh model, 2016.

[18] M. Ceze, K. Fidkowski, Anisotropic *hp*-adaptation framework for functional prediction, AIAA Journal 51 (2) (2013) 492–509.

[19] V. Dolejší, Anisotropic *hp*-adaptive method based on interpolation error estimates in the H^1-seminorm, Appl. Math. 60 (6) (2015) 597–616.

[20] P. G. Ciarlet, The Finite Elements Method for Elliptic Problems, North-Holland, Amsterdam, New York, Oxford, 1979.

[21] R. Becker, R. Rannacher, An optimal control approach to a-posteriori error estimation in finite element methods, Acta Numerica 10 (2001) 1–102.

[22] M. Jourlin, I. Fillere, J.-M. Becker, M.-J. Labourne, Shape and metrices, in: D. Chetverikov, W. Kropatsch (Eds.), Computer Analysis of Images and Patterns: 5th International Conference, CAIP '93 Budapest, Hungary, September 13-15, 1993, Lecture Notes in Computer Science, Springer-Verlag, 1993, pp. 254–258.

[23] P. Laug, H. Borouchaki, BL2D-V2: isotropic or anisotropic 2D mesher, INRIA, https://www.rocq.inria.fr/gamma/Patrick.Laug/logiciels/ bl2d-v2/INDEX.html (2002).

[24] V. Dolejší, ANGENER – software package, Charles University Prague, Faculty of Mathematics and Physics, www.karlin.mff.cuni.cz/ dolejsi/angen.html (2000).

[25] D. Venditti, D. Darmofal, Grid adaptation for functional outputs: Application to two-dimensional inviscid flows, J. Computat. Phys. 176 (1) (2002) 40–69.

[26] V. Dolejší, G. May, A. Rangarajan, Continuous *hp*-mesh model, (in preparation).

[27] V. Dolejší, *hp*-DGFEM for nonlinear convection-diffusion problems, Math. Comput. Simul. 87 (2013) 87–118.

[28] C. Clavero, J. L. Gracia, J. C. Jorge, A uniformly convergent alternating direction (HODIE) finite difference scheme for 2D time-dependent convection-diffusion problems, IMA J. Numer. Anal. 26 (2006) 155–172.

[29] W. F. Mitchell, M. A. McClain, A comparison of *hp*-adaptive strategies for elliptic partial differential equations, ACM Trans. Math. Software 41 (1) (2014) Art. 2, 39.

[30] I. Babuška, M. Suri, The *hp*-version of the finite element method with quasiuniform meshes, M^2AN Math. Model. Numer. Anal. 21 (1987) 199–238.

[31] L. Demkowicz, W. Rachowicz, P. Devloo, A fully automatic *hp*-adaptivity, J. Sci. Comput. 17 (1-4) (2002) 117–142.

[32] T. Knopp, G. Lube, G. Rapin, Stabilized finite element methods with shock capturing for advection–diffusion problems, Comput. Methods Appl. Mech. Engrg. 191 (2002) 2997–3013.