

An a-posteriori error estimate for hp -adaptive DG methods for convection–diffusion problems on anisotropically refined meshes

Stefano Giani^a, Dominik Schötzau^{b,*}, Liang Zhu^b

^a School of Mathematical Sciences, University of Nottingham, University Park, Nottingham, NG7 2RD, UK

^b Mathematics Department, University of British Columbia, 1984 Mathematics Road, Vancouver, BC, V6T 1Z2, Canada

ARTICLE INFO

Keywords:

Discontinuous Galerkin methods
Error estimation
 hp -adaptivity
Convection–diffusion problems

ABSTRACT

We prove an a-posteriori error estimate for hp -adaptive discontinuous Galerkin methods for the numerical solution of convection–diffusion equations on anisotropically refined rectangular elements. The estimate yields global upper and lower bounds of the errors measured in terms of a natural norm associated with diffusion and a semi-norm associated with convection. The anisotropy of the underlying meshes is incorporated in the upper bound through an alignment measure. We present a series of numerical experiments to test the feasibility of this approach within a fully automated hp -adaptive refinement algorithm.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

We derive and numerically test a residual-based a-posteriori error estimate for hp -version discontinuous Galerkin (DG) methods for the convection–diffusion model problem:

$$\begin{aligned} -\varepsilon \Delta u + \underline{a}(x) \cdot \nabla u &= f(x) \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \Gamma. \end{aligned} \quad (1)$$

Here, Ω is a bounded Lipschitz polygonal domain in \mathbb{R}^2 with boundary $\Gamma = \partial\Omega$. The parameter $\varepsilon > 0$ is the (constant) diffusion coefficient, the function $\underline{a}(x) \in W^{1,\infty}(\Omega)^2$ a given flow field, and $f(x)$ a source term in $L^2(\Omega)$. We assume that

$$\nabla \cdot \underline{a} = 0 \quad \text{in } \Omega. \quad (2)$$

For simplicity, we shall also assume that $\|\underline{a}\|_{L^\infty(\Omega)}$ and the length scale of Ω are of order one so that ε^{-1} is the Péclet number of the problem. The standard weak form of the convection–diffusion equation (1) is to find $u \in H_0^1(\Omega)$ such that

$$A(u, v) = \int_{\Omega} (\varepsilon \nabla u \cdot \nabla v + \underline{a} \cdot \nabla uv) dx = \int_{\Omega} f v dx \quad \forall v \in H_0^1(\Omega). \quad (3)$$

Under assumption (2), the variational problem (3) is uniquely solvable.

This paper is a continuation of our work on hp -adaptive DG methods for diffusion and convection–diffusion problems. This work was initiated in [1], where an energy norm a-posteriori error estimate was derived for hp -version DG methods for diffusion problems in two dimensions. The key technical tool was the introduction of an hp -version averaging operator, inspired by that of [2] for h -version DG methods. In [3], related averaging techniques were used in the numerical analysis of continuous interior penalty hp -elements. Extensions to linear elasticity in mixed form, quasi-linear elliptic problems and three-dimensional diffusion equations were presented in [4–6], respectively. In [7], the same averaging approach was

* Corresponding author.

E-mail addresses: stefano.giani@nottingham.ac.uk (S. Giani), schoetzau@math.ubc.ca (D. Schötzau), zhuliang@math.ubc.ca (L. Zhu).

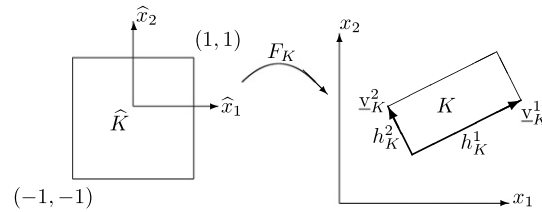


Fig. 1. Anisotropic directions of rectangle K .

pursued to derive an error estimator for hp -adaptive DG methods for convection–diffusion equations on isotropically refined meshes. This estimator has the distinct feature that it is robust in the Péclet number of the problem with respect to a suitably defined error measure (i.e., it is reliable and efficient with constants that are independent of the parameter ε).

The purpose of this paper is to extend the work [7] to anisotropically refined meshes, and to present an estimator η which yields global upper and lower bounds of the error measured in terms of a natural norm associated with diffusion and a semi-norm associated with convection. In particular, our error measure contains the standard DG energy norm and a variant of the dual norm introduced in [8] to measure convective effects. The constant in the lower bound is independent of ε and the mesh size, but weakly depending on the polynomial degrees, as in many hp -version error estimators for diffusion problems. In the upper bound, we use an alignment measure to incorporate the anisotropy of the underlying meshes in the reliability constant; see [9–11] and the references therein. As a consequence, the upper bound depends on the elemental aspect ratios and is not fully robust in the Péclet number, in contrast to the case of isotropic elements considered in [7]. Our analysis is valid for 1-irregularly refined rectangular elements with arbitrarily large aspect ratios, and is based on the hp -version averaging operator of [7], but with anisotropically scaled approximation properties.

We present a series of numerical experiments to test the feasibility of this approach within a fully automated hp -adaptive algorithm. Our tests indicate that internal and boundary layers are correctly captured and resolved at exponential rates of convergence in the number of degrees of freedom. We further observe that as soon as a reasonable h -resolution of the layers is achieved, the alignment measure is of moderate size, and the ratios of the error estimators and the energy errors are practically independent of the diffusion parameter ε and the mesh size. In all the tests, our new hp -version anisotropic refinement strategy outperforms similar strategies based on isotropic mesh refinement by orders of magnitude.

Let us also point out that in [12,13], a duality-based a-posteriori approach was successfully proposed and studied for hp -adaptive DG methods for convection–diffusion problems on anisotropically refined meshes and with anisotropically enriched elemental polynomial degrees.

The outline of the rest of the paper is as follows. In Section 2, we introduce hp -adaptive discontinuous Galerkin methods for the discretization of the convection–diffusion problem (1). In Section 3, we state and discuss our a-posteriori error estimates. The proof of these estimates is carried out in Section 4. In Section 5, we present a series of numerical tests illustrating the performance of a fully automated hp -adaptive algorithm. Finally, in Section 6, we end with some concluding remarks.

Throughout the paper, we shall frequently use the symbols \lesssim and \gtrsim to denote bounds that are valid up to positive constants, independently of the local mesh sizes, the elemental aspect ratios, the elemental polynomial degrees, and the parameter ε .

2. Interior penalty discretization

In this section, we introduce an hp -version interior penalty DG finite element method for the discretization of Eq. (1) on anisotropically refined meshes.

2.1. Elements and meshes

We consider (a family of) partitions \mathcal{T} of Ω into disjoint rectangular elements $\{K\}$. Each element is the image of the reference square $\hat{K} = (-1, 1)^2$ under an affine elemental mapping F_K . We allow for 1-irregularly refined meshes, where each elemental edge may contain at most one hanging node located in the middle of the edge. For each rectangle $K \in \mathcal{T}$, we denote by \underline{v}_K^1 and \underline{v}_K^2 its two anisotropic directions, as shown in Fig. 1. With the direction vectors, we associate the matrix

$$\mathbf{M}_K = [\underline{v}_K^1, \underline{v}_K^2]. \quad (4)$$

The lengths of the direction vectors are denoted by h_K^1 and h_K^2 , respectively. Then we define the minimum and maximum diameters of an element K by

$$h_{\min, K} = \min\{h_K^1, h_K^2\}, \quad h_{\max, K} = \max\{h_K^1, h_K^2\}. \quad (5)$$

We denote by $\mathcal{N}(K)$ the set of four vertices of K , and define $\mathcal{N}(\mathcal{T}) = \cup_{K \in \mathcal{T}} \mathcal{N}(K)$. We further split the set of all nodes into interior nodes and boundary nodes, that is, we write $\mathcal{N}(\mathcal{T}) = \mathcal{N}_I(\mathcal{T}) \cup \mathcal{N}_B(\mathcal{T})$. We denote by $\mathcal{E}(K)$ the set of four elemental edges of an element K . The length of an elemental edge is denoted by h_E , i.e., $h_E = h_K^i$ if E is parallel to \underline{v}_K^i , $i = 1, 2$.

The non-empty intersection $E = \partial K \cap \partial K'$ of two neighboring elements $K, K' \in \mathcal{T}$ is called an interior edge of \mathcal{T} . The set of all interior edges is denoted by $\mathcal{E}_I(\mathcal{T})$. Analogously, the non-empty intersection $E = \partial K \cap \Gamma$ of an element $K \in \mathcal{T}$ with the boundary Γ is called a boundary edge of \mathcal{T} . The set of all boundary edges of \mathcal{T} is denoted by $\mathcal{E}_B(\mathcal{T})$. Moreover, we set $\mathcal{E}(\mathcal{T}) = \mathcal{E}_I(\mathcal{T}) \cup \mathcal{E}_B(\mathcal{T})$.

Let now $E \in \mathcal{E}(\mathcal{T})$ be part of an elemental edge of element K . Then we denote by $h_{E,K}^\perp$ the width of element K perpendicular to E . That is, $h_{E,K}^\perp = h_K^{3-i}$ if E is parallel to \underline{v}_K^i , $i = 1, 2$. We then make the following bounded local variation assumption: there is a constant $\rho_1 \geq 1$ independent of the particular mesh in the mesh family, such that

$$\rho_1^{-1} \leq h_{E,K}^\perp / h_{E,K'}^\perp \leq \rho_1, \quad (6)$$

for all edges $E \in \mathcal{E}_I(\mathcal{T})$ shared by elements K and K' . Moreover, for $E \in \mathcal{E}(\mathcal{T})$ we define

$$h_E^\perp = \begin{cases} \min\{h_{E,K}^\perp, h_{E,K'}^\perp\}, & E \in \mathcal{E}_I(\mathcal{T}), E = \partial K \cap \partial K', \\ h_{E,K}^\perp, & E \in \mathcal{E}_B(\mathcal{T}), E = \partial K \cap \Gamma, \end{cases} \quad (7)$$

and

$$h_{\min,E} = \begin{cases} \min\{h_{\min,K}, h_{\min,K'}\}, & E \in \mathcal{E}_I(\mathcal{T}), E = \partial K \cap \partial K', \\ h_{\min,K}, & E \in \mathcal{E}_B(\mathcal{T}), E = \partial K \cap \Gamma. \end{cases} \quad (8)$$

Remark 1. Assumption (6) and the fact that the meshes are 1-irregular imply that there is a constant $C \geq 1$ independent of the particular mesh in the mesh family, such that

$$C^{-1} \leq h_E^\perp / h_{E,K}^\perp \leq C, \quad C^{-1} \leq h_{\min,E} / h_{\min,K} \leq C, \quad (9)$$

for any edge E which is part of an elemental edge of K .

2.2. Polynomial degrees and finite element spaces

With each element $K \in \mathcal{T}$, we associate a polynomial degree $p_K \geq 1$. We store these degrees in the vector $\mathbf{p} = \{p_K : K \in \mathcal{T}\}$, and set $|\mathbf{p}| = \max_{K \in \mathcal{T}} p_K$. We assume that \mathbf{p} is also of bounded local variation: there is a second constant $\rho_2 \geq 1$ independent of the particular mesh in the mesh family, such that

$$\rho_2^{-1} \leq p_K / p_{K'} \leq \rho_2, \quad (10)$$

for any pairs of neighboring elements K, K' in \mathcal{T} . For $E \in \mathcal{E}(\mathcal{T})$, we introduce the edge polynomial degree p_E by

$$p_E = \begin{cases} \max\{p_K, p_{K'}\}, & E = \partial K \cap \partial K' \in \mathcal{E}_I(\mathcal{T}), \\ p_K, & E = \partial K \cap \Gamma \in \mathcal{E}_B(\mathcal{T}). \end{cases} \quad (11)$$

The hp -version discontinuous Galerkin finite element space is now given by

$$\mathbf{S}_{\mathbf{p}}(\mathcal{T}) = \{v \in L^2(\Omega) : v|_K \circ F_K \in \mathcal{Q}_{p_K}(\widehat{K}), K \in \mathcal{T}\},$$

with $\mathcal{Q}_{p_K}(\widehat{K})$ denoting the set of all polynomials on the reference square \widehat{K} of degree at most p_K in each coordinate direction.

2.3. Discretization

We consider the following discontinuous Galerkin method for the approximation of the convection–diffusion problem (1): Find $u_{hp} \in \mathbf{S}_{\mathbf{p}}(\mathcal{T})$ such that

$$A_{hp}(u_{hp}, v) = \int_{\Omega} f v \, dx \quad (12)$$

for all $v \in \mathbf{S}_{\mathbf{p}}(\mathcal{T})$, with the bilinear form A_{hp} given by

$$\begin{aligned} A_{hp}(u, v) = & \sum_{K \in \mathcal{T}} \int_K (\varepsilon \nabla u \cdot \nabla v + \underline{a} \cdot \nabla u v) \, dx - \sum_{E \in \mathcal{E}(\mathcal{T})} \int_E \{\{\varepsilon \nabla u\}\} \cdot \llbracket v \rrbracket \, ds - \sum_{E \in \mathcal{E}(\mathcal{T})} \int_E \{\{\varepsilon \nabla v\}\} \cdot \llbracket u \rrbracket \, ds \\ & + \sum_{E \in \mathcal{E}(\mathcal{T})} \int_E \gamma \frac{\varepsilon p_E^2}{h_E^\perp} \llbracket u \rrbracket \cdot \llbracket v \rrbracket \, ds - \sum_{K \in \mathcal{T}} \int_{\partial K_{\text{in}} \cap \Gamma_{\text{in}}} \underline{a} \cdot \underline{n}_K u v \, ds + \sum_{K \in \mathcal{T}} \int_{\partial K_{\text{in}} \setminus \Gamma} \underline{a} \cdot \underline{n}_K (u^e - u) v \, ds. \end{aligned}$$

Here, the operators $\{\{\cdot\}\}$ and $\llbracket \cdot \rrbracket$ denote the usual averages and jumps of piecewise smooth functions across edges of \mathcal{T} ; see [7, Section 2.2] for their explicit definitions. Note that for a piecewise smooth function, the operator ∇ has to be understood as the broken gradient. Furthermore, we denote by u^e the trace of u on an elemental boundary taken from

the exterior, and by Γ_{in} and ∂K_{in} the inflow parts of Γ and $K \in \mathcal{T}$, respectively:

$$\Gamma_{\text{in}} = \{x \in \Gamma : \underline{a}(x) \cdot \underline{n}(x) < 0\}, \quad \partial K_{\text{in}} = \{x \in \partial K : \underline{a}(x) \cdot \underline{n}_K(x) < 0\}.$$

Finally, the constant $\gamma > 0$ is the interior penalty parameter.

The variational problem (12) is uniquely solvable, provided that the parameter γ is chosen sufficiently large, independently of the local mesh sizes, the elemental aspect ratios, the elemental polynomial degrees, and the parameter ε ; see, e.g., [7,14,15] and the references therein.

3. A-posteriori error estimates

In this section, our main results are presented and discussed.

3.1. Norms

We begin by introducing the standard energy norm associated with the discontinuous Galerkin discretization of the diffusion term:

$$\begin{aligned} \|v\|_{E,\mathcal{T}}^2 &= \sum_{K \in \mathcal{T}} \varepsilon \|\nabla v\|_{L^2(K)}^2 + \text{ejump}_{\mathbf{p},\mathcal{T}}(v)^2, \\ \text{ejump}_{\mathbf{p},\mathcal{T}}(v)^2 &= \sum_{E \in \mathcal{E}(\mathcal{T})} \varepsilon \gamma \frac{p_E^2}{h_E^\perp} \|\llbracket v \rrbracket\|_{L^2(E)}^2. \end{aligned} \quad (13)$$

Under assumption (2) and for γ sufficiently large, the DG form A_{hp} is coercive over the finite element space $S_{\mathbf{p}}(\mathcal{T})$ with respect to the energy norm.

To measure the effects of convection, we use a variant of the dual norm introduced in [8], namely the following semi-norm

$$|v|_\star = \sup_{w \in H_0^1(\Omega) \setminus \{0\}} \frac{\int_\Omega (\underline{a}v) \cdot \nabla w \, dx}{\|w\|_{E,\mathcal{T}}}. \quad (14)$$

In the sequel, we shall refer to $|\cdot|_\star$ as the convective semi-norm. Note that since $\|\underline{a}\|_{L^\infty(\Omega)}$ is assumed to be of order one, we have

$$|v|_\star \lesssim \varepsilon^{-1/2} \|v\|_{L^2(\Omega)}. \quad (15)$$

Finally, we shall introduce the following semi-norm involving the inter-elemental jumps:

$$\text{ojump}_{\mathbf{p},\mathcal{T}}(v)^2 = \sum_{E \in \mathcal{E}(\mathcal{T})} \left(\varepsilon p_E^2 \frac{h_E^\perp}{h_{\min,E}^2} + \frac{h_E^\perp}{\varepsilon p_E} \right) \|\llbracket v \rrbracket\|_{L^2(E)}^2. \quad (16)$$

In (16), the expressions weighting the L^2 -norms of the jumps are related to diffusion and convection as well. Indeed, the weights $\varepsilon \gamma p_E^2 \frac{h_E^\perp}{h_{\min,E}^2}$ are needed to account for the anisotropy of the meshes with respect to diffusive terms; they are motivated by the scaling properties of the averaging operator in Theorem 12. In the case of isotropic elements, they coincide with the weights in (13) (up to the interior penalty parameter γ); see also [7]. On the other hand, the weights $\varepsilon^{-1} p_E^{-1} h_E^\perp$ represent cell Péclet numbers in the direction perpendicular to E , and are associated with convection.

In what follows, it will be convenient to define

$$|v|_{0,\mathcal{T}}^2 = |v|_\star^2 + \text{ojump}_{\mathbf{p},\mathcal{T}}(v)^2. \quad (17)$$

In addition to the standard energy norm (13), we shall use the semi-norm (17) as part of our error measure.

3.2. Error estimators and data approximation

Let now $u_{hp} \in S_{\mathbf{p}}(\mathcal{T})$ be the discontinuous Galerkin approximation obtained by (12). Moreover, let f_{hp} and \underline{a}_{hp} denote piece polynomial approximations in $S_{\mathbf{p}}(\mathcal{T})$ and $S_{\mathbf{p}}(\mathcal{T})^2$ to the right-hand side f and the flow field \underline{a} , respectively. For example, these approximations can be taken as L^2 -projections into $S_{\mathbf{p}}(\mathcal{T})$ and $S_{\mathbf{p}}(\mathcal{T})^2$.

For each element $K \in \mathcal{T}$, we then introduce a local error indicator η_K , which is given by the sum of the three terms

$$\eta_K^2 = \eta_{R_K}^2 + \eta_{E_K}^2 + \eta_{J_K}^2. \quad (18)$$

The first term η_{R_K} is the interior residual defined by

$$\eta_{R_K}^2 = \varepsilon^{-1} p_K^{-2} h_{\min,K}^2 \|f_{hp} + \varepsilon \Delta u_{hp} - \underline{a}_{hp} \cdot \nabla u_{hp}\|_{L^2(K)}^2. \quad (19)$$

The second term η_{E_K} is the edge residual given by

$$\eta_{E_K}^2 = \frac{1}{2} \sum_{E \in \mathcal{E}(K)} \frac{h_{\min,K}^2}{\varepsilon p_E h_{E,K}^\perp} \|\llbracket \varepsilon \nabla u_{hp} \rrbracket\|_{L^2(E \setminus \Gamma)}^2. \quad (20)$$

The last residual η_{J_K} measures the error in the jumps of the approximate solution u_{hp} :

$$\begin{aligned} \eta_{J_K}^2 &= \frac{1}{2} \sum_{E \in \mathcal{E}(K)} \left(\frac{\varepsilon \gamma^2 p_E^3}{h_{E,K}^\perp} + \frac{\varepsilon h_{E,K}^\perp p_E^2}{h_{\min,K}^2} + \frac{h_{E,K}^\perp}{\varepsilon p_E} \right) \|\llbracket u_{hp} \rrbracket\|_{L^2(E \setminus \Gamma)}^2 \\ &\quad + \sum_{E \in \mathcal{E}(K)} \left(\frac{\varepsilon \gamma^2 p_E^3}{h_{E,K}^\perp} + \frac{\varepsilon h_{E,K}^\perp p_E^2}{h_{\min,K}^2} + \frac{h_{E,K}^\perp}{\varepsilon p_E} \right) \|\llbracket u_{hp} \rrbracket\|_{L^2(E \cap \Gamma)}^2. \end{aligned} \quad (21)$$

Note that the residual η_{J_K} contains the usual diffusive jumps as in (13) (but weighted with p_E^3 rather than p_E^2 as in [5,7]), along with the additional jump terms appearing in (16).

We also introduce the local data approximation term

$$\Theta_K^2 = \varepsilon^{-1} p_K^{-2} h_{\min,K}^2 \left(\|f - f_{hp}\|_{L^2(K)}^2 + \|(\underline{a} - \underline{a}_{hp}) \cdot \nabla u_{hp}\|_{L^2(K)}^2 \right),$$

and define our (global) error estimator and data approximation term by

$$\eta^2 = \sum_{K \in \mathcal{T}} \eta_K^2, \quad \Theta^2 = \sum_{K \in \mathcal{T}} \Theta_K^2. \quad (22)$$

3.3. A-posteriori estimates

The error estimator η in (22) is reliable up to a so-called alignment measure $\mathcal{M}(v, \mathcal{T})$. This notion was originally introduced in [10]; see also [9,11].

Definition 2. Let $v \in H^1(\Omega)$ be an arbitrary non-constant function and \mathcal{T} a triangulation of Ω . The alignment measure $\mathcal{M}(v, \mathcal{T})$ is then defined by

$$\mathcal{M}(v, \mathcal{T}) = \frac{\left(\sum_{K \in \mathcal{T}} h_{\min,K}^{-2} \|\mathbf{M}_K \nabla v\|_{L^2(K)}^2 \right)^{1/2}}{\|\nabla v\|_{L^2(\Omega)}}.$$

The expression $\mathcal{M}(v, \mathcal{T})$ measures how well the possibly anisotropic function v is aligned with the mesh \mathcal{T} . It also appears naturally in anisotropic interpolation estimates. We note that

$$1 \leq \mathcal{M}(v, \mathcal{T}) \lesssim \max_{K \in \mathcal{T}} \frac{h_{\max,K}}{h_{\min,K}} \quad \forall v \in H_0^1(\Omega).$$

Hence, for isotropic meshes the alignment measure is always of order one.

We are now ready to state our upper bound.

Theorem 3. Let u be the solution of (1) and $u_{hp} \in \mathbf{S}_p(\mathcal{T})$ its DG approximation obtained by (12). Let the error estimator η and the data approximation error Θ be defined by (22). Then we have the a-posteriori error bound

$$\|u - u_{hp}\|_{E,\mathcal{T}} + |u - u_{hp}|_{0,\mathcal{T}} \lesssim \mathcal{M}(v, \mathcal{T})(\eta + \Theta). \quad (23)$$

Here, $v \in H_0^1(\Omega)$ is the test function defined in the inf-sup condition (33).

Remark 4. We emphasize that the function $v \in H_0^1(\Omega)$ appearing in $\mathcal{M}(v, \mathcal{T})$ in the bound (23) is not the solution of problem (1). Instead, it is a test function related to the conforming part of the error, analogously to the analysis of [10] for the Poisson problem. As such, it is not possible to easily estimate or evaluate $\mathcal{M}(v, \mathcal{T})$ in a more explicit manner. However, we observe numerically that $\mathcal{M}(v, \mathcal{T})$ becomes of moderate size once anisotropic solution behavior is sufficiently well resolved. For additional discussions on the alignment measure, we refer the reader to [10,11] and the references therein.

Remark 5. Note that estimate (23) does not provide an upper bound for the L^2 -errors. However, estimate (15) implies that

$$|u - u_{hp}|_\star \lesssim \varepsilon^{-1/2} \|u - u_{hp}\|_{L^2(\Omega)}.$$

Our numerical experiments indicate that the estimators η overestimate this weighted L^2 -errors (and thus the L^2 -errors for small ε) for sufficiently well resolved layers. They also confirm that the standard L^2 -errors converge exponentially, with convergence plots that are qualitatively very similar to those in the energy errors; see Section 5.

The proof of [Theorem 3](#) is presented in [Section 4](#). It is based on using an hp -version anisotropic averaging operator as in [\[7\]](#) and a uniform inf–sup condition as in [\[8\]](#) (see [Lemma 7](#)).

Our next theorem states a lower bound.

Theorem 6. *Let u be the solution of [\(1\)](#) and $u_{hp} \in S_p(\mathcal{T})$ its DG approximation obtained by [\(12\)](#). Let the error estimator η and the data approximation error Θ be defined by [\(22\)](#). Then for any $\delta \in (0, \frac{1}{2})$, we have the bound*

$$\eta \lesssim |\mathbf{p}|^{\delta+1} \|u - u_{hp}\|_{E,\mathcal{T}} + |\mathbf{p}|^{2\delta+1} |u - u_{hp}|_{0,\mathcal{T}} + |\mathbf{p}|^{2\delta+\frac{1}{2}} \Theta.$$

As in [\[1,6,7,16\]](#), the efficiency bound in [Theorem 6](#) is suboptimal with respect to the polynomial degree due to the use of inverse estimates (which are suboptimal in the polynomial order). The proof of [Theorem 6](#) follows along the same lines, taking into account anisotropic scaling. For the sake of brevity, we omit it, and instead refer to [\[7\]](#) and [\[17, Section 5.4\]](#) for details.

4. Proofs

In this section, we present the proof of [Theorem 3](#).

4.1. Stability

The following uniform inf–sup condition for the form A is the crucial stability result in our analysis; it holds with an absolute constant.

Lemma 7. *Assume [\(2\)](#). Then we have*

$$\inf_{u \in H_0^1(\Omega) \setminus \{0\}} \sup_{v \in H_0^1(\Omega) \setminus \{0\}} \frac{A(u, v)}{(\|u\|_{E,\mathcal{T}} + |u|_\star) \|v\|_{E,\mathcal{T}}} \geq \frac{1}{3}.$$

For the proof, we refer to [\[18, Lemma 4.4\]](#); see also [\[8\]](#).

4.2. Auxiliary forms

Next, we split the discontinuous Galerkin form A_{hp} into two parts, and define

$$\begin{aligned} \tilde{A}_{hp}(u, v) &= \sum_{K \in \mathcal{T}} \int_K (\varepsilon \nabla u \cdot \nabla v + \underline{a} \cdot \nabla uv) dx + \sum_{E \in \mathcal{E}(\mathcal{T})} \int_E \frac{\varepsilon \gamma p_E^2}{h_E^\perp} \llbracket u \rrbracket \cdot \llbracket v \rrbracket ds \\ &\quad - \sum_{K \in \mathcal{T}} \int_{\partial K_{in} \cap \Gamma_{in}} \underline{a} \cdot \underline{n}_K uv ds + \sum_{K \in \mathcal{T}} \int_{\partial K_{in} \setminus \Gamma} \underline{a} \cdot \underline{n}_K (u^e - u) v ds, \\ K_{hp}(u, v) &= - \sum_{E \in \mathcal{E}(\mathcal{T})} \int_E \{\{\varepsilon \nabla u\}\} \cdot \llbracket v \rrbracket ds - \sum_{E \in \mathcal{E}(\mathcal{T})} \int_E \{\{\varepsilon \nabla v\}\} \cdot \llbracket u \rrbracket ds. \end{aligned}$$

We shall use these auxiliary forms to express both the continuous form A in [\(3\)](#) and the discontinuous Galerkin form A_{hp} in [\(12\)](#). Indeed, we have

$$A(u, v) = \tilde{A}_{hp}(u, v), \quad u, v \in H_0^1(\Omega), \quad (24)$$

$$A_{hp}(u, v) = \tilde{A}_{hp}(u, v) + K_{hp}(u, v), \quad u, v \in S_p(\mathcal{T}). \quad (25)$$

4.3. Anisotropic interpolation

We will need the following anisotropic interpolation bounds.

Lemma 8. *For $v \in H_0^1(\Omega)$, there exists a function $v_{hp} \in S_p(\mathcal{T})$ such that*

$$\begin{aligned} p_K^2 \|v - v_{hp}\|_{L^2(K)}^2 &\lesssim \|\mathbf{M}_K \nabla v\|_{L^2(K)}^2, \\ \|\mathbf{M}_K \nabla (v - v_{hp})\|_{L^2(K)}^2 &\lesssim \|\mathbf{M}_K \nabla v\|_{L^2(K)}^2, \\ \sum_{E \in \mathcal{E}(K)} h_{E,K}^\perp p_E \|v - v_{hp}\|_{L^2(E)}^2 &\lesssim \|\mathbf{M}_K \nabla v\|_{L^2(K)}^2, \end{aligned} \quad (26)$$

for any $K \in \mathcal{T}$.

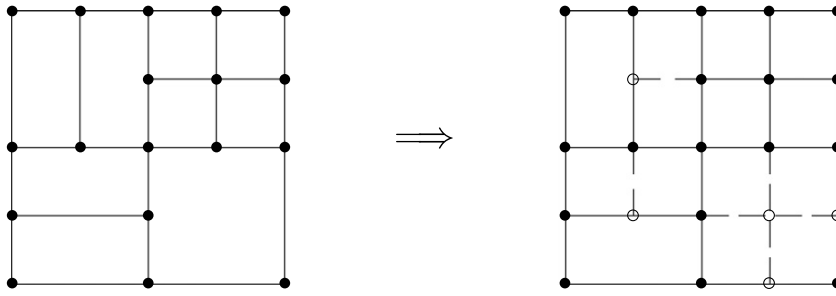


Fig. 2. The construction of the auxiliary mesh $\tilde{\mathcal{T}}$ from \mathcal{T} .

Proof. The first two inequalities follow from those in [5, Lemma 3.7] and anisotropic scaling. Next, consider an elemental edge E of K . By using the anisotropically scaled multiplicative trace inequality,

$$\|v - v_{hp}\|_{L^2(E)}^2 \lesssim \frac{1}{h_{E,K}^\perp} \|\mathbf{M}_K \nabla(v - v_{hp})\|_{L^2(K)} \|v - v_{hp}\|_{L^2(K)} + \frac{1}{h_{E,K}^\perp} \|v - v_{hp}\|_{L^2(K)}^2,$$

the weighted Cauchy–Schwarz inequality, the fact that $p_E \geq 1$ and the previous two estimates, we find that

$$\begin{aligned} \|v - v_{hp}\|_{L^2(E)}^2 &\lesssim \frac{1}{h_{E,K}^\perp p_E} \|\mathbf{M}_K \nabla(v - v_{hp})\|_{L^2(K)}^2 + \frac{p_E}{h_{E,K}^\perp} \|v - v_{hp}\|_{L^2(K)}^2 \\ &\lesssim \frac{1}{h_{E,K}^\perp p_E} \left(\|\mathbf{M}_K \nabla v\|_{L^2(K)}^2 + \|\mathbf{M}_K \nabla v\|_{L^2(K)}^2 \right), \end{aligned}$$

which shows the third inequality. \square

From Lemma 8 and the definition of the alignment measure, we immediately obtain global interpolation bounds.

Lemma 9. For $v \in H_0^1(\Omega)$, there exists a function $v_{hp} \in S_{\mathbf{p}}(\mathcal{T})$ such that

$$\begin{aligned} \sum_{K \in \mathcal{T}} \frac{p_K^2}{h_{\min,K}^2} \|v - v_{hp}\|_{L^2(K)}^2 &\lesssim \mathcal{M}(v, \mathcal{T})^2 \|\nabla v\|_{L^2(\Omega)}^2, \\ \sum_{K \in \mathcal{T}} \sum_{E \in \mathcal{E}(K)} \frac{h_{E,K}^\perp p_E}{h_{\min,K}^2} \|v - v_{hp}\|_{L^2(E)}^2 &\lesssim \mathcal{M}(v, \mathcal{T})^2 \|\nabla v\|_{L^2(\Omega)}^2. \end{aligned} \quad (27)$$

4.4. Averaging

We refer to averaging as the approximation of a discontinuous finite element function by a continuous one. This can be achieved by assigning to each conforming degree of freedom the value obtained by averaging over all the values of the discontinuous function taken elementwise at the corresponding degree of freedom. This procedure affects in particular vertex and edge degrees of freedom, but not interior ones. While averaging is relatively straightforward for conforming meshes, see also [1], it introduces some technicalities when dealing with hanging nodes.

Here, we shall make use of the averaging operator constructed and analyzed in [7], but scaled anisotropically. To handle 1-irregular meshes, it involves an auxiliary 1-irregular mesh $\tilde{\mathcal{T}}$ of rectangles, obtained from \mathcal{T} as follows.

Let $K \in \mathcal{T}$. If all four elemental edges are edges of the mesh \mathcal{T} , that is, if $\mathcal{E}(K) \subseteq \mathcal{E}(\mathcal{T})$, we leave K untouched. Otherwise, at least one of the elemental edges of K , say E , contains hanging nodes. In this case, we replace K by the two or four rectangles obtained by bisecting the elemental edges of K . This construction is illustrated in Fig. 2. Clearly, the mesh $\tilde{\mathcal{T}}$ is a refinement of \mathcal{T} ; it is also 1-irregular. We denote by $\mathcal{R}(K)$ the elements in $\tilde{\mathcal{T}}$ that have been generated inside $K \in \mathcal{T}$. If K has not been refined, then $\mathcal{R}(K) = \{K\}$. Otherwise, the set $\mathcal{R}(K)$ consists of two or four newly created elements.

Then, we introduce the following auxiliary discontinuous Galerkin finite element space on the mesh $\tilde{\mathcal{T}}$:

$$S_{\tilde{\mathbf{p}}}(\tilde{\mathcal{T}}) = \{v \in L^2(\Omega) : v|_{\tilde{K}} \circ F_{\tilde{K}} \in \mathcal{Q}_{p_{\tilde{K}}}(\tilde{K}), \tilde{K} \in \tilde{\mathcal{T}}\},$$

where the auxiliary polynomial degree vector $\tilde{\mathbf{p}}$ is defined by $p_{\tilde{K}} = p_K$ for $\tilde{K} \in \mathcal{R}(K)$. We have the inclusion $S_{\mathbf{p}}(\mathcal{T}) \subseteq S_{\tilde{\mathbf{p}}}(\tilde{\mathcal{T}})$. As in (13) and (17), we set

$$\begin{aligned} \|v\|_{E,\tilde{\mathcal{T}}}^2 &= \sum_{\tilde{K} \in \tilde{\mathcal{T}}} \varepsilon \|\nabla v\|_{L^2(\tilde{K})}^2 + \text{ejump}_{\tilde{\mathbf{p}},\tilde{\mathcal{T}}}(v)^2, \\ |v|_{0,\tilde{\mathcal{T}}}^2 &= |v|_\star^2 + \text{ojump}_{\tilde{\mathbf{p}},\tilde{\mathcal{T}}}(v)^2, \end{aligned} \quad (28)$$

where the jump weights are defined analogously to (7), (11), but with respect to the auxiliary mesh $\tilde{\mathcal{T}}$ and degree vector $\tilde{\mathbf{p}}$. Obviously, we have

$$\|v\|_{E,\mathcal{T}} = \|v\|_{E,\tilde{\mathcal{T}}}, \quad |v|_{0,\mathcal{T}} = |v|_{0,\tilde{\mathcal{T}}},$$

for all $v \in H_0^1(\Omega)$. As in [7, Lemmas 4.2 and 4.3], the following results hold.

Lemma 10. Let $v \in S_{\tilde{\mathbf{p}}}(\tilde{\mathcal{T}}) + H_0^1(\Omega)$ be such that $\llbracket v \rrbracket|_E = \llbracket w \rrbracket|_E$ for all $E \in \mathcal{E}(\tilde{\mathcal{T}})$, for a function $w \in S_{\mathbf{p}}(\mathcal{T}) + H_0^1(\Omega)$. Then we have

$$\begin{aligned} \text{ejump}_{\mathbf{p},\mathcal{T}}(w) &\lesssim \text{ejump}_{\tilde{\mathbf{p}},\tilde{\mathcal{T}}}(v) \lesssim \text{ejump}_{\mathbf{p},\mathcal{T}}(w), \\ \text{ojump}_{\mathbf{p},\mathcal{T}}(w) &\lesssim \text{ojump}_{\tilde{\mathbf{p}},\tilde{\mathcal{T}}}(v) \lesssim \text{ojump}_{\mathbf{p},\mathcal{T}}(w). \end{aligned}$$

Lemma 11. For $v \in S_{\mathbf{p}}(\mathcal{T}) + H_0^1(\Omega)$, we have the bounds

$$\|v\|_{E,\mathcal{T}} \lesssim \|v\|_{E,\tilde{\mathcal{T}}}, \quad |v|_{0,\mathcal{T}} \lesssim |v|_{0,\tilde{\mathcal{T}}}.$$

Let $S_{\tilde{\mathbf{p}}}^c(\tilde{\mathcal{T}})$ be the conforming subspace of $S_{\tilde{\mathbf{p}}}(\tilde{\mathcal{T}})$ given by

$$S_{\tilde{\mathbf{p}}}^c(\tilde{\mathcal{T}}) = S_{\tilde{\mathbf{p}}}(\tilde{\mathcal{T}}) \cap H_0^1(\Omega).$$

We are now ready to state the following result regarding the averaging of a DG function. Due to the possible presence of hanging nodes, the averaged function will belong to the conforming space $S_{\tilde{\mathbf{p}}}^c(\tilde{\mathcal{T}})$ on the auxiliary mesh $\tilde{\mathcal{T}}$.

Theorem 12. There is an averaging operator $I_{hp} : S_{\mathbf{p}}(\mathcal{T}) \rightarrow S_{\tilde{\mathbf{p}}}^c(\tilde{\mathcal{T}})$ that satisfies

$$\sum_{\tilde{K} \in \tilde{\mathcal{T}}} \|v - I_{hp}v\|_{L^2(\tilde{K})}^2 \lesssim \sum_{E \in \mathcal{E}(\mathcal{T})} \int_E (p_E)^{-2} h_E^\perp \llbracket v \rrbracket^2 ds, \quad (29)$$

$$\sum_{\tilde{K} \in \tilde{\mathcal{T}}} \|\nabla(v - I_{hp}v)\|_{L^2(\tilde{K})}^2 \lesssim \sum_{E \in \mathcal{E}(\mathcal{T})} \int_E p_E^2 h_E^\perp h_{\min,E}^{-2} \llbracket v \rrbracket^2 ds. \quad (30)$$

The proof of Theorem 12 follows along the lines of [7, Section 5], but with the key Lemmas 5.3 and 5.4 there scaled anisotropically (which is readily achieved). For details, we refer to [17, Section 5.5].

4.5. Proof of Theorem 3

After these preliminary results, we now present the proof of Theorem 3. We follow [1,18], and decompose the discontinuous Galerkin solution into a conforming part and a remainder:

$$u_{hp} = u_{hp}^c + u_{hp}^r \quad \text{with } u_{hp}^c = I_{hp}u_{hp} \in S_{\tilde{\mathbf{p}}}^c(\tilde{\mathcal{T}}) \subset H_0^1(\Omega), \quad (31)$$

where I_{hp} is the averaging operator of Theorem 12. The remainder is then given by $u_{hp}^r = u_{hp} - u_{hp}^c = u_{hp} - I_{hp}u_{hp} \in S_{\tilde{\mathbf{p}}}(\tilde{\mathcal{T}})$. By Lemma 11 and the triangle inequality, we obtain

$$\begin{aligned} \|u - u_{hp}\|_{E,\mathcal{T}} + |u - u_{hp}|_{0,\mathcal{T}} &\lesssim \|u - u_{hp}\|_{E,\tilde{\mathcal{T}}} + |u - u_{hp}|_{0,\tilde{\mathcal{T}}} \\ &\lesssim \|u - u_{hp}^c\|_{E,\tilde{\mathcal{T}}} + |u - u_{hp}^c|_{0,\tilde{\mathcal{T}}} + \|u_{hp}^r\|_{E,\tilde{\mathcal{T}}} + |u_{hp}^r|_{0,\tilde{\mathcal{T}}} \\ &= \|u - u_{hp}^c\|_{E,\mathcal{T}} + |u - u_{hp}^c|_{0,\mathcal{T}} + \|u_{hp}^r\|_{E,\tilde{\mathcal{T}}} + |u_{hp}^r|_{0,\tilde{\mathcal{T}}}. \end{aligned} \quad (32)$$

It is now sufficient to show that both the conforming part $u - u_{hp}^c$ and the remainder u_{hp}^r can be bounded by the estimator η and the data approximation term Θ . We begin by bounding u_{hp}^r .

Lemma 13. There holds

$$\|u_{hp}^r\|_{E,\tilde{\mathcal{T}}} + |u_{hp}^r|_{0,\tilde{\mathcal{T}}} \lesssim \eta.$$

Proof. Since $\llbracket u_{hp}^r \rrbracket|_E = \llbracket u_{hp} \rrbracket|_E$ for all $E \in \mathcal{E}(\tilde{\mathcal{T}})$ and $u_{hp} \in S_{\mathbf{p}}(\mathcal{T})$, the definition of the jump residual η_{JK} and Lemma 10 yield

$$\begin{aligned} \|u_{hp}^r\|_{E,\tilde{\mathcal{T}}}^2 + |u_{hp}^r|_{0,\tilde{\mathcal{T}}}^2 &= \sum_{\tilde{K} \in \tilde{\mathcal{T}}} \varepsilon \|\nabla u_{hp}^r\|_{L^2(\tilde{K})}^2 + |u_{hp}^r|_\star^2 + \text{ejump}_{\tilde{\mathbf{p}},\tilde{\mathcal{T}}}(u_{hp}^r)^2 + \text{ojump}_{\tilde{\mathbf{p}},\tilde{\mathcal{T}}}(u_{hp}^r)^2 \\ &\lesssim \sum_{\tilde{K} \in \tilde{\mathcal{T}}} \varepsilon \|\nabla u_{hp}^r\|_{L^2(\tilde{K})}^2 + |u_{hp}^r|_\star^2 + \sum_{K \in \mathcal{T}} \eta_{JK}^2. \end{aligned}$$

Hence, only the volume terms and $|u_{hp}^r|_*$ need to be bounded further. [Theorem 12](#) and the equivalences (9), (10) yield

$$\varepsilon \sum_{\tilde{K} \in \tilde{\mathcal{T}}} \|\nabla u_{hp}^r\|_{L^2(\tilde{K})}^2 \lesssim \sum_{E \in \mathcal{E}(\mathcal{T})} \int_E \varepsilon p_E^2 h_E^\perp h_{\min,E}^{-2} \llbracket u_{hp} \rrbracket^2 ds \lesssim \sum_{K \in \mathcal{T}} \eta_{J_K}^2.$$

To estimate $|u_{hp}^r|_*$, we use the bound (15), [Theorem 12](#), the fact that $p_E \geq 1$, and the relations (9), (10). We obtain

$$|u_{hp}^r|_*^2 \lesssim \frac{1}{\varepsilon} \sum_{\tilde{K} \in \tilde{\mathcal{T}}} \|u_{hp}^r\|_{L^2(\tilde{K})}^2 \lesssim \sum_{E \in \mathcal{E}(\mathcal{T})} \frac{h_E^\perp}{\varepsilon p_E^2} \|\llbracket u_{hp} \rrbracket\|_{L^2(E)}^2 \lesssim \sum_{K \in \mathcal{T}} \eta_{J_K}^2.$$

This finishes the proof. \square

To bound the conforming errors in (32), we establish the following auxiliary result.

Lemma 14. For any $v \in H_0^1(\Omega)$, we have

$$\int_{\Omega} f(v - v_{hp}) dx - \tilde{A}_{hp}(u_{hp}, v - v_{hp}) + K_{hp}(u_{hp}, v_{hp}) \lesssim \mathcal{M}(v, \mathcal{T}) (\eta + \Theta) \|v\|_{E, \mathcal{T}},$$

where $v_{hp} \in S_p(\mathcal{T})$ is the hp -interpolant of v in [Lemma 8](#).

Proof. Integration by parts of the diffusive volume terms readily yields

$$\int_{\Omega} f(v - v_{hp}) dx - \tilde{A}_{hp}(u_{hp}, v - v_{hp}) + K_{hp}(u_{hp}, v_{hp}) = T_1 + T_2 + T_3 + T_4 + T_5,$$

where

$$\begin{aligned} T_1 &= \sum_{K \in \mathcal{T}} \int_K (f + \varepsilon \Delta u_{hp} - \underline{a} \cdot \nabla u_{hp})(v - v_{hp}) dx, \\ T_2 &= - \sum_{E \in \mathcal{E}_I(\mathcal{T})} \int_E \llbracket \varepsilon \nabla u_{hp} \rrbracket \{v - v_{hp}\} ds, \\ T_3 &= - \sum_{E \in \mathcal{E}(\mathcal{T})} \int_E \{ \varepsilon \nabla v_{hp} \} \cdot \llbracket u_{hp} \rrbracket ds, \\ T_4 &= \sum_{K \in \mathcal{T}} \int_{\partial K_{in} \setminus \Gamma} \underline{a} \cdot \underline{n}_K (u_{hp} - u_{hp}^e)(v - v_{hp}) ds + \sum_{K \in \mathcal{T}} \int_{\partial K_{in} \cap \Gamma_{in}} \underline{a} \cdot \underline{n}_K u_{hp} (v - v_{hp}) ds, \\ T_5 &= - \sum_{E \in \mathcal{E}(\mathcal{T})} \int_E \frac{\varepsilon \gamma p_E^2}{h_E^\perp} \llbracket u_{hp} \rrbracket \cdot \{v - v_{hp}\} ds. \end{aligned}$$

To bound T_1 , we first add and subtract the data approximations. From the weighted Cauchy–Schwarz inequality and the approximation properties in (27), we obtain

$$T_1 \lesssim \mathcal{M}(v, \mathcal{T}) \left(\sum_{K \in \mathcal{T}} (\eta_{R_K}^2 + \Theta_K^2) \right)^{\frac{1}{2}} \|v\|_{E, \mathcal{T}}.$$

Similarly, by the Cauchy–Schwarz inequality and (27), we have

$$\begin{aligned} T_2 &\lesssim \left(\sum_{K \in \mathcal{T}} \sum_{E \in \mathcal{E}(K)} \frac{h_{\min,K}^2}{\varepsilon p_E h_{E,K}^\perp} \|\llbracket \varepsilon \nabla u_{hp} \rrbracket\|_{L^2(E \setminus \Gamma)}^2 \right)^{\frac{1}{2}} \left(\sum_{K \in \mathcal{T}} \sum_{E \in \mathcal{E}(K)} \frac{\varepsilon p_E h_{E,K}^\perp}{h_{\min,K}^2} \|v - v_{hp}\|_{L^2(E)}^2 \right)^{\frac{1}{2}} \\ &\lesssim \mathcal{M}(v, \mathcal{T}) \left(\sum_{K \in \mathcal{T}} \eta_{E_K}^2 \right)^{\frac{1}{2}} \|v\|_{E, \mathcal{T}}. \end{aligned}$$

To estimate T_3 , we employ the Cauchy–Schwarz inequality and the trace inequality in [3, Lemma 3.1]. This results in

$$\begin{aligned} T_3 &\lesssim \left(\sum_{E \in \mathcal{E}(\mathcal{T})} \frac{\varepsilon p_E^2}{h_E^\perp} \|\llbracket u_{hp} \rrbracket\|_{L^2(E)}^2 \right)^{\frac{1}{2}} \left(\sum_{K \in \mathcal{T}} \sum_{E \in \mathcal{E}(K)} \frac{\varepsilon h_E^\perp}{p_E^2} \|\nabla v_{hp}\|_{L^2(E)}^2 \right)^{\frac{1}{2}} \\ &\lesssim \left(\sum_{K \in \mathcal{T}} \eta_{J_K}^2 \right)^{\frac{1}{2}} \left(\sum_{K \in \mathcal{T}} \varepsilon \|\nabla v_{hp}\|_{L^2(K)}^2 \right)^{\frac{1}{2}} \lesssim \left(\sum_{K \in \mathcal{T}} \eta_{J_K}^2 \right)^{\frac{1}{2}} \|v\|_{E, \mathcal{T}}. \end{aligned}$$

For T_4 , we use the boundedness of $\|\underline{a}\|_{L^\infty(\Omega)}$, and apply again the Cauchy–Schwarz inequality and (27). We get

$$\begin{aligned} T_4 &\lesssim \left(\sum_{K \in \mathcal{T}} \sum_{E \in \mathcal{E}(K)} \frac{h_{\min,K}^2}{\varepsilon p_E h_{E,K}^\perp} \|\llbracket u_{hp} \rrbracket\|_{L^2(E)}^2 \right)^{\frac{1}{2}} \left(\sum_{K \in \mathcal{T}} \sum_{E \in \mathcal{E}(K)} \frac{\varepsilon p_E h_{E,K}^\perp}{h_{\min,K}^2} \|v - v_{hp}\|_{L^2(E)}^2 \right)^{\frac{1}{2}} \\ &\lesssim \mathcal{M}(v, \mathcal{T}) \left(\sum_{K \in \mathcal{T}} \eta_{JK}^2 \right)^{\frac{1}{2}} \|v\|_{E, \mathcal{T}}. \end{aligned}$$

Finally, we have

$$\begin{aligned} T_5 &\lesssim \left(\sum_{K \in \mathcal{T}} \sum_{E \in \mathcal{E}(K)} \frac{\varepsilon \gamma^2 h_{\min,K}^2 p_E^3}{(h_{E,K}^\perp)^3} \|\llbracket u_{hp} \rrbracket\|_{L^2(E)}^2 \right)^{\frac{1}{2}} \left(\sum_{K \in \mathcal{T}} \sum_{E \in \mathcal{E}(K)} \frac{\varepsilon p_E h_{E,K}^\perp}{h_{\min,K}^2} \|v - v_{hp}\|_{L^2(E)}^2 \right)^{\frac{1}{2}} \\ &\lesssim \mathcal{M}(v, \mathcal{T}) \left(\sum_{K \in \mathcal{T}} \eta_{JK}^2 \right)^{\frac{1}{2}} \|v\|_{E, \mathcal{T}}. \end{aligned}$$

The above estimates for T_1 through T_5 yield the assertion. \square

Now, we bound the norms of the conforming part $u - u_{hp}^c$ in (32).

Lemma 15. *There holds:*

$$\|u - u_{hp}^c\|_{E, \mathcal{T}} + |u - u_{hp}^c|_{0, \mathcal{T}} \lesssim \mathcal{M}(v, \mathcal{T})(\eta + \Theta).$$

Proof. Since $u - u_{hp}^c \in H_0^1(\Omega)$, we have $|u - u_{hp}^c|_{0, \mathcal{T}} = |u - u_{hp}^c|_\star$. The inf–sup condition in Lemma 7 ensures the existence of a test function $v \in H_0^1(\Omega)$ such that

$$\|u - u_{hp}^c\|_{E, \mathcal{T}} + |u - u_{hp}^c|_{0, \mathcal{T}} \lesssim A(u - u_{hp}^c, v) \quad \text{and} \quad \|v\|_{E, \mathcal{T}} \leq 1. \quad (33)$$

Then, property (24) shows that

$$A(u - u_{hp}^c, v) = \int_{\Omega} f v \, dx - A_{hp}(u_{hp}^c, v) = \int_{\Omega} f v \, dx - \tilde{A}_{hp}(u_{hp}^c, v).$$

By employing the fact that $v \in H_0^1(\Omega)$ and integrating by parts the convection term, one finds that

$$\tilde{A}_{hp}(u_{hp}^c, v) = \tilde{A}_{hp}(u_{hp}, v) + R,$$

with

$$R = \sum_{\tilde{K} \in \tilde{\mathcal{T}}} \int_{\tilde{K}} (-\varepsilon \nabla u_{hp}^r + \underline{a} u_{hp}^r) \cdot \nabla v \, dx.$$

From the DG method (12) and property (25), it follows that

$$\int_{\Omega} f v_{hp} \, dx = \tilde{A}_{hp}(u_{hp}, v_{hp}) + K_{hp}(u_{hp}, v_{hp}),$$

where $v_{hp} \in S_p(\mathcal{T})$ is the hp -version interpolant of v in Lemma 8. Combining the above results yields

$$A(u - u_{hp}^c, v) = \int_{\Omega} f(v - v_{hp}) \, dx - \tilde{A}_{hp}(u_{hp}, v - v_{hp}) + K_{hp}(u_{hp}, v_{hp}) - R.$$

The estimate in Lemma 14 now shows that

$$|A(u - u_{hp}^c, v)| \lesssim \mathcal{M}(v, \mathcal{T})(\eta + \Theta) \|v\|_{E, \mathcal{T}} + |R|. \quad (34)$$

It remains to bound $|R|$. From the Cauchy–Schwarz inequality, the definition of the convective semi-norm $|\cdot|_\star$, the conformity of v and Lemma 13, we conclude that

$$|R| \lesssim (\|u_{hp}^r\|_{E, \tilde{\mathcal{T}}} + |u_{hp}^r|_{0, \tilde{\mathcal{T}}}) \|v\|_{E, \mathcal{T}} \lesssim \eta \|v\|_{E, \mathcal{T}}. \quad (35)$$

Eqs. (33) through (35) imply the desired result. \square

The proof of Theorem 3 now is a consequence of inequality (32), Lemmas 13 and 15.

5. Numerical experiments

We present a series of numerical examples where we use the error indicator η in (22) to drive a fully automated hp -adaptive refinement strategy. All computations are performed using the AptoFEM software package; see [19] for details. The resulting systems of linear equations are solved by exploiting the parallel multifrontal solver MUMPS; see [20–22], for example.

In our numerics below, we compare an anisotropic hp -adaptive scheme against an isotropic one, which is obtained by restricting the estimator η to isotropically refined meshes. In the isotropic case, we recall that Theorem 3 is valid without an alignment measure; cf. [7]. On the other hand, the adaptive resolution of boundary layers using isotropic refinement is generally much less robust and may be prohibitively expensive. In both schemes the meshes are adapted by marking the elements for refinement according to the size of the local error indicators η_K ; this is achieved by employing the fixed fraction strategy, see [23], with refinement fraction set to 25% and derefinement fraction to 10%. That is, the top 25% fraction of elements with the largest indicators η_K is marked for refinement, and the bottom 10% one with the smallest indicators for derefinement. For each marked element, the schemes automatically decide whether the local mesh size h_K or the local degree p_K should be adjusted accordingly. The choice to perform either h - or p -refinement is based on estimating the local smoothness of the (unknown) analytical solution. To this end, we employ the hp -adaptive strategy developed in [24], where the local regularity of the analytical solution is estimated from truncated local Legendre expansions of the computed numerical solution; see also [25,26]. In the anisotropic hp -scheme, we also need to decide whether to perform isotropic or anisotropic h -refinement. To make this decision, we denote by E_K^1, E_K^2 the two sets containing the edges of K parallel to either \underline{v}_K^1 or \underline{v}_K^2 , and then define

$$\eta_{E_K^i}^2 = \eta_{E_K^i}^2|_{E_K^i} + \eta_{J_K^i}^2|_{E_K^i} \quad i = 1, 2.$$

Then the choice between isotropic or anisotropic h -refinement is made by comparing $\eta_{E_K^1}$ to $\eta_{E_K^2}$: if $\eta_{E_K^1} > 10 \eta_{E_K^2}$, then the element K is refined anisotropically along the direction \underline{v}_K^1 . On the other hand, if $\eta_{E_K^2} > 10 \eta_{E_K^1}$, then the element K is refined along the direction \underline{v}_K^2 . If none of these two conditions is met, the element K is refined isotropically. The derefinement procedure is the same for both schemes, and consists in simply undoing the last refinement made to the element.

In all our tests, we set the stabilization parameter to $\gamma = 10$. The approximate right-hand side f_{hp} is taken as the L^2 -projection of f onto $S_p(\mathcal{T})$. The flow fields considered are constant or polynomial vector fields. Hence, the volume residuals η_{RK} can always be integrated exactly by taking $\underline{a}_{hp} = \underline{a}$. We then neglect the data approximation term Θ in (22).

5.1. Example 1

We take $\Omega = (0, 1)^2$, choose the constant convection $\underline{a} = (1, 1)^\top$, and select the right-hand side f so that the solution to problem (1) is given by

$$u(x_1, x_2) = \left(\frac{e^{(x_1-1)/\varepsilon} - 1}{e^{-1/\varepsilon} - 1} + x_1 - 1 \right) \left(\frac{e^{(x_2-1)/\varepsilon} - 1}{e^{-1/\varepsilon} - 1} + x_2 - 1 \right).$$

The solution is analytic, but has boundary layers along the coordinate directions $x_1 = 1$ and $x_2 = 1$; their widths are both of order $\mathcal{O}(\varepsilon)$. This problem is well-suited to test whether the indicator η is able to pick up the steep gradients near these boundaries using anisotropic refinement.

We test this problem for $\varepsilon = 10^{-3}$, $\varepsilon = 10^{-4}$ and $\varepsilon = 10^{-6}$. For $\varepsilon = 10^{-3}$, we begin the test with a uniform mesh of size 4×4 and the uniform polynomial degree $p_K = 2$, and for $\varepsilon = 10^{-4}$, $\varepsilon = 10^{-6}$, with an 8×8 mesh and $p_K = 2$. In Fig. 3, we show the convergence of the estimators η , along with the energy norm errors $\|u - u_{hp}\|_{E,\mathcal{T}}$, the weighted L^2 -errors $\varepsilon^{-1/2} \|u - u_{hp}\|_{L^2(\Omega)}$ (which by Remark 5 bound the errors in the convective semi-norm $|u - u_{hp}|_\star$), and the jump errors $\text{ojump}_{p,\mathcal{T}}(u - u_{hp})$. We notice that the estimators provide upper bounds for the energy and jump errors, in agreement with Theorem 3. They also overestimate the weighted L^2 -errors (which is not guaranteed by Theorem 3). On the basis of the a-priori analysis in [27] or [28, Section 3.4.6, p. 118], we plot the errors against $N^{\frac{1}{2}}$, where N is the number of degrees of freedom. In the asymptotic regime and in a semi-logarithmic scale, all the curves are roughly straight lines, indicating exponential convergence in $N^{\frac{1}{2}}$. We observe that the asymptotic regime is achieved once the layers are sufficiently well resolved.

In Fig. 4, we compute the effectivity indices with respect to the DG energy norm errors, that is, the quantities $\eta/\|u - u_{hp}\|_{E,\mathcal{T}}$. Again, note that the estimators η actually bound a stronger norm; see Theorem 3. After a few iterations, the numerical values seem to settle in around 5, for all values of ε considered. This indicates that the alignment measure eventually becomes of moderate size once the layers are correctly captured. In this regime, the effectivity indices are relatively uniform in the number of iterations, similar to a pure diffusion problem.

In Fig. 5, we compare the DG energy norm errors obtained for the isotropic and anisotropic algorithms. Once the layers are properly captured, we expect exponential convergence in both cases. However, resolving the layers is more costly for the isotropic scheme. Indeed, for $\varepsilon = 10^{-3}$ and $\varepsilon = 10^{-4}$, it can be seen that both methods converge exponentially, but the

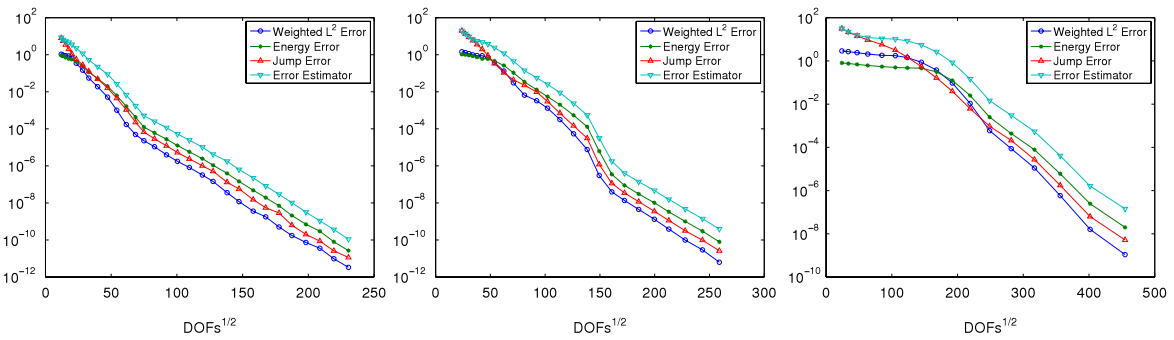


Fig. 3. Example 1: error plots for $\varepsilon = 10^{-3}$ (left), $\varepsilon = 10^{-4}$ (middle), and $\varepsilon = 10^{-6}$ (right).

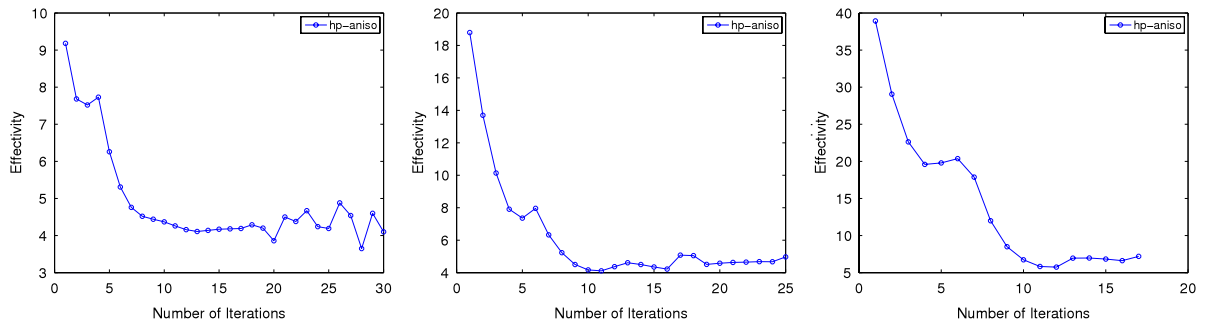


Fig. 4. Example 1: effectivity indices for $\varepsilon = 10^{-3}$ (left), $\varepsilon = 10^{-4}$ (middle), and $\varepsilon = 10^{-6}$ (right).

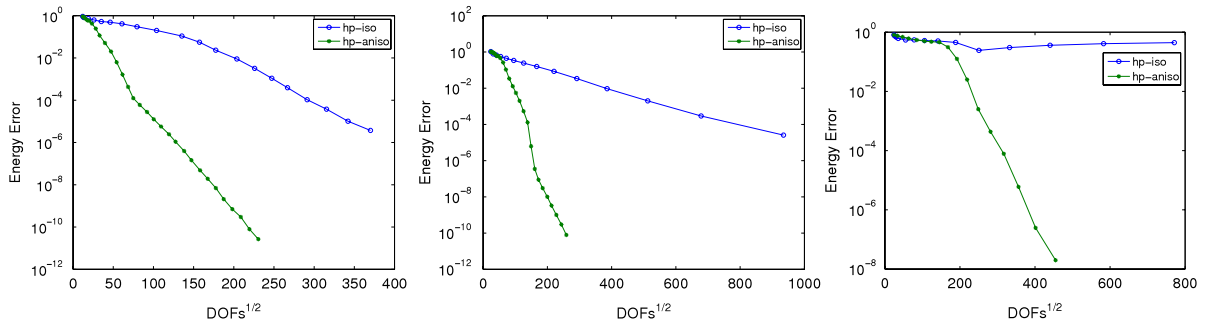


Fig. 5. Example 1: comparison of the DG energy errors for isotropic and anisotropic refinement for $\varepsilon = 10^{-3}$ (left), $\varepsilon = 10^{-4}$ (middle), and $\varepsilon = 10^{-6}$ (right).

anisotropic *hp*-algorithm outperforms the isotropic one by orders of magnitude. This is more pronounced for $\varepsilon = 10^{-4}$. In the case $\varepsilon = 10^{-6}$, the isotropic scheme is not able anymore to properly resolve the layers using a reasonable amount of degrees of freedom. As a result, the convergence plot stagnates while the anisotropic *hp*-scheme still converges exponentially.

As discussed in Remark 5, our estimator does not control the L^2 -norm errors. Nevertheless, the numerical results in Fig. 6 indicate that the L^2 -norm convergence is qualitatively very similar to the energy norm convergence depicted in Fig. 5. As before, we observe exponential convergence rates, the anisotropic schemes yield much smaller errors than the isotropic ones, and the isotropic curve stagnates for $\varepsilon = 10^{-6}$.

In Fig. 7, we show the final adapted meshes for both schemes and for $\varepsilon = 10^{-3}$. The colors indicate the order of polynomials used in each element; they are ranging between 2 and 11. In both cases, the adaptive procedure correctly captures the location and orientation of the boundary layers, and the meshes are refined accordingly. Particularly in the anisotropic case, we notice that relatively large polynomial degrees are applied near the boundaries. This is consistent with the theoretical results in [27] or [28, Section 3.4.6, p. 118]. Indeed, since the solution is analytic and once the layers are resolved, *p*-refinement is the most effective refinement strategy.

In Fig. 8, we show the final anisotropically adapted mesh for $\varepsilon = 10^{-4}$. Due to the presence of the strong layers, most of the adaptivity is performed very close to the right and upper boundaries of the domain. In order to better appreciate the adaptation of the mesh, we have magnified the region $(0.75, 1) \times (0.75, 1)$ in the upper-right corner of the domain. We observe strong anisotropic refinement along the layers. Qualitatively similar meshes are obtained for $\varepsilon = 10^{-6}$.

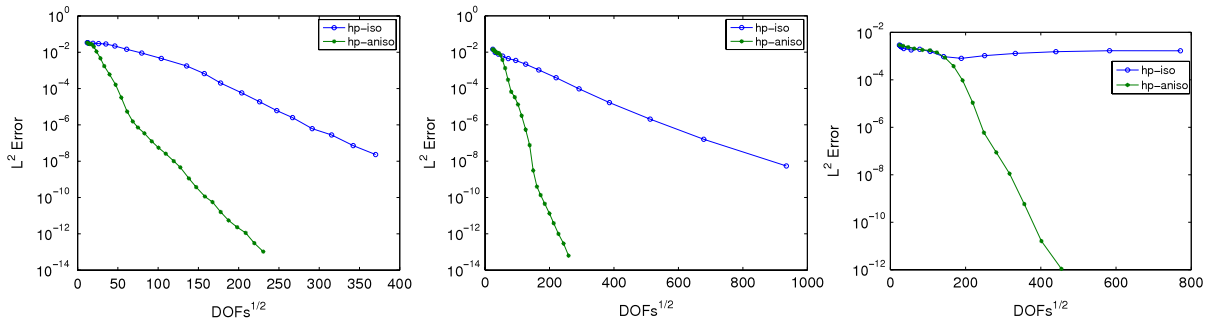


Fig. 6. Example 1: comparison of the L^2 -errors for isotropic and anisotropic refinement for $\varepsilon = 10^{-3}$ (left), $\varepsilon = 10^{-4}$ (middle), and $\varepsilon = 10^{-6}$ (right).

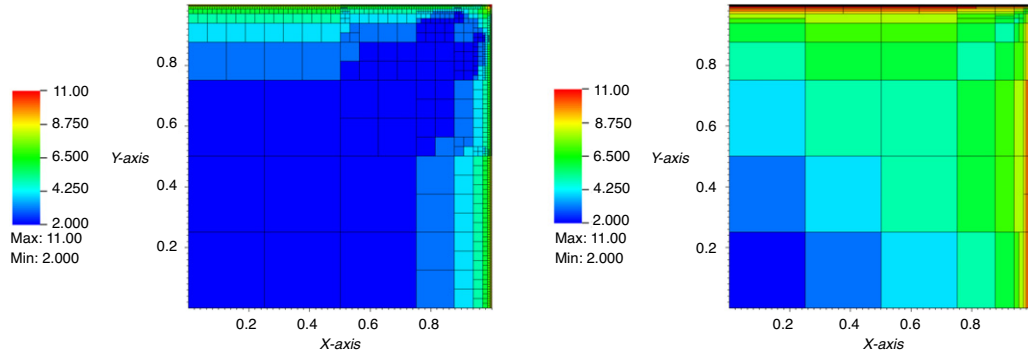


Fig. 7. Example 1 with $\varepsilon = 10^{-3}$: final adapted mesh with isotropic refinement (left) and anisotropic refinement (right).

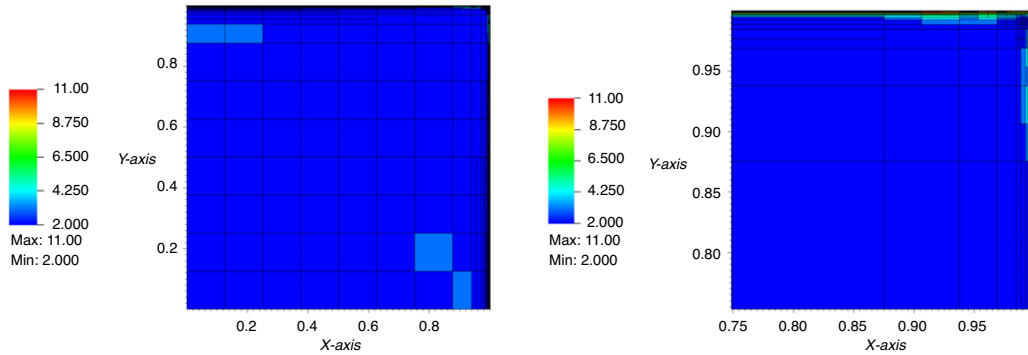


Fig. 8. Example 1 with $\varepsilon = 10^{-4}$: final adapted mesh with anisotropic refinement (left) and zoom into $(0.75, 1) \times (0.75, 1)$ (right).

5.2. Example 2

Next, we consider an example with an internal layer. In the domain $\Omega = (-1, 1)^2$, we take $\underline{a} = (1, 1)^\top$. We choose f and the inhomogeneous Dirichlet boundary conditions such that the solution to (1) is given by

$$u(x_1, x_2) = \arctan\left(\frac{x_1}{\varepsilon}\right)(1 - x_2^2).$$

For small values of ε , the solution u has an internal layer at $x_1 = 0$.

The estimator η can be readily extended to take into account the inhomogeneous boundary conditions. We run this problem for $\varepsilon = 10^{-3}$ and $\varepsilon = 10^{-4}$. For $\varepsilon = 10^{-3}$, we begin the test with a uniform mesh of 4×4 and the uniform polynomial degree $p_K = 2$, and for $\varepsilon = 10^{-4}$, with a 16×16 mesh and $p_K = 2$.

We present the same plots as in Example 1. In Fig. 9, we show the various error quantities. Again, we roughly see straight lines in a semi-logarithmic plot, indicating exponential convergence in $N^{1/2}$, although for $\varepsilon = 10^{-4}$ the convergence

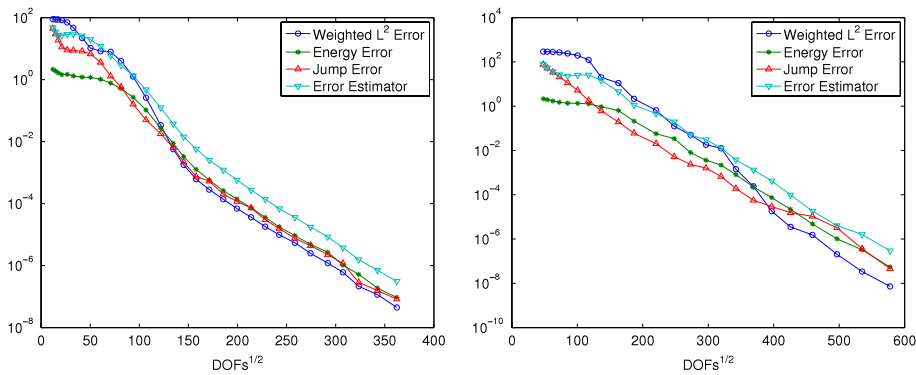


Fig. 9. Example 2: error plots for $\varepsilon = 10^{-3}$ (left) and $\varepsilon = 10^{-4}$ (right).

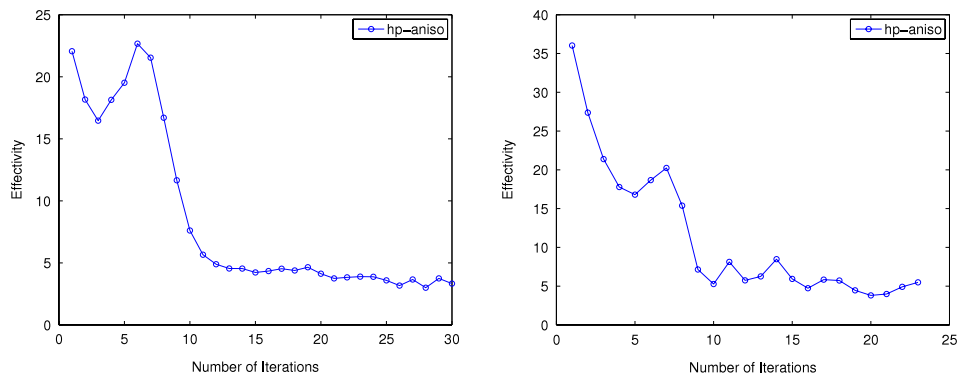


Fig. 10. Example 2: effectivity indices for $\varepsilon = 10^{-3}$ (left) and $\varepsilon = 10^{-4}$ (right).

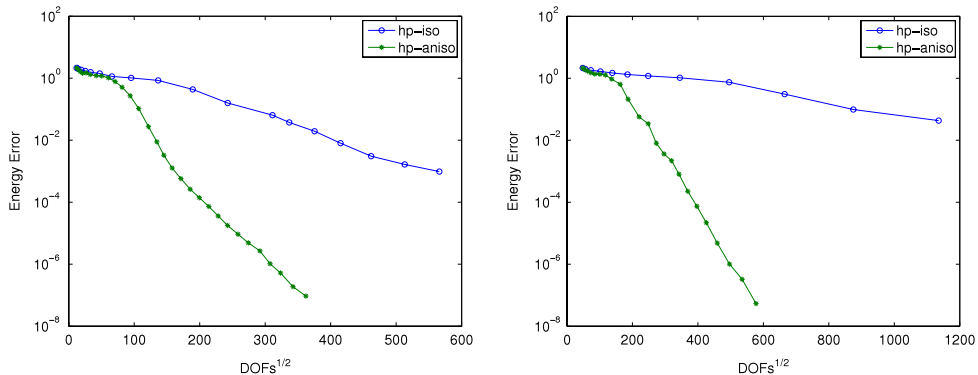


Fig. 11. Example 2: comparison of the DG energy errors for isotropic and anisotropic refinement for $\varepsilon = 10^{-3}$ (left) and $\varepsilon = 10^{-4}$ (right).

behavior particularly for the jump errors is much less clean. We observe that the estimator is overestimating the energy and jump errors, but not the weighted L^2 -norm errors. This is not a contradiction to our theoretical results, since the $\varepsilon^{-1/2}$ -weighted L^2 -norm provides only an upper bound of the convective semi-norm of the errors; cf. Remark 5. The effectivity indices are depicted in Fig. 10. Again, they start out large, but eventually converge to a reasonable value of around 5.

In Figs. 11 and 12, we show the energy norm and L^2 -norm errors for both the isotropic and anisotropic hp -algorithms. We can draw essential the same conclusions as in Example 1. The anisotropic version is clearly superior to the isotropic one; this is again more pronounced for the smaller value of $\varepsilon = 10^{-4}$. In this case, the isotropic L^2 -error curve reaches a value of around 10^{-2} with over a million degrees of freedom, while the anisotropic plot decreases to the level of 10^{-10} using less than 360,000 degrees of freedom. Fig. 13 shows the final adapted meshes for isotropic and anisotropic adaptivity for $\varepsilon = 10^{-3}$. Again, the final meshes are quite different, with the anisotropic one more effectively adapted to the layers.

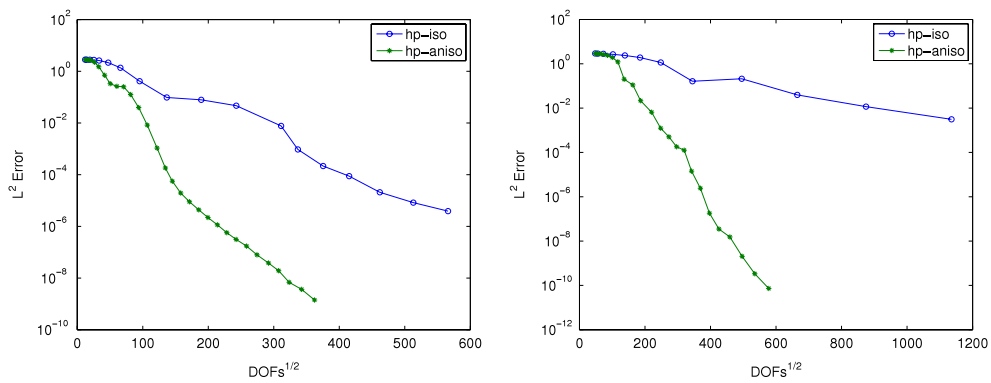


Fig. 12. Example 2: comparison of the L^2 -errors for isotropic and anisotropic refinement for $\varepsilon = 10^{-3}$ (left) and $\varepsilon = 10^{-4}$ (right).

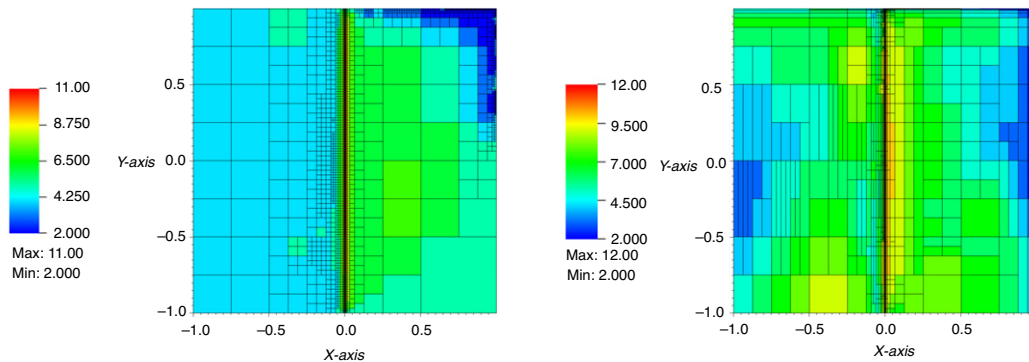


Fig. 13. Example 2 with $\varepsilon = 10^{-3}$: final adapted meshes with isotropic refinement (left) and anisotropic refinement (right).

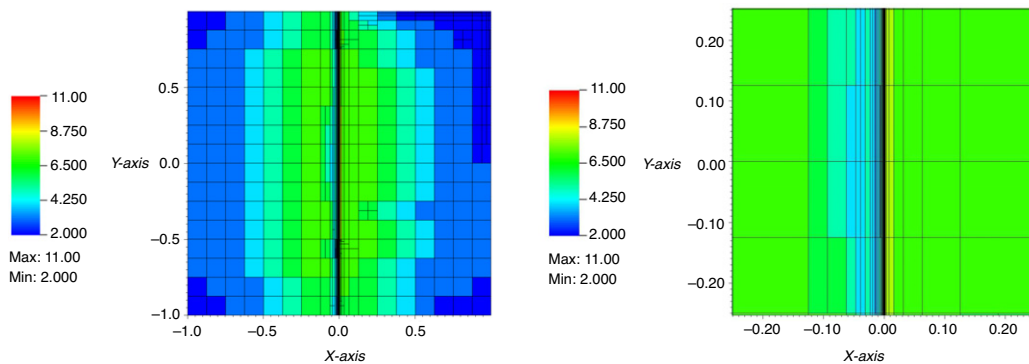


Fig. 14. Example 2 with $\varepsilon = 10^{-4}$: final adapted mesh with anisotropic refinement (left) and zoom into $(-0.25, 0.25) \times (-0.25, 0.25)$ (right).

Finally, in Fig. 14 we show the final adapted anisotropic mesh for $\varepsilon = 10^{-4}$, as well as a zoom into the central region $(-0.25, 0.25) \times (-0.25, 0.25)$ to better visualize the anisotropically refined elements.

5.3. Example 3

Next, we consider a problem, where the wind is not aligned with the mesh. We take $\Omega = (-1, 1)^2$, $\underline{a} = (-\sin \frac{\pi}{6}, \cos \frac{\pi}{6})^\top$, $f = 0$ and consider the boundary conditions $u = 0$ on $x_1 = -1$ and $x_2 = 1$, as well as

$$u = \tanh\left(\frac{1-x_2}{\varepsilon}\right) \quad \text{on } x_1 = 1, \quad u = \frac{1}{2}\left(\tanh\left(\frac{x_1}{\varepsilon}\right) + 1\right) \quad \text{on } x_2 = -1.$$

The boundary data is almost discontinuous near the point $(0, -1)$, and causes u to have an internal layer of width $\mathcal{O}(\sqrt{\varepsilon})$ along the line $x_2 + \sqrt{3}x_1 = -1$, with values $u = 0$ to the left and $u = 1$ to the right, as well as a boundary layer along the

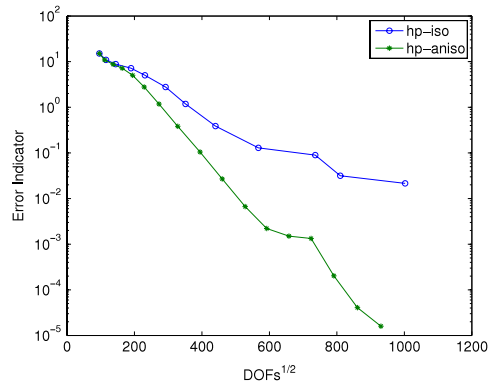


Fig. 15. Example 3 with $\varepsilon = 2.5 \times 10^{-4}$: convergence of error estimators for isotropic and anisotropic refinement.

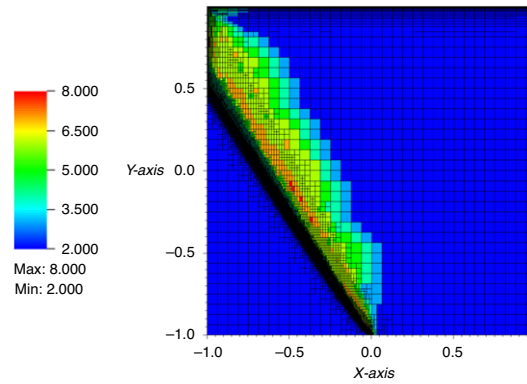


Fig. 16. Example 3 with $\varepsilon = 2.5 \times 10^{-4}$: final adapted mesh with anisotropic refinement.

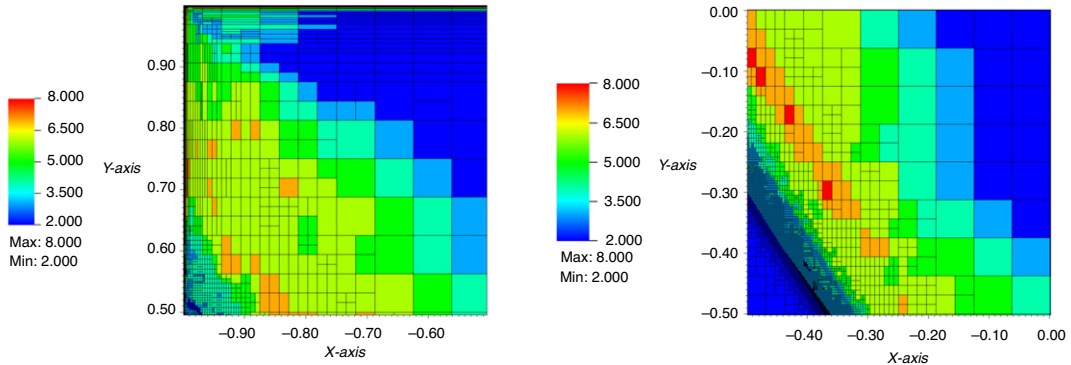


Fig. 17. Example 3 with $\varepsilon = 2.5 \times 10^{-4}$: zooms into the upper-left corner $(-1, -0.5) \times (0.5, 1)$ (left) and the central region $(-0.5, 0) \times (-0.5, 0)$ (right).

outflow boundary. There is no exact solution available to this problem. We test this problem with $\varepsilon = 2.5 \times 10^{-4}$, and start the algorithm for $p_K = 2$ on a uniform mesh of 16×16 elements.

In Fig. 15, we plot the values of the error indicators η for the isotropic and anisotropic hp -methods. We observe exponential convergence for the indicators in both algorithms, with the curves being closer together than in the previous tests. The reason for this is that in this example, the internal layer is not aligned with a coordinate direction. Hence, it cannot be anisotropically captured by the Cartesian meshes generated by our anisotropic code. As a result, the internal layer is only resolved isotropically, while anisotropic refinement is employed in the outflow layer. This is clearly visible in Fig. 16, where we show the final anisotropically adapted mesh. In Fig. 17 we show magnifications of the upper-left corner $(-1, -0.5) \times (0.5, 1)$ and the central region $(-0.5, 0) \times (-0.5, 0)$ of the adapted mesh. We note that designing fully automated ways to properly align meshes is a crucial aspect of anisotropic hp -adaptivity, which we do not address in this paper.

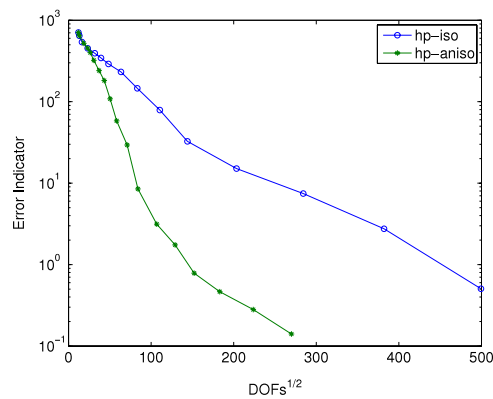


Fig. 18. Example 4 with $\varepsilon = 10^{-6}$: convergence of error estimators for isotropic and anisotropic refinements.

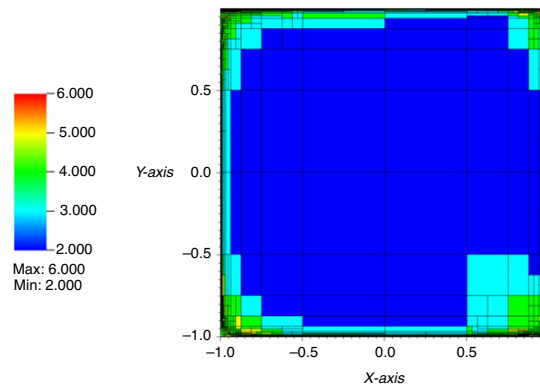


Fig. 19. Example 4 with $\varepsilon = 10^{-6}$: final adapted mesh with anisotropic refinement.

5.4. Example 4

Finally, we test our algorithm for an example with variable convection. In the square $\Omega = (-1, 1)^2$, we take the recirculating flow field $\underline{a} = (2y(1 - x^2), -2x(1 - y^2))^T$, set $f = 0$, and impose the inhomogeneous boundary conditions

$$u = \tanh\left(\frac{1 - x_2}{\varepsilon}\right) \quad \text{on } x_1 = -1, \quad u = \tanh\left(\frac{1 - x_1}{\varepsilon}\right) \quad \text{on } x_2 = -1,$$

and $u = 0$ on $x_1 = 1$ and $x_2 = 1$. In this problem, all the boundaries are characteristic, and the nearly discontinuous boundary conditions introduce boundary layers near them. Again, there is no exact solution available. We test the example with $\varepsilon = 10^{-6}$, and start our hp -algorithm with a uniform mesh of 16×16 elements and the uniform degree $p_K = 2$. Since the convection field \underline{a} is polynomial, we have evaluated the residual η_{R_K} exactly by using a Gauss quadrature rule of sufficiently high order on each element K . Hence, we have $\underline{a}_{hp} = \underline{a}$ in our computations.

In Fig. 18, we show the convergence of the estimators η obtained for the isotropic and anisotropic schemes. Both plots start with rather large values, and it takes over 10 adaptive iterations until the layers are reasonably well resolved. After 16 iterations, the estimated errors are still relatively large. Nevertheless, the anisotropic algorithm reaches an estimated error value of roughly 10^{-1} with less than 90,000 degrees of freedom, whereas the isotropic error value still is of order one with $N = 250,000$. Fig. 19 shows the final anisotropically adapted mesh. We observe strong anisotropic refinement along the layers, again with high polynomial degrees in the elements close to the boundaries. In the interior of the domain, our algorithm has selected biquadratic approximations on relatively large elements.

In Fig. 20 we show magnifications of the central left region $(-1, -0.5) \times (-0.5, 0)$ and the upper-left corner $(-1, -0.75) \times (0.75, 1)$ of the adapted mesh. In the first plot (left), anisotropic mesh refinement is strongly applied near the left boundary $x_1 = -1$. In the second plot (right), we observe a combination of isotropic and anisotropic elements on the left boundary, while anisotropic refinement is dominating again on the upper boundary $x_2 = 1$.

6. Conclusions

We have derived an a-posteriori error estimator for DG discretizations of convection–diffusion problems on anisotropically refined meshes. We have proved its reliability, up to an alignment measure which takes into account the

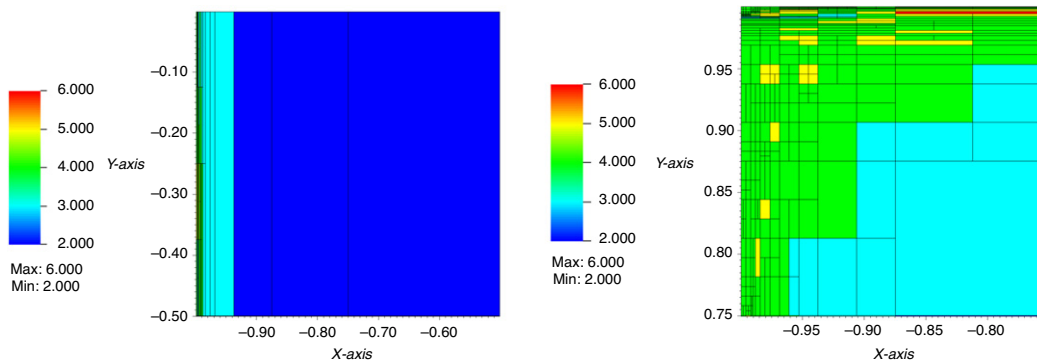


Fig. 20. Example 4 with $\varepsilon = 10^{-6}$: zooms into $(-1, -0.5) \times (-0.5, 0)$ (left) and $(-1, -0.75) \times (0.75, 1)$ (right).

possible anisotropy of the underlying meshes. The proof is based on the hp -version averaging operator of [7], appropriately scaled to anisotropic elements.

While the introduction of an alignment measure may not be completely satisfactory from a theoretical point of view, our numerical experiments indicate that it becomes of moderate size as soon as boundary layers have been sufficiently resolved, and that in this regime the effectivity indices behave practically uniformly. Our tests further indicate that anisotropic hp -adaptive DG algorithms are superior to isotropic ones by orders of magnitude, provided that the layers are properly aligned with the meshes.

Let us also mention a number of possible extensions of our work. An important item is the use of anisotropic polynomial degrees in the adaptive algorithms, which may be desirable to resolve boundary layers most effectively. Since our DG method is based on tensor-product polynomial spaces with respect to master element coordinates, anisotropic polynomial degrees can be incorporated in the numerical scheme with only minor modifications. Regarding the theoretical analysis, one of the key difficulties will be the construction and analysis of a suitable averaging operator in this setting. This will be addressed elsewhere.

Another valuable direction for future research is the extension of our analysis to non-affinely mapped quadrilaterals. For example, this seems possible for the class of anisotropic boundary-layer meshes introduced in [29, Section 3.2], where the elemental mappings F_K are factored into $F_K = \tilde{F}_K \circ G_K$. The mapping G_K is a combination of a dilation and a translation which maps the reference element \hat{K} into an anisotropic rectangle of the form $(0, h_x) \times (0, h_y)$, while \tilde{F}_K is a smooth mapping whose derivatives are uniformly bounded.

Finally, the generalization of our approach to three-dimensional problems is possible by anisotropically scaling the approximation estimates for the three-dimensional hp -averaging operator constructed in [6]. This is the subject of the ongoing research.

Acknowledgments

The first author acknowledges the financial support of the EPSRC under grant EP/H005498. The second author was supported in part by the National Sciences and Engineering Council of Canada (NSERC).

References

- [1] P. Houston, D. Schötzau, T.P. Wihler, Energy norm a-posteriori error estimation of hp -adaptive discontinuous Galerkin methods for elliptic problems, *Math. Models Methods Appl. Sci.* 17 (2007) 33–62.
- [2] O.A. Karakashian, F. Pascal, A-posteriori error estimates for a discontinuous Galerkin approximation of second-order elliptic problems, *SIAM J. Numer. Anal.* 41 (2003) 2374–2399.
- [3] E. Burman, A. Ern, Continuous interior penalty hp -finite element methods for advection and advection–diffusion equations, *Math. Comp.* 76 (2007) 1119–1140.
- [4] P. Houston, D. Schötzau, T. Wihler, An hp -adaptive discontinuous Galerkin FEM for nearly incompressible linear elasticity, *Comput. Methods Appl. Mech. Engrg.* 195 (2006) 3224–3246.
- [5] P. Houston, E. Süli, T. Wihler, A-posteriori error analysis of hp -version discontinuous Galerkin finite-element methods for second-order quasi-linear elliptic PDEs, *IMA J. Numer. Anal.* 28 (2008) 245–273.
- [6] L. Zhu, S. Giani, P. Houston, D. Schötzau, Energy norm a-posteriori error estimation for hp -adaptive discontinuous Galerkin methods for elliptic problems in three dimensions, *Math. Models Methods Appl. Sci.* 21 (2011) 267–306.
- [7] L. Zhu, D. Schötzau, A robust a-posteriori error estimate for hp -adaptive DG methods for convection–diffusion equations, *IMA J. Numer. Anal.* (2011) 971–1005.
- [8] R. Verfürth, Robust a-posteriori error estimates for stationary convection–diffusion equations, *SIAM J. Numer. Anal.* 43 (2005) 1766–1782.
- [9] E. Creusé, S. Nicaise, Anisotropic a-posteriori error estimation for the mixed discontinuous Galerkin approximation of the Stokes problem, *Numer. Methods Partial Differential Equations* 22 (2006) 449–483.
- [10] G. Kunert, An a-posteriori residual error estimator for the finite element method on anisotropic tetrahedral meshes, *Numer. Math.* 86 (2000) 471–490.
- [11] G. Kunert, A local problem error estimator for anisotropic tetrahedral finite element meshes, *SIAM J. Numer. Anal.* 39 (2001) 668–689.
- [12] E. Georgoulis, E. Hall, P. Houston, Discontinuous Galerkin methods on hp -anisotropic meshes I: a priori error analysis, *Int. J. Comput. Sci. Math.* 1 (2007) 221–244.

- [13] E. Georgoulis, E. Hall, P. Houston, Discontinuous Galerkin methods on hp -anisotropic meshes II: a-posteriori error analysis and adaptivity, *Appl. Numer. Math.* 59 (2009) 2179–2194.
- [14] E. Georgoulis, E. Hall, P. Houston, Discontinuous Galerkin methods for advection–diffusion–reaction problems on anisotropically refined meshes, *SIAM J. Sci. Comput.* 30 (2007/08) 246–271.
- [15] D. Schötzau, C. Schwab, A. Toselli, Mixed hp -DGFEM for incompressible flows II: geometric edge meshes, *IMA J. Numer. Anal.* 24 (2004) 273–308.
- [16] J. Melenk, B. Wohlmuth, On residual-based a-posteriori error estimation in hp -FEM, *Adv. Comput. Math.* 15 (2001) 311–331.
- [17] L. Zhu, Robust a-posteriori error estimation for discontinuous Galerkin methods for convection–diffusion problems, Ph.D. Thesis, The University of British Columbia, 2010. Electronic version downloadable at: <https://circle.ubc.ca>.
- [18] D. Schötzau, L. Zhu, A robust a-posteriori error estimator for discontinuous Galerkin methods for convection–diffusion equations, *Appl. Numer. Math.* 59 (2009) 2236–2255.
- [19] S. Giani, E. Hall, P. Houston, AptoFEM: users manual, Technical Report, University of Nottingham, 2009.
- [20] P.R. Amestoy, I.S. Duff, J. Koster, J.-Y. L'Excellent, A fully asynchronous multifrontal solver using distributed dynamic scheduling, *SIAM J. Matrix Anal. Appl.* 23 (2001) 15–41.
- [21] P.R. Amestoy, I.S. Duff, J.-Y. L'Excellent, Multifrontal parallel distributed symmetric and unsymmetric solvers, *Comput. Methods Appl. Mech. Engrg.* 184 (2000) 501–520.
- [22] P.R. Amestoy, A. Guermouche, J.-Y. L'Excellent, S. Pralet, Hybrid scheduling for the parallel solution of linear systems, *Parallel Comput.* 32 (2006) 136–156.
- [23] P. Houston, E. Süli, Adaptive finite element approximation of hyperbolic problems, in: T. Barth, H. Deconinck (Eds.), *Error Estimation and Adaptive Discretization Methods in Computational Fluid Dynamics*, in: *Lect. Notes Comput. Sci. Engrg.*, vol. 25, Springer-Verlag, 2002, pp. 269–344.
- [24] P. Houston, E. Süli, A note on the design of hp -adaptive finite element methods for elliptic partial differential equations, *Comput. Methods Appl. Mech. Engrg.* 194 (2005) 229–243.
- [25] T. Eibner, J.M. Melenk, An adaptive strategy for hp -FEM based on testing for analyticity, *Comput. Mech.* 39 (2007) 575–595.
- [26] P. Houston, B. Senior, E. Süli, Sobolev regularity estimation for hp -adaptive finite element methods, in: F. Brezzi, A. Buffa, S. Corsaro, A. Murli (Eds.), *Numerical Mathematics and Advanced Applications ENUMATH 2001*, Springer, Heidelberg, 2003, pp. 631–656.
- [27] C. Schwab, M. Suri, The p and hp versions of the finite element method for problems with boundary layers, *Math. Comp.* 65 (1996) 1403–1429.
- [28] C. Schwab, p - and hp -Finite Element Methods: Theory and Application to Solid and Fluid Mechanics, Oxford University Press, 1998.
- [29] K. Gerdes, J. Melenk, D. Schötzau, C. Schwab, The hp -version of the streamline–diffusion finite element method in two space dimensions, *Math. Models Methods Appl. Sci.* 11 (2001) 301–337.