hp-adaptive discontinuous Galerkin methods for non-stationary convection-diffusion problems

A. Cangiani[a], E. H. Georgoulis[a], S. Giani[c], S. Metcalfe[b]

[a] *Department of Mathematics, University of Leicester, Leicester LE1 7RH*
[b] *Department of Chemistry, University of York, York, YO10 5DD*
[c] *Department of Engineering, Durham University*
Lower Mountjoy, South Road, Durham, DH1 3LE, UK.
T: +44 (0) 191 334 2397, F: +44 (0) 191 334 2408, E: stefano.giani@durham.ac.uk

Abstract

An a posteriori error estimator for the error in the $(L^2(H^1)+L^\infty(L^2))$-type norm for an interior penalty discontinuous Galerkin (dG) spatial discretisation and backward Euler temporal discretisation of linear non-stationary convection-diffusion initial/boundary value problems is derived, allowing for anisotropic elements. The proposed error estimator is used to drive an *hp*-space-time adaptive algorithm wherein directional mesh refinement is employed to give rise to highly anisotropic elements able to accurately capture layers. The performance of the *hp*-space-time adaptive algorithm is assessed via a number of standard test problems characterised by sharp and/or moving layers.

Keywords:
Discontinuous Galerkin, unsteady convection-diffusion, a posteriori error estimation, adaptive finite element methods, anisotropic meshes.

1. Introduction

Convection-dominated convection-diffusion type problems often admit solutions involving steep gradients commonly referred to as boundary or interior layers. The accurate and efficient numerical resolution of such steep layers is a classical challenge in adaptive finite element methods as their exact location and strength may not, in general, be known a priori. For the special cases where the location of boundary or interior layers is known, structured grids can be successfully employed [30]. For non-stationary convection-diffusion equations, the situation is more challenging as the layers' location and/or strength may change over time, necessitating the use of adaptive algorithms to resolve the layers with an acceptable computational cost. Adaptive algorithms are usually based on suitable a posteriori error estimators or indicators, which range from ad hoc quantities to mathematically rigorous error bounds.

Rigorous a posteriori error estimation for stationary linear equations is now relatively well understood: for pure diffusion problems, we refer to the textbooks [36, 1] and for dG methods in particular to the papers [23, 8, 20]. For stationary convection-diffusion equations, the quest for robust a posteriori error estimators has seen recent advancements in various contexts [37, 39, 25, 31, 32, 33]. A posteriori error estimators for non-stationary linear convection-diffusion equations are also available for various discretisations [21, 9, 38, 14, 5, 6, 35, 17, 27, 11].

A posteriori error estimators for (spatial) dG methods for non-stationary parabolic problems can be found in [18, 15, 34], while convection-diffusion problems are considered in [12].

This work is concerned with the extension of the $L^2(H^1)+L^\infty(L^2)$-type a posteriori error estimators from [11] norm to include *hp*-version sensitivity and anisotropic mesh capabilities. We analyse a fully discrete method based on backward Euler time stepping coupled with the interior penalty dG discretisation in space. The derivation of the a posteriori bound is based on the elliptic reconstruction technique [28, 18] which allows the use of the robust elliptic error estimator from [32, 19].

The a posteriori estimator is used to drive an adaptive algorithm with *hp*-version and anisotropic refinement capabilities, thereby generalising the respective adaptive algorithm from [11]. First, the adaptive algorithm follows the recipe proposed in [11] to drive temporal and spatial refinement and coarsening. Then, for each element marked for refinement, it decides whether to apply h or p adaptivity. To this end, it follows the *hp*-adaptive strategy developed [22] based on estimating the local smoothness of the (unknown) analytical solution from truncated local Legendre expansions of the computed numerical solution. Finally, when h refinement is selected, a choice on whether to perform isotropic or anisotropic h refinement is made. This is based on exploiting the anisotropic nature of the novel a posteriori error estimator as detailed in Section 6.

We show numerically that *hp*-adaptivity results to exponential convergence rates and that anisotropic adaptivity is able to allow for the onset of exponential convergence under less numerical degrees of freedom than respective standard (isotropic) refinement procedures.

The following standard notation will be used throughout the reminder of the paper. For $T > 0$ and X a real Banach space with norm $\|\cdot\|_X$, we define the spaces $L^p(0,T;X)$, for $1 \leq p \leq +\infty$, that consist of all measurable functions $v : [0,T] \to X$ for which:

$$\begin{aligned} \|v\|_{L^p(0,T;X)} &:= \Big(\int_0^T \|v(t)\|_X^p dt \Big)^{1/p} < +\infty, \quad \text{for} \quad 1 \leq p < +\infty, \\ \|v\|_{L^\infty(0,T;X)} &:= \operatorname*{ess\,sup}_{0 \leq t \leq T} \|v(t)\|_X < +\infty, \quad \text{for} \quad p = +\infty. \end{aligned} \tag{1}$$

We also define $H^1(0,T,X) := \{u \in L^2(0,T;X) : u_t \in L^2(0,T;X)\}$. Finally, we denote by $C(0,T;X)$ and $C^{0,1}(0,T;X)$, respectively, the spaces of continuous and Lipschitz-continuous functions $v : [0,T] \to X$ such that:

$$\begin{aligned} ||v||_{C(0,T;X)} &:= \max_{0 \leq t \leq T} ||v(t)||_X < \infty, \\ ||v||_{C^{0,1}(0,T;X)} &:= \max \Big\{ ||v||_{C(0,T;X)}, ||\frac{\partial v}{\partial t}||_{C(0,T,X)} \Big\} < \infty. \end{aligned} \tag{2}$$

The symbols $\lesssim$ and $\gtrsim$ will be used to denote inequalities that are true up to a positive constant that is independent of the data (including the diffusion coefficient ε), cf. below and the continuous and discrete solutions.

Additionally, as they will be used frequently throughout this work, we will denote the L^2-inner product on the space domain Ω by $(\cdot,\cdot)$ and the L^2-norm on Ω by $||\cdot||$.

2. Model problem

Let $\Omega \subset \mathbb{R}^2$ be a bounded Lipschitz polygon with boundary $\partial\Omega$. We consider the model problem of finding $u : \Omega \times [0,T] \to \mathbb{R}$ such that

$$\begin{aligned} \frac{\partial u}{\partial t} - \varepsilon \Delta u + \mathbf{a} \cdot \nabla u + bu &= f \qquad \text{in } \Omega \times (0,T], \\ u &= 0 \qquad \text{on } \partial\Omega \times (0,T], \\ u(\cdot,0) &= u_0 \qquad \text{in } \bar{\Omega}, \end{aligned} \tag{3}$$

under the following assumptions: $0 < \varepsilon \leq 1$, $f \in C(0,T;L^2(\Omega))$, $\mathbf{a} \in C(0,T;W^{1,\infty}(\Omega))^2$, $b \in C(0,T;L^\infty(\Omega))$ and $u_0 \in L^2(\Omega)$. We further assume that

$$b - \frac{1}{2}\nabla \cdot \mathbf{a} \geq \beta \quad \text{a.e. in } \Omega \times [0,T], \qquad ||b - \nabla \cdot \mathbf{a}||_{C(0,T;L^\infty(\Omega))} \leq c_* \beta, \tag{4}$$

for some constants $\beta \geq 0$ and $c_* \geq 0$. For simplicity, we also assume that $\mathbf{a} \equiv \mathbf{a}(\mathbf{x},\mathbf{t})$ and Ω are of order one, possibly up to rescaling, so that ε^{-1} can be taken as the Péclet number of the problem.

The weak form of (3) then reads: find $u \in L^2(0,T;H_0^1(\Omega)) \cap H^1(0,T;L^2(\Omega))$ such that $u(\cdot,0) = u_0$ and for each $t \in (0,T]$ we have

$$\int_\Omega \frac{\partial u}{\partial t} v \, dx + \int_\Omega (\varepsilon \nabla u \cdot \nabla v + \mathbf{a} \cdot \nabla u v + buv) \, dx = \int_\Omega f v \, dx \qquad \forall v \in H_0^1(\Omega). \tag{5}$$

Note that under the regularity assumptions above, we have that $u \in C(0,T;L^2(\Omega))$.

3. Discontinuous Galerkin method

We consider a family of meshes $\zeta = \{K\}$ covering the computational domain Ω with K denoting a generic element. We assume that each subdivision ζ is constructed via affine mappings $F_K : \hat{K} \to K$ with non-singular Jacobian where $\hat{K}$ is the reference square, however, we do not assume that the subdivisions are shape-regular and hence each element K can be characterised by two length scales as detailed below. For simplicity, we also assume that the mesh is axiparallel. We denote by $\mathcal{E}(\zeta)$ the set of all edges in the triangulation ζ and by $\mathcal{E}^{int}(\zeta) \subset \mathcal{E}(\zeta)$ the subset of all interior edges. Each mesh is allowed to contain at most one hanging node per edge and so a mesh edge may coincide with either a part of or all of an edge of some $K \in \zeta$. This assumption is necessary because the error estimator presented in Section 5 is based on the analysis in [19] which exploits such assumption. DG methods have been extended to allow for arbitrarily small as well as an arbitrary number of edges, see eg. [10] and the references therein, but the proof of posteriori error bounds allowing for this generality is still an open problem. We further denote by $\mathcal{E}(K)$ the set of edges of K and note that as just discussed we may have $\mathcal{E}(K) \neq \{E \in \mathcal{E}(\zeta) : E \subset \partial K\}$. We further assume that whenever an edge $E \in \mathcal{E}(K)$ has a hanging node that this bisects E.

3.1. Mesh sizes

Each element $K \in \zeta$ can be characterised by two vectors $\underline{\mathrm{v}}_K^1$ and $\underline{\mathrm{v}}_K^2$ reflecting the two anisotropic directions such that $\underline{\mathrm{v}}_K^1 \perp \underline{\mathrm{v}}_K^2$. The lengths of these vectors are denoted by h_K^1 and h_K^2, respectively, and we set $h_{\min,K} := \min\{h_K^1, h_K^2\}$ and $h_{\max,K} := \max\{h_K^1, h_K^2\}$. We then define the matrix

$$\mathbf{M}_K = [\underline{\mathrm{v}}_K^1, \underline{\mathrm{v}}_K^2]. \tag{6}$$

Note that $\mathbf{M}_K$ is orthogonal and we have

$$\mathbf{M}_K^\top \mathbf{M}_K = \begin{bmatrix} (h_K^1)^2 & 0 \\ 0 & (h_K^2)^2 \end{bmatrix}.$$

Given an edge $E \in \mathcal{E}(\zeta)$, if $E \in \mathcal{E}(K)$ or E is a part of an elemental edge of K, we define a local function of the edge E, viz.,

$$\mathrm{h}_{E,K}^\perp = h_K^{3-i}, \qquad \text{if } E \text{ is parallel to } \underline{\mathrm{v}}_K^i,\ i = 1, 2,$$

Furthermore, if $E \in \mathcal{E}^{int}(\zeta)$ we assume that

$$\mathrm{h}_{E,K}^\perp \sim \mathrm{h}_{E,K'}^\perp, \qquad \text{if } E = \bar{K} \cap \bar{K}',\ K, K' \in \zeta. \tag{7}$$

Note that this assumption does not imply any restriction on the aspect ratio of the elements. Indeed, given any edges $E, E' \in \mathcal{E}(K)$ with $E \cap E' \neq \emptyset$, the ratio $h_E/h_{E'}$ is not constrained by the above where $h_E = h_K^i$ if E is parallel to $\underline{\mathrm{v}}_K^i$.

With this notation at hand, we define $h_E^\perp$ for $E \in \mathcal{E}(\zeta)$ as follows:

$$\mathrm{h}_E^\perp := \begin{cases} \min\{\mathrm{h}_{E,K}^\perp, \mathrm{h}_{E,K'}^\perp\}, & E \in \mathcal{E}^{int}(\zeta),\ E = \partial K \cap \partial K', \\ \mathrm{h}_{E,K}^\perp, & E \in \mathcal{E}(\zeta) \setminus \mathcal{E}^{int}(\zeta),\ E = \partial K \cap \partial\Omega. \end{cases}$$

Next, we define $h_{\min,E}$ for $E \in \mathcal{E}(\zeta)$ by setting

$$h_{\min,E} := \begin{cases} \min\{h_{\min,K}, h_{\min,K'}\}, & E \in \mathcal{E}^{int}(\zeta),\ E = \partial K \cap \partial K', \\ h_{\min,K}, & E \in \mathcal{E}(\zeta) \setminus \mathcal{E}^{int}(\zeta),\ E = \partial K \cap \partial\Omega, \end{cases}$$

Note that assumption (7) implies that given any edge $E \in \mathcal{E}(\zeta)$ and any element $K \in \zeta$, if $E \in \mathcal{E}(K)$ or E is a part of an elemental edge of K, we have

$$\mathrm{h}_E^\perp \sim \mathrm{h}_{E,K}^\perp, \qquad h_{\min,E} \sim h_{\min,K}. \tag{8}$$

Remark 3.1. *In the present setting of proof of context (i.e., that combination of hp-adaptivity with anisotropic mesh refinement can lead to computational savings), we opted for the simplifying assumption of axiparallel elements. We stress, however, that with minor modifications only the analysis presented below can be extended to valid distorted elements constructed via element mappings as described in detail in [16, 17].*

3.2. Polynomial degrees

We associate to each element $K \in \zeta$ a polynomial degree $p_K \geq 1$ and we introduce the degree vector $\mathbf{p} = \{ p_K : K \in \zeta \}$ and set $|\mathbf{p}| = \max_{K\in\zeta} p_K$. We assume that $\mathbf{p}$ is of bounded local variation, that is, that there exists a constant $\varrho \geq 1$ independent of the particular mesh in a sequence of meshes such that given any pair of neighboring elements $K, K' \in \zeta$ we have

$$\varrho^{-1} \leq p_K/p_{K'} \leq \varrho. \tag{9}$$

We can then introduce the edge polynomial degree p_E for $E \in \mathcal{E}(\zeta)$ given by

$$p_E = \begin{cases} \max\{p_K, p_{K'}\}, & E \in \mathcal{E}^{int}(\zeta),\ E = \partial K \cap \partial K', \\ p_K, & E \in \mathcal{E}(\zeta) \setminus \mathcal{E}^{int}(\zeta),\ E = \partial K \cap \partial\Omega. \end{cases} \tag{10}$$

For a given partition ζ of Ω and a given degree vector $\mathbf{p}$ on ζ, we define the *hp*-version discontinuous Galerkin finite element space by

$$V_{\mathbf{p}}(\zeta) = \{ v \in L^2(\Omega) : v|_K \circ F_K \in \mathcal{Q}_{p_K}(\widehat{K}),\ K \in \zeta \}, \tag{11}$$

with $\mathcal{Q}_{p_K}(\widehat{K})$ denoting the set of all polynomials on the reference square $\widehat{K}$ of degree no more than p_K.

3.3. Bilinear forms

We first introduce some notation. The outward unit normal to the boundary ∂K of an element K is denoted by $\mathbf{n}_K$. Given an edge $E \in \mathcal{E}^{int}(\zeta)$ shared by two elements K and K', a vector field $\mathbf{v} \in [H^{1/2}(\Omega)]^2$ and a scalar field $v \in H^{1/2}(\Omega)$, we define jumps and averages of $\mathbf{v}$ and v across E by:

$$\begin{aligned} \{\mathbf{v}\} &= \frac{1}{2}(\mathbf{v}|_{\bar{K}} + \mathbf{v}|_{\bar{K}'}), & [\mathbf{v}] &= \mathbf{v}|_{\bar{K}} \cdot \mathbf{n}_K + \mathbf{v}|_{\bar{K}'} \cdot \mathbf{n}_{K'}, \\ \{v\} &= \frac{1}{2}(v|_{\bar{K}} + v|_{\bar{K}'}), & [v] &= v|_{\bar{K}}\mathbf{n}_K + v|_{\bar{K}'}\mathbf{n}_{K'}. \end{aligned} \tag{12}$$

If $E \subset \partial\Omega$, we set $\{\mathbf{v}\} = \mathbf{v}$, $[\mathbf{v}] = \mathbf{v} \cdot \mathbf{n}$, $\{v\} = v$ and $[v] = v\mathbf{n}$, with $\mathbf{n}$ denoting the outward unit normal to the boundary $\partial\Omega$.

We define the outflow part of the boundary $\partial\Omega$ at the time t by

$$\partial\Omega^t_{out} = \{x \in \partial\Omega : \mathbf{a}(x,t) \cdot \mathbf{n}(x) \geq 0\}, \tag{13}$$

Similarly, we define the outflow parts of the boundary of an element K at time t by

$$\partial K^t_{out} = \{x \in \partial K : \mathbf{a}(x,t) \cdot \mathbf{n}_K(x) \geq 0\}. \tag{14}$$

We consider a full discretisation of problem (5) based on backward Euler time stepping and a variable mesh *hp*-dG discretisation in space. To this end, we introduce a subdivision of $[0, T]$ into n time steps of size $\tau_1, \tau_2, ..., \tau_n$ such that $\sum_{j=1}^n \tau_j = T$ for some $n \geq 1$ and set $t^0 = 0$ and $t^k := \sum_{j=1}^k \tau_j$. We denote an initial triangulation by ζ^0 and associate to each time step $k > 0$ a triangulation ζ^k which is assumed to have been obtained from ζ^{k-1} by locally refining and coarsening of ζ^{k-1}. This assumption allows us to avoid the well known problem of degradation of the finite element solution related to mesh change, *cf.* [13, 7]. To each mesh ζ^k, we associate a polynomial degree vector $\mathbf{p}$ and the finite element space $V^k_{\mathbf{p}} = V_{\mathbf{p}}(\zeta^k)$ given by (11). We also set $f(., t^k) = f^k$, $\mathbf{a}(., t^k) = \mathbf{a}^k$, and $b(., t^k) = b^k$ for brevity.

The fully-discrete dG method then reads as follows. Set u^0_h to be a projection of u_0 onto $V^0_{\mathbf{p}}$, e.g., the orthogonal L^2-projection; then, for $k = 0$, ..., $n-1$, find $u^{k+1}_h \in V^{k+1}_{\mathbf{p}}$, such that

$$(\frac{u^{k+1}_h - u^k_h}{\tau_{k+1}}, v^{k+1}_h) + B(t^{k+1}; u^{k+1}_h, v^{k+1}_h) + K_h(u^{k+1}_h, v^{k+1}_h) = (f^{k+1}, v^{k+1}_h) \qquad \forall v^{k+1}_h \in V^{k+1}_{\mathbf{p}}, \tag{15}$$

with

$$B(t;w,v) := \sum_{K\in\zeta}\int_K (\varepsilon\nabla w - \mathbf{a}w)\cdot\nabla v + (b-\nabla\cdot\mathbf{a})wv\,dx + \sum_{E\in\mathcal{E}(\zeta)}\frac{\gamma\varepsilon p_E^2}{h_E^\perp}\int_E [w]\cdot[v]\,ds$$
$$+ \sum_{K\in\zeta}\Big(\int_{\partial K_{out}^t\cap\partial\Omega_{out}^t}\mathbf{a}\cdot\mathbf{n}_K wv\,ds + \int_{\partial K_{out}^t\setminus\partial\Omega}\mathbf{a}\cdot\mathbf{n}_K w(v|_{\bar K} - v|_{\bar K'})\,ds\Big), \tag{16}$$
$$K_h(w,v) := -\sum_{E\in\mathcal{E}(\zeta)}\int_E \{\varepsilon\nabla w\}\cdot[v] + \{\varepsilon\nabla v\}\cdot[w]\,ds,$$

where $\bar K\cap\bar K' = E\subset\partial K_{out}^t\setminus\partial\Omega$. Here, $\gamma>0$ denotes the discontinuity penalty parameter which must be chosen large enough so that the operator $B+K_h$ is coercive, cf. Section 4 below. For simplicity, we assume $\gamma>1$ so that subsequent constants are independent of it.

Note that the bilinear form B is well-defined for $u,v\in H_0^1(\Omega)$ and $t\in(0,T]$. In particular, we have

$$B(t;u,v) = \int_\Omega(\varepsilon\nabla u\cdot\nabla v + \mathbf{a}\cdot\nabla uv + buv)\,dx. \tag{17}$$

Hence, the weak problem (5) may be rewritten as: find $u\in L^2(0,T;H_0^1(\Omega))\cap H^1(0,T;L^2(\Omega))$ such that $u(\cdot,0)=u_0$ and for each $t\in(0,T]$ we have

$$(\frac{\partial u}{\partial t},v) + B(t;u,v) = (f,v)\qquad \forall v\in H_0^1(\Omega). \tag{18}$$

4. Error bounds for the stationary problem

We introduce the following mesh-dependent quantities:

$$\begin{aligned}
\mathrm{ejump}_{\mathbf{p},\zeta}(\mathrm{u}) :=& \Big(\sum_{E\in\mathcal{E}(\zeta)}\frac{\gamma\varepsilon p_E^2}{h_E^\perp}|||[u]|||_{L^2(E)}^2\Big)^{1/2},\\
|||u|||_\zeta :=& \Big(\sum_{K\in\zeta}(\varepsilon||\nabla u||_{L^2(K)}^2 + \beta||u||_{L^2(K)}^2 + \mathrm{ejump}_{\mathbf{p},\zeta}(\mathrm{u})^2\Big)^{1/2},\\
\mathrm{ojump}_{\mathbf{p},\zeta}(\mathrm{u}) :=& \Big(\sum_{E\in\mathcal{E}(\zeta)}\Big(\frac{\varepsilon p_E^2 h_E^\perp}{h_{\min,\mathrm{E}}^2} + \frac{\beta h_E^\perp}{p_E^2} + \frac{h_E^\perp}{\varepsilon p_E}\Big)|||[u]|||_{L^2(E)}^2\Big)^{1/2}\\
|u|_{A,\zeta} :=& \Big(\Big(\sup_{v\in H_0^1(\Omega)\setminus\{0\}}\frac{\int_\Omega \mathbf{a}u\cdot\nabla v\,dx}{|||v|||_\zeta}\Big)^2 + \mathrm{ojump}_{\mathbf{p},\zeta}(\mathrm{u})^2\Big)^{1/2},
\end{aligned} \tag{19}$$

noting that $|||\cdot|||_\zeta$ and $|\cdot|_{A,\zeta}$ define norms on $H_0^1(\Omega)+V_{\mathbf{p}}(\zeta)$.

For any $t\in(0,T]$, the following properties holds in standard fashion; their proof is omitted for brevity. The bilinear form $B(t;\cdot,\cdot)$ is coercive on $H_0^1(\Omega)$, that is,

$$B(t;v,v)\geq |||v|||_\zeta^2 \tag{20}$$

for all $v\in H_0^1(\Omega)$. It is also continuous in the following sense

$$B(t;w,v)\lesssim (|||w|||_\zeta + |w|_{A,\zeta})|||v|||_\zeta \tag{21}$$

for all $w\in H_0^1(\Omega)+V_{\mathbf{p}}(\zeta)$ and $v\in H_0^1(\Omega)$. Finally, for γ large enough, the discrete bilinear form $B+K_h$ is coercive in $V_{\mathbf{p}}(\zeta)$ with respect to the $|||\cdot|||_\zeta$ norm, viz.,

$$B(t;v_h,v_h) + K_h(v_h,v_h)\gtrsim |||v_h|||_\zeta^2 \tag{22}$$

for all $v_h\in V_{\mathbf{p}}(\zeta)$.

Next, given $K \in \zeta$ and $E \subset \partial K$, we introduce the following notation

$$\begin{aligned}\alpha_K &:= \min(h_{\min,K}\varepsilon^{-\frac{1}{2}}p_K^{-1}, \beta^{-\frac{1}{2}}),\\ \alpha_E &:= \min(h_{\min,K}^2\varepsilon^{-\frac{1}{2}}p_E^{-1}(h_E^{\perp})^{-1}, \beta^{-\frac{1}{2}}),\\ \alpha_T &:= \min(\varepsilon^{-\frac{1}{2}}, \beta^{-\frac{1}{2}}).\end{aligned}$$

We then have the following result whose proof is completely analogous to that of [19, Theorem 3] and is therefore omitted for brevity, cf. also [32].

Theorem 4.1. *For a given $t \in (0,T]$, let $u^s \in H_0^1(\Omega)$ be such that*

$$B(t; u^s, v) = (f, v) \qquad \forall v \in H_0^1(\Omega),$$

and consider $u_h^s \in V_{\mathbf{p}}(\zeta)$ such that

$$B(t; u_h^s, v_h) + K_h(u_h^s, v_h) = (f, v_h) \qquad \forall v_h \in V_{\mathbf{p}}(\zeta).$$

Then, the following a posteriori bound holds:

$$\begin{aligned}&(|||u^s - u_h^s|||_\zeta + |u^s - u_h^s|_{A,\zeta})^2\\ &\lesssim \sum_{K\in\zeta} \alpha_K^2 ||f + \varepsilon\Delta u_h^s - \mathbf{a}\cdot\nabla u_h^s - b u_h^s||_{L^2(K)}^2\\ &\quad + \frac{1}{2}\sum_{K\in\zeta}\sum_{E\in\mathcal{E}^{int}(K)} \varepsilon^{\frac{3}{2}}\alpha_E |||[\nabla u_h^s]|||_{L^2(E)}^2\\ &\quad + \frac{1}{2}\sum_{K\in\zeta}\sum_{E\in\mathcal{E}(K)} \Big(\frac{\gamma^2\varepsilon p_E^3}{h_{E,K}^{\perp}} + \frac{\varepsilon h_{E,K}^{\perp} p_E^2}{h_{\min,\mathrm{K}}^2} + \frac{\beta h_{E,K}^{\perp}}{p_E^2} + \frac{h_{E,K}^{\perp}}{\varepsilon p_E}\Big) |||[u_h^s]|||_{L^2(E\setminus\partial\Omega)}^2\\ &\quad + \sum_{K\in\zeta}\sum_{E\in\mathcal{E}(K)} \Big(\frac{\gamma^2\varepsilon p_E^3}{h_{E,K}^{\perp}} + \frac{\varepsilon h_{E,K}^{\perp} p_E^2}{h_{\min,\mathrm{K}}^2} + \frac{\beta h_{E,K}^{\perp}}{p_E^2} + \frac{h_{E,K}^{\perp}}{\varepsilon p_E}\Big) |||[u_h^s]|||_{L^2(E\cap\partial\Omega)}^2.\end{aligned}$$

□

5. An a posteriori error bound

For the proof of a posteriori error bounds we shall make use of the elliptic reconstruction framework [28, 26, 18].

Definition 5.1. *We define the* elliptic reconstruction $w^k \in H_0^1(\Omega)$ *to be the unique solution of the elliptic problem*

$$B(t^k; w^k, v) = (A^k, v) \qquad \forall v \in H_0^1(\Omega),$$

where $A^k \in V_{\mathbf{p}}^k$ is the Riesz representer satisfying

$$B(t^k; u_h^k, v_h^k) + K_h(u_h^k, v_h^k) = (A^k, v_h^k) \qquad \forall v_h^k \in V_{\mathbf{p}}^k.$$

Remark 5.2. *For $k \geq 1$, we obtain from the numerical method that*

$$A^{k+1} = I_h^{k+1} f^{k+1} - \frac{u_h^{k+1} - I_h^{k+1} u_h^k}{\tau_{k+1}},$$

where I_h^{k+1} is the L^2-projection operator onto $V_{\mathbf{p}}^{k+1}$.

For each time index k, we decompose the dG solution u_h^k into a conforming part $u_{h,c}^k \in H_0^1(\Omega) \cap V_{\mathbf{p}}^k$ and a non-conforming part $u_{h,d}^k \in V_{\mathbf{p}}^k$ such that $u_h^k = u_{h,c}^k + u_{h,d}^k$. Given $t \in (t^k, t^{k+1}]$, we define $u_h(t)$ to be the linear interpolant with respect to t of the values u_h^k and u_h^{k+1}, viz.,

$$u_h(t) := l_k(t)u_h^k + l_{k+1}(t)u_h^{k+1}, \tag{23}$$

with $\{l_k, l_{k+1}\}$ denoting the standard linear Lagrange interpolation basis defined on the interval $[t^k, t^{k+1}]$. We define $u_{h,c}(t)$ and $u_{h,d}(t)$ analogously. We can then decompose the error $e = u - u_h = e_c - u_{h,d}$ where $e_c = u - u_{h,c}$. It will also be useful to define $\theta^k = w^k - u_h^k$.

Lemma 5.3. *Given $t \in (t^k, t^{k+1}]$ we have*

$$(\frac{\partial e}{\partial t}, v) + B(t; e, v) = (f - f^{k+1}, v) + (f^{k+1} - \frac{\partial u_h}{\partial t}, v) - B(t; u_h, v) \qquad \forall v \in H_0^1(\Omega).$$

Proof. This follows from (18) by adding and subtracting various terms. □

We now introduce the error estimators. We begin by defining the *spatial estimator*

$$\begin{aligned}\eta_S^2 :=& ||e(0)||^2 + \frac{1}{3}\sum_{j=0}^{n-1}\tau_{j+1}(\eta_{S_1,j}^2 + \eta_{S_1,j+1}^2) + \sum_{j=0}^{n-1}\tau_{j+1}\eta_{S_2,j+1}^2 + \max_{0\le j\le n}\eta_{S_3,j}^2 \\ &+ \min\Big\{\Big(\sum_{j=0}^{n-1}\tau_{j+1}\eta_{S_4,j+1}\Big)^2, \alpha_T^2\sum_{j=0}^{n-1}\tau_{j+1}\eta_{S_4,j+1}^2\Big\},\end{aligned} \tag{24}$$

where

$$\begin{aligned}\eta_{S_1,j}^2 =& \sum_{K\in\zeta^j}\alpha_K^2||A^j + \varepsilon\Delta u_h^j - \mathbf{a}^j\cdot\nabla u_h^j - b^j u_h^j||_{L^2(K)}^2 + \sum_{K\in\zeta^j}\sum_{E\in\mathcal{E}^{int}(K)}\varepsilon^{\frac{3}{2}}\alpha_E|||[\nabla u_h^j]|||_{L^2(E)}^2 \\ &+ \sum_{K\in\zeta^j}\sum_{E\in\mathcal{E}(K)}\Big(\frac{\gamma^2\varepsilon p_E^3}{h_{E,K}^{\perp}} + \frac{\varepsilon h_{E,K}^{\perp}p_E^2}{h_{\min,\mathrm{K}}^2} + \frac{\beta h_{E,K}^{\perp}}{p_E^2} + \frac{h_{E,K}^{\perp}}{\varepsilon p_E}\Big)|||[u_h^j]|||_{L^2(E)}^2, \\ \eta_{S_2,j+1}^2 =& \sum_{K\in\zeta^j\cup\zeta^{j+1}}\alpha_K^2||f^{j+1} - I_h^{j+1}f^{j+1} + \frac{u_h^j - I_h^{j+1}u_h^j}{\tau_{j+1}}||_{L^2(K)}^2, \\ \eta_{S_3,j}^2 =& \sum_{K\in\zeta^j}\sum_{E\in\mathcal{E}(K)}\frac{h_{E,K}^{\perp}}{p_E^2}|||[u_h^j]|||_{L^2(E)}^2, \\ \eta_{S_4,j+1}^2 =& \sum_{K\in\zeta^j\cup\zeta^{j+1}}\sum_{E\in\mathcal{E}(K)}\frac{h_{E,K}^{\perp}}{p_E^2}|||[\frac{u_h^{j+1} - u_h^j}{\tau_{j+1}}]|||_{L^2(E)}^2.\end{aligned} \tag{25}$$

The *time (or temporal) estimator*, is given by

$$\eta_T^2 := \sum_{j=0}^{n-1}\int_{t^j}^{t^{j+1}}\eta_{T_1,j+1}^2\,dt + \min\Big\{\Big(\sum_{j=0}^{n-1}\int_{t^j}^{t^{j+1}}\eta_{T_2,j+1}\,dt\Big)^2, \alpha_T^2\sum_{j=0}^{n-1}\int_{t^j}^{t^{j+1}}\eta_{T_2,j+1}^2\,dt\Big\}, \tag{26}$$

where

$$\begin{aligned}\eta_{T_1,j+1}^2 =& \sum_{K\in\zeta^j\cup\zeta^{j+1}}\frac{1}{\varepsilon}||l_{j+1}(\mathbf{a}^{j+1} - \mathbf{a})u_h^{j+1} + l_j(\mathbf{a}^j - \mathbf{a})u_h^j||_{L^2(K)}^2, \\ \eta_{T_2,j+1}^2 =& \sum_{K\in\zeta^j\cup\zeta^{j+1}}||f - f^{j+1} + l_j(A^{j+1} - A^j) + l_{j+1}(b^{j+1} - b - \nabla\cdot\mathbf{a}^{j+1} + \nabla\cdot\mathbf{a})u_h^{j+1} \\ &+ l_j(b^j - b - \nabla\cdot\mathbf{a}^j + \nabla\cdot\mathbf{a})u_h^j||_{L^2(K)}^2.\end{aligned} \tag{27}$$

We are now ready to state our error bound for the $(L^2(H^1)+L^\infty(L^2))$-type norm

$$||e||_\star := \left(||e||^2_{L^\infty(0,T;L^2(\Omega))} + \int_0^T |||e|||^2 dt \right)^{1/2}, \tag{28}$$

where the norm $|||\cdot|||$ is taken over the mesh $\zeta^k \cup \zeta^{k+1}$ for $t \in (t^k, t^{k+1})$.

Theorem 5.4. *The error e of the fully-discrete method satisfies the bound*

$$||e||_\star \lesssim \sqrt{\eta_S^2 + \eta_T^2}. \tag{29}$$

Proof. From Lemma 5.3 and Definition 5.1 we have

$$\begin{aligned}(\frac{\partial e}{\partial t}, v) + B(t; e, v) =& (f - f^{k+1}, v) + (f^{k+1} - \frac{\partial u_h}{\partial t} - A^{k+1}, v) + B(t^{k+1}; \theta^{k+1}, v) \\ &+ B(t^{k+1}; u_h^{k+1}, v) - B(t; u_h, v),\end{aligned} \tag{30}$$

which upon straightforward manipulation and using Remark 5.2 gives

$$\begin{aligned}&(\frac{\partial e}{\partial t}, v) + B(t; e, v) \\ &= l_{k+1} B(t^{k+1}; \theta^{k+1}, v) + l_k B(t^k; \theta^k, v) + (f - f^{k+1} + l_k(A^{k+1} - A^k), v) \\ &\quad + l_{k+1} B(t^{k+1}; u_h^{k+1}, v) + l_k B(t^k; u_h^k, v) - B(t; u_h, v) + (f^{k+1} - \frac{\partial u_h}{\partial t} - A^{k+1}, v).\end{aligned} \tag{31}$$

The final term can be rewritten using Remark 5.2, properties and bounds of the L^2-projection and the Cauchy-Schwarz inequality yielding

$$(f^{k+1} - \frac{\partial u_h}{\partial t} - A^{k+1}, v) = (f^{k+1} - \frac{\partial u_h}{\partial t} - A^{k+1}, v - I_h^{k+1} v) \lesssim \eta_{S_2,k+1} |||v|||. \tag{32}$$

Using the definition of the bilinear form B, the Cauchy-Schwarz inequality, integration by parts and a standard hp-version inverse estimate [16, 17], viz.,

$$\|v\|^2_{L^2(E)} \lesssim \frac{p_E^2}{h_E^\perp} \|v\|^2_{L^2(K)}, \tag{33}$$

the remaining four terms give rise to the time estimator:

$$\begin{aligned}&(f - f^{k+1} + l_k(A^{k+1} - A^k), v) + l_{k+1} B(t^{k+1}; u_h^{k+1}, v) + l_k B(t^k; u_h^k, v) \\ &- B(t; u_h, v) \lesssim \eta_{T_1,k+1} |||v||| + \eta_{T_2,k+1} ||v||.\end{aligned} \tag{34}$$

Setting $v = e_c$ and using (20), (21), the Cauchy-Schwarz inequality and Young's inequality yields

$$\begin{aligned}\frac{d}{dt}(||e_c||^2) + |||e_c|||^2 \lesssim& ||\frac{\partial u_{h,d}}{\partial t}|| \, ||e_c|| + l_{k+1}^2 (|||\theta^{k+1}||| + |\theta^{k+1}|_A)^2 + l_k^2 (|||\theta^k||| + |\theta^k|_A)^2 \\ &+ |||u_{h,d}|||^2 + |u_{h,d}|_A^2 + \eta_{T_2,k+1} ||e_c|| + \eta_{T_1,k+1}^2 + \eta_{S_2,k+1}^2.\end{aligned} \tag{35}$$

The θ terms on the right-hand side of the last inequality give rise to the space estimator via elliptic reconstruction and Theorem 4.1, viz.,

$$(|||\theta^k||| + |\theta^k|_A)^2 \lesssim \eta_{S_1,k}^2, \qquad (|||\theta^{k+1}||| + |\theta^{k+1}|_A)^2 \lesssim \eta_{S_1,k+1}^2 \tag{36}$$

Finally, we recall the conforming-nonconforming stability bounds from [19] which are extensions to anisotropic elements of the respective bounds by Karakashian and Pascal [23, 24]:

$$\begin{aligned}|||u_{h,d}|||^2 + |u_{h,d}|_A^2 &\lesssim \sum_{K \in \zeta^k \cup \zeta^{k+1}} \sum_{E \in \mathcal{E}(K)} \Big(\frac{\gamma^2 \varepsilon p_E^3}{h_{E,K}^\perp} + \frac{\varepsilon h_{E,K}^\perp p_E^2}{h_{\min,K}^2} + \frac{\beta h_{E,K}^\perp}{p_E^2} + \frac{h_{E,K}^\perp}{\varepsilon p_E} \Big) |||[u_h]|||^2_{L^2(E)}, \\ ||\frac{\partial u_{h,d}}{\partial t}||^2 &\lesssim \sum_{K \in \zeta^k \cup \zeta^{k+1}} \sum_{E \in \mathcal{E}(K)} \frac{h_{E,K}^\perp}{p_E^2} |||[\frac{\partial u_h}{\partial t}]|||^2_{L^2(E)}, \\ ||u_{h,d}||^2 &\lesssim \sum_{K \in \zeta^k \cup \zeta^{k+1}} \sum_{E \in \mathcal{E}(K)} \frac{h_{E,K}^\perp}{p_E^2} |||[u_h]|||^2_{L^2(E)}.\end{aligned} \tag{37}$$

The proof then follows from (35), (36), (37) along with standard arguments [11]. □

6. An adaptive algorithm

The a posteriori error bounds presented above will be used to drive an *hp*-space-time adaptive algorithm. The algorithm we propose is an extension to the one presented in [11] allowing also for *hp*-adaptivity. In the algorithm in [11], the space-time adaptivity is regulated by four parameters: $\texttt{ttol}^+$, $\texttt{ttol}^-$, $\texttt{stol}^+$, $\texttt{stol}^-$. The quantities $\texttt{ttol}^+$ and $\texttt{ttol}^-$ are the thresholds for applying refinement and derefinement in time, respectively. Similarly, $\texttt{stol}^+$ and $\texttt{stol}^-$ are the thresholds for applying refinement and derefinement in space. The *hp*-space-time adaptive algorithm is based on using different parts of the a posteriori estimator from Theorem 5.4 to drive space-time adaptivity.

Algorithm 1 Algorithm to apply anisotropic adaptivity.

Input: $\zeta^j, \eta^2_{S_1,j}, u_h, \texttt{stol}^+, \texttt{stol}^-$.
Output: ζ^{j+1}.
for $\forall K \in \zeta^j$ **do**
  if $\eta^2_{S_1,j}|_K > \texttt{stol}^+$ **then**
    if $u_h|_K$ is smooth enough **then**
      Mark element K for refinement in p
    else
      $\eta^2_{S_1,E^1_K} := \eta^2_{S_1,j}|_{E^1_K}$
      $\eta^2_{S_1,E^2_K} := \eta^2_{S_1,j}|_{E^2_K}$
      if $\eta_{S_1,E^1_K} > 10\eta_{S_1,E^2_K}$ **then**
        Mark K for anisotropic h-refinement in the direction $\underline{\mathrm{v}}^1_K$
      else if $\eta_{S_1,E^2_K} > 10\eta_{S_1,E^1_K}$ **then**
        Mark K for anisotropic h-refinement in the direction $\underline{\mathrm{v}}^2_K$
      else
        Mark K for isotropic refinement in h
      end if
    end if
  else if $\eta^2_{S_1,j}|_K < \texttt{stol}^-$ **then**
    if $u_h|_K$ is smooth enough **then**
      Mark K for derefinement in p lowering the order by 1
    else
      Mark K for derefinement in h undoing the last h-refinement done to the element
    end if
  end if
end for
Apply h-smoothing to guarantee at most one hanging node per edge
Apply p-smoothing to guarantee (9)
Adapt the mesh ζ^j to create ζ^{j+1}

For simplicity, we use in the *hp*-space-time adaptive algorithm the quantity $\hat{\eta}_{T,j+1}$ given by

$$\hat{\eta}^2_{T,j+1} = \int_{t^j}^{t^{j+1}} \eta^2_{T_1,j+1}\,dt + \min\{\alpha_T, T\} \int_{t^j}^{t^{j+1}} \eta^2_{T_2,j+1}\,dt, \tag{38}$$

to drive the time refinement. The sum of all terms $\hat{\eta}^2_{T,j+1}$ can bound η^2_T from above and, hence, this can be used to drive temporal refinement and coarsening subject to two temporal thresholds $\texttt{ttol}^+$ and $\texttt{ttol}^-$ on each time interval. We refer the reader to [29, 11] for a discussion on the delicate matter of the relative choice of the various thresholds. Both mesh refinement and coarsening are driven by the term $\eta_{S_1,j}$. The size of the elemental contributions to $\eta_{S_1,j}$ determines whether the elements are to be

refined, coarsened or neither depending on two spatial thresholds $\texttt{stol}^+$ and $\texttt{stol}^-$, i.e. all elements with $\eta^2_{S_1,j}|_K > \texttt{stol}^+$ are marked for refinement and all elements with $\hat{\eta}^2_{S_1,j} < \texttt{stol}^-$ are marked for coarsening. When *hp*-adaptivity is used, the algorithm needs to decide for each element marked for refinement whether to apply either h or p adaptivity. The choice is based on estimating the local smoothness of the analytical solution. To this end, we employ the *hp*-adaptive strategy developed in [22] where the local regularity of the analytical solution is estimated from truncated local Legendre expansions of the computed numerical solution. In the case that anisotropic *hp*-adaptivity is used, the method also has to decide whether to perform isotropic or anisotropic h refinement. To make this decision, we denote by E^1_K and E^2_K the two sets containing the edges of K parallel to either $\underline{\mathrm{v}}^1_K$ or $\underline{\mathrm{v}}^2_K$ and we define

$$\eta^2_{S_1,E^i_K} = \sum_{E\in E^i_K} \varepsilon^{\frac{3}{2}}\alpha_E |||[\nabla u^j_h]|||^2_{L^2(E)} + \sum_{E\in E^i_K}\Big(\frac{\gamma^2\varepsilon p^3_E}{h^{\perp}_{E,K}} + \frac{\varepsilon h^{\perp}_{E,K}p^2_E}{h^2_{\min,\mathrm{K}}} + \frac{\beta h^{\perp}_{E,K}}{p^2_E} + \frac{h^{\perp}_{E,K}}{\varepsilon p_E}\Big)|||[u^j_h]|||^2_{L^2(E)},$$

$i = 1, 2$. The choice between isotropic or anisotropic h refinement is then made by comparing η_{S_1,E^1_K} and η_{S_1,E^2_K}: if $\eta_{S_1,E^1_K} > 10\eta_{S_1,E^2_K}$, then the element K is refined anisotropically along the direction $\underline{\mathrm{v}}^1_K$; on the other hand, if $\eta_{S_1,E^2_K} > 10\eta_{S_1,E^1_K}$, then the element K is refined along the direction $\underline{\mathrm{v}}^2_K$. If neither of these conditions is met then the element K is refined isotropically. The pseudocode of the full anisotropic adaptivity procedure is reported in Algorithm 1.

7. Numerical experiments

We shall investigate numerically the presented a posteriori bounds and the performance of the adaptive algorithm through an implementation based on the `AptoFEM` software package. The resulting discrete systems of linear equations are solved by exploiting the Multifrontal Massively Parallel Solver (MUMPS) [2, 3, 4]. All the numerical experiments have been performed using an Intel® Core i7 PC with 16 GB of RAM and 3.60 GHz. The codes only take the default optimization of the machine, i.e., they are not parallel. Since the basic adaptive algorithm that we use is already analysed in [11], this section focuses solely on the advantages of the *hp*-error estimator in space for non-stationary problems.

If we let λ_k denote the total number of degrees of freedom on the union mesh $\zeta^k \cup \zeta^{k+1}$ then the weighted degrees of freedom of the problem is given by

$$\text{Weighted Average DoFs} := \frac{1}{T}\sum_{j=0}^{n-1}\tau_{j+1}\lambda_j. \tag{39}$$

7.1. Anisotropic adaptivity

In the section we present numerical results highlighting the advantages of *hp*-adaptivity in terms of average DoFS for a model problem. In the sequel, we shall use the terminology isoh, isohp, anisoh and anisohisop to refer to the dG method with isotropic h-adaptivity, isotropic *hp*-adaptivity, anisotropic h-adaptivity and anisotropic h-adaptivity/isotropic p-adaptivity, respectively.

Let $\Omega = (0,1)^2$, $\mathbf{a} = (1,1)^T$, $b = 0$, $u_0 = 0$, $T = 10$ and select the function f so that the exact solution to problem (5) is given by

$$u(x,y,t) = (1-e^{-t})\bigg(\frac{e^{(x-1)/\varepsilon}-1}{e^{-1/\varepsilon}-1} + x - 1\bigg)\bigg(\frac{e^{(y-1)/\varepsilon}-1}{e^{-1/\varepsilon}-1} + y - 1\bigg). \tag{40}$$

The solution exhibits boundary layers at the outflow boundary of the domain of width $\mathcal{O}(\varepsilon)$ as well as a temporal boundary layer.

We begin by fixing the temporal threshold to $\texttt{ttol}^+ = 0.001$. This value is small enough to ensure that the temporal contribution to the error is very small in comparison to the spatial contribution. The spatial threshold is then gradually reduced in order to observe the spatial effectivity indices for this problem. The same test is run several times for different values of ε and using different adaptive techniques. The results are reported in Figure 1.

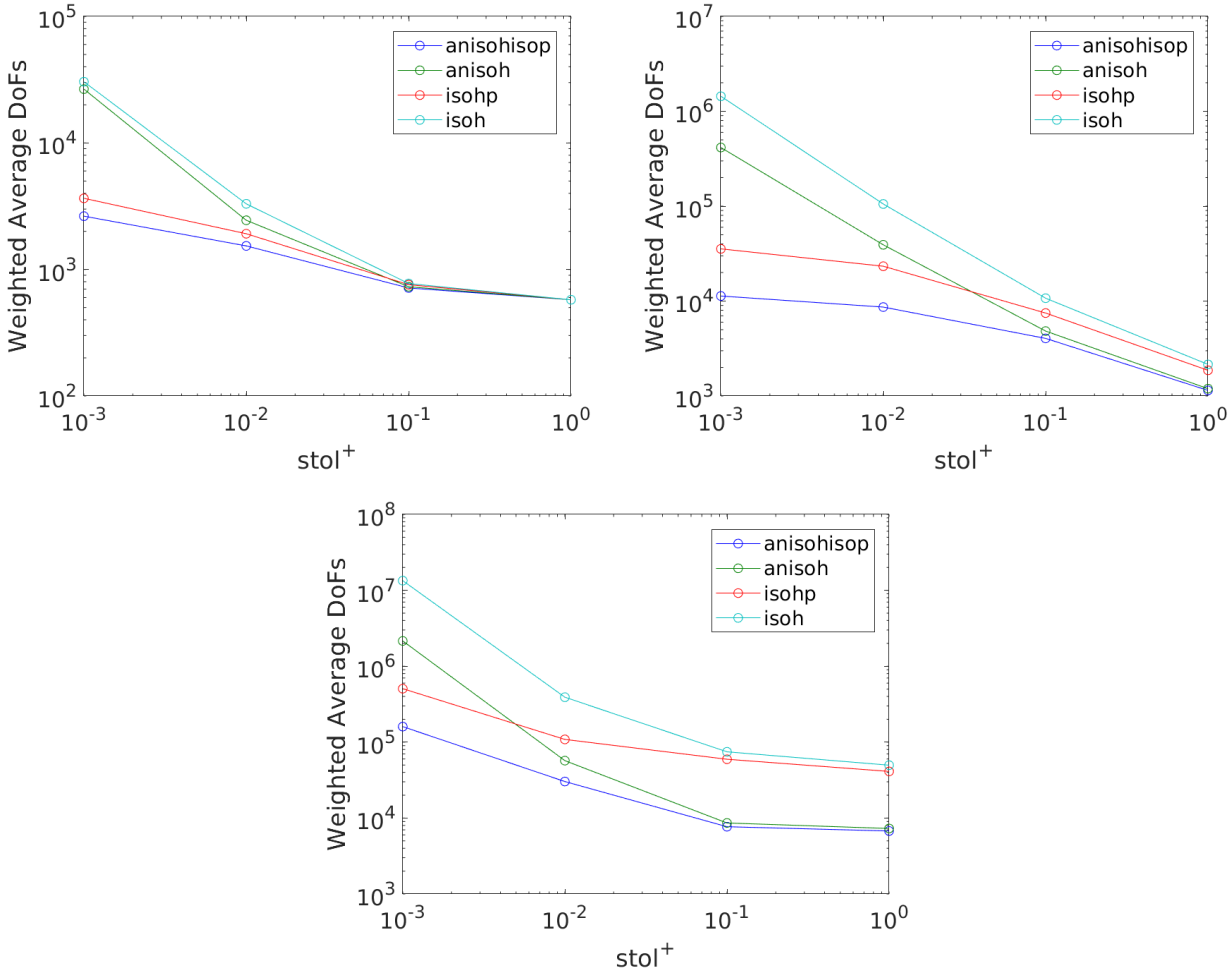

Figure 1: Average DoFS as function of $\texttt{stol}^+$ for $\varepsilon = 0.1$ (top left), $\varepsilon = 0.01$ (top right) and $\varepsilon = 0.001$ (bottom).

As can be seen from the figure, for all values of ε and for all adaptive techniques, the reduction in $\texttt{stol}^+$ leads to the Average DoFS increasing – this is obvious since more degrees of freedom are necessary when a smaller value of the threshold for the error estimator is required. However, the curves representing the different adaptive techniques are further apart for smaller values of ε suggesting that the efficiency in resolving the boundary layers is different for different adaptive strategies. Perhaps unsurprisingly, the one that seems to perform well in all situations is the anisohisop adaptive technique which combines anisotropic adaptivity in h with isotropic adaptivity in p. The former is used to resolve the boundary layer while the latter efficiently controls the error in the regions were the solution is smoother.

7.2. Faces only error estimator

In this section we explore the effectiveness of a variant of the error estimator $\eta_{S_1,j}$. In particular, we consider the error estimator $\eta_{S_1,j}$ with the interior elemental residual term removed:

$$\begin{aligned}\bar{\eta}^2_{S_1,j} := &\sum_{K\in\zeta^j}\sum_{E\in\mathcal{E}^{int}(K)} \varepsilon^{\frac{3}{2}}\alpha_E |||[\nabla u_h^j]|||^2_{L^2(E)} \\ &+ \sum_{K\in\zeta^j}\sum_{E\in\mathcal{E}(K)} \Big(\frac{\gamma^2\varepsilon p_E^3}{h^{\perp}_{E,K}} + \frac{\varepsilon h^{\perp}_{E,K} p_E^2}{h^2_{\min,K}} + \frac{\beta h^{\perp}_{E,K}}{p_E^2} + \frac{h^{\perp}_{E,K}}{\varepsilon p_E}\Big) |||[u_h^j]|||^2_{L^2(E)}.\end{aligned} \tag{41}$$

We compare the two error estimators by using them to drive the adaptivity applied to a series of test problems. Firstly, we consider the problem of the previous subsection with $\varepsilon = 0.01$, $T = 2$ and

$\tau = 0.01$. Figure 2 shows the values of the two error estimators for $\texttt{stol}^+ = 0.01$ is used to advance the solution in time. In both cases, the anisohisop adaptive technique is used and the results obtained with the reduced estimator $\bar{\eta}_{S_1,j}$ are marked 'anisohisop faces only'. We observe that the two curves are very similar except for a shift, however, considering the efficiency of the two error estimators, cf. Figure 2, it is clear that $\bar{\eta}_{S_1,j}$ is superior since it has efficiency values closer to 1. The reason for the shift can be found in the construction of the face only error estimator that makes it lower in value than the standard one on the same mesh. Hence, more iterations are needed before the estimated error reaches the tolerance and the mesh is refined. When the mesh is refined, the error estimator drops. Comparing the final adapted meshes, reported in Figure 3, we can clearly see that both error estimators detect and resolve the boundary layers. In both meshes, there is a similar distribution of the polynomial degrees over the domain. The highest degrees are sitting where the boundary layers are and the lower degrees are far away from the boundary layers. Such distribution is the results of two mechanisms: the error estimator targeting for refinement the regions where the error is most concentrated and the estimation of the local smoothness of the computed solutions used to decide between h- or p-refinement. Clearly, for this example, the error is concentrated along the boundary layers. Therefore, the regions far enough form the boundary layers are never marked by the error estimator and never considered for refinement. Consequently, the reason for low polynomial degrees far away from the boundary layers is not due to a lack of regularity of the solution in such regions, rather the presence of a small accuracy error. The variation of the polynomial degrees is higher near the boundary layers since those regions have been heavily refined. Locally, the solution may be smooth even next to a boundary layer, hence a combination of both h- or p-refinement it is to be expected.

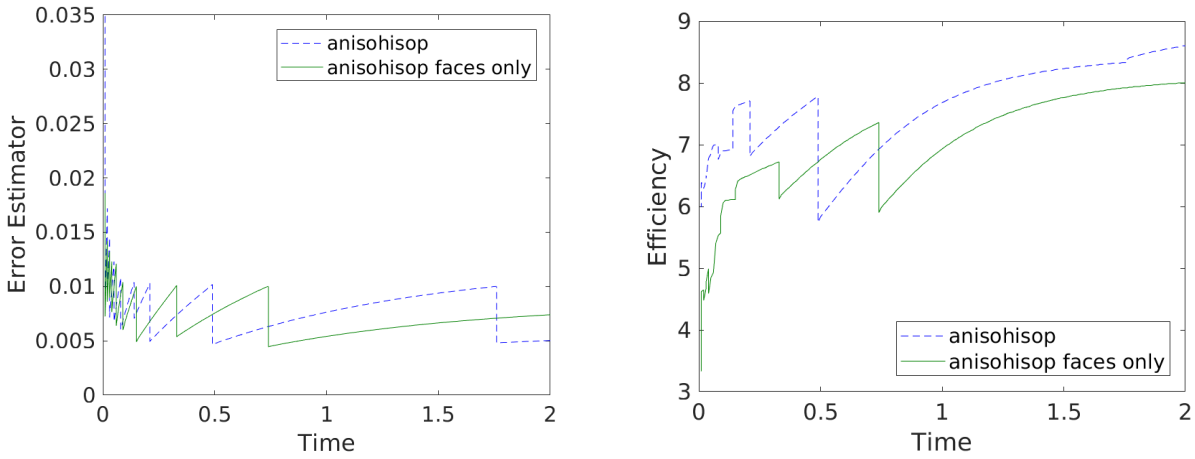

Figure 2: Comparison between the error estimators $\eta_{S_1,j}$ (anisohisop) and $\bar{\eta}_{S_1,j}$ (anisohisop faces only): values of the estimators (left) and efficiency of the estimators (right).

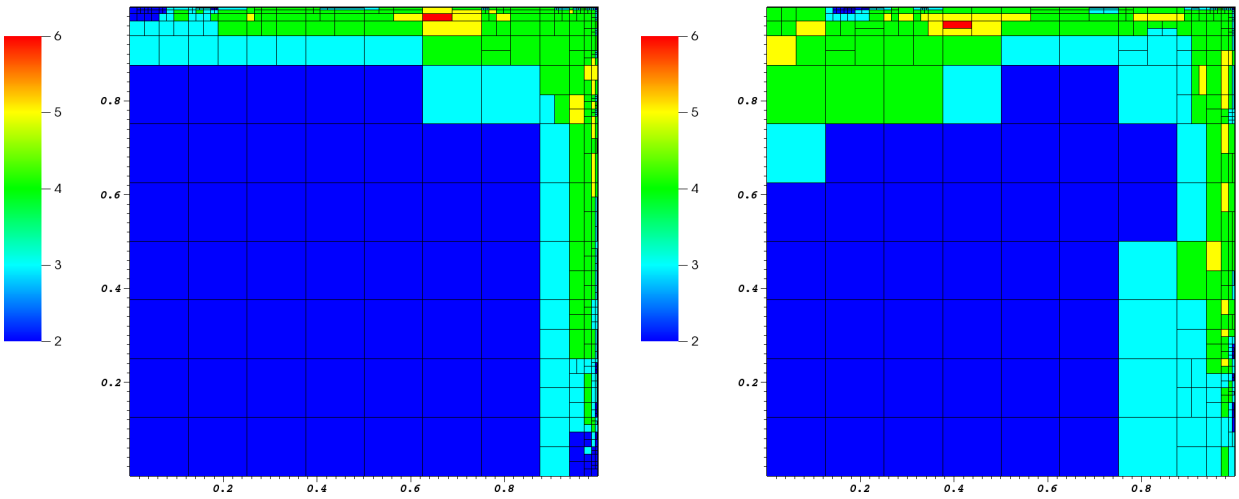

Figure 3: Final ansitropically adapted mesh using $\eta_{S_1,j}$ (left) and $\bar{\eta}_{S_1,j}$ (right). The order of the polynomials on each element is expressed using different colours.

We also compare the results of using the error estimator $\bar{\eta}_{S_1,j}$ with anisohisop adaptivity and isohp adaptivity. To this end, we consider a test problem with an internal layer that is not aligned with the mesh (see Figure 4). In particular, we set $\Omega = (-0.5, 0.5)^2$, $\mathbf{a} = (1,1)^T$, $b = 0$, $u_0 = 0$, $T = 1$, $\varepsilon = 0.01$ and we select the function f so that the exact solution to problem (5) is given by

$$u(x,y,t) = (1 - e^{-t}) \arctan\left(\frac{y-x}{\varepsilon\sqrt{2}}\right)\left(1 - \frac{x+y}{\sqrt{2}}^2\right).$$

To drive the refinement in space we set $\texttt{stol}^+ = 0.01$. Since the focus of this numerical experiment is the space error estimator $\bar{\eta}_{S_1,j}$, to limit the temporal contribution to the error τ is set to 0.01. Also, to make the comparison more fair, we switch off the refinement in time setting $\texttt{ttol}^+ = 10000$ and $\texttt{ttol}^- = 0$. In this way the time steps are the same for both simulations.

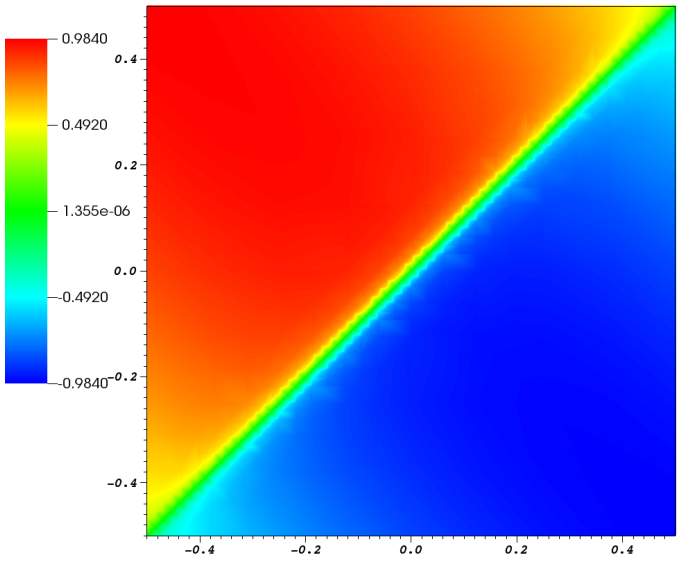

Figure 4: Numerical solution at $t = 1$.

In Figure 5, we report the value of the error estimator using either anisohisop adaptivity or isohp adaptivity. As the layer is not aligned with the mesh, anisotropic adaptivity is never chosen by the algorithm leading to no difference in the error estimator values. Additionally, the meshes at final time look identical as can be seen in Figure 6.

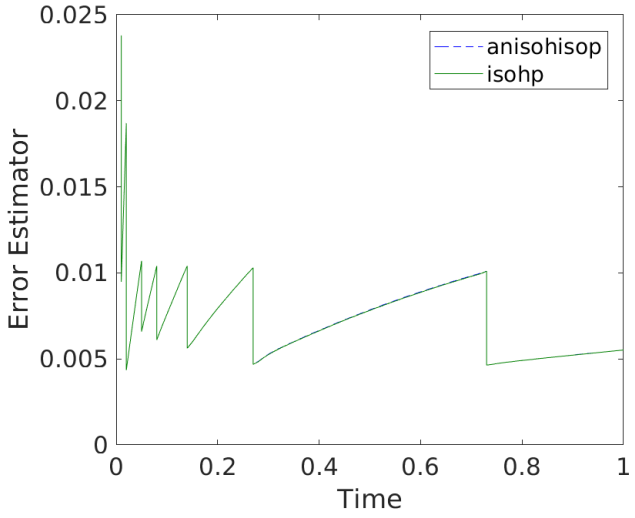

Figure 5: Comparison between anisohisop adaptivity and isohp adaptivity.

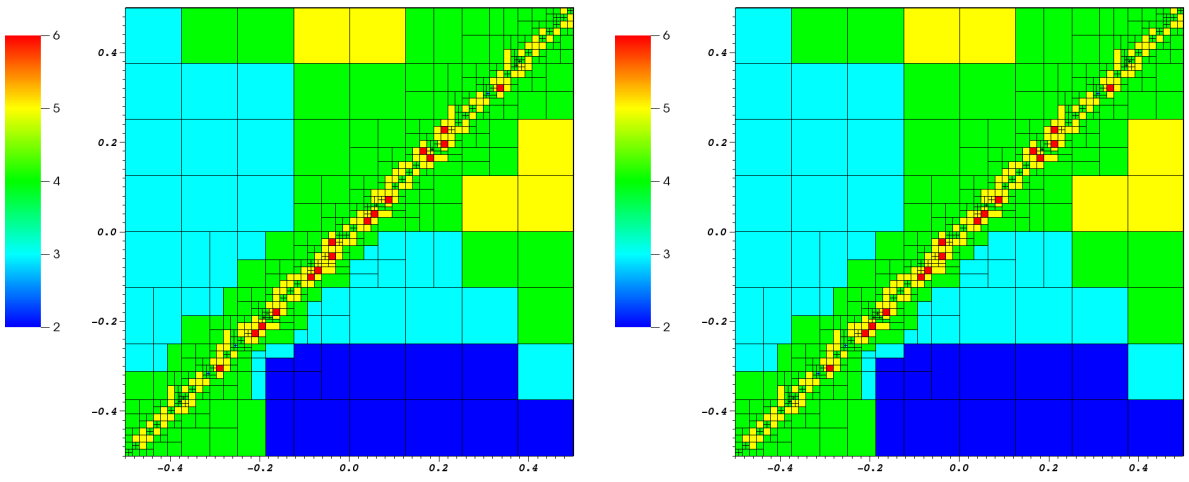

Figure 6: Anisotropically (left) and isotropically (right) adapted meshes at final time. The order of the polynomials on each element is expressed using different colours.

Finally, we again use the error estimator $\bar{\eta}_{S_1,j}$ to drive the adaptivity but this time we apply the algorithm to a problem for which the analytical solution is not known. To that end, we let $\Omega = (0,2)\times(0,1)$, $\mathbf{a} = (1,0)^T$, $b = 0$, $T = 1$, $\varepsilon = 0.0001$ and the initial value be given by $u_0(x,y) = \exp(-200(x-0.2)^2 + (y-0.5)^2)$. Setting $\texttt{stol}^+ = 0.01$ and using anisohisop adaptivity, we obtain the three meshes reported in Figure 7. As for the previous set of simulations, in this test problem the focus is on the capability of the space error estimator. In particular we test its ability to follow a non-stationary solution, therefore τ is set to 0.01 and the refinement in time is switched off by setting $\texttt{ttol}^+ = 10000$ and $\texttt{ttol}^- = 0$.

We observe that the estimator successfully tracks the location of the solution and removes irrelevant elements. Furthermore, contrary to the previous example, we notice that the algorithm does select anisotropic adaptivity aligning the elements shapes with the solution's layers, hence achieving the required resolution with fewer degrees of freedom.

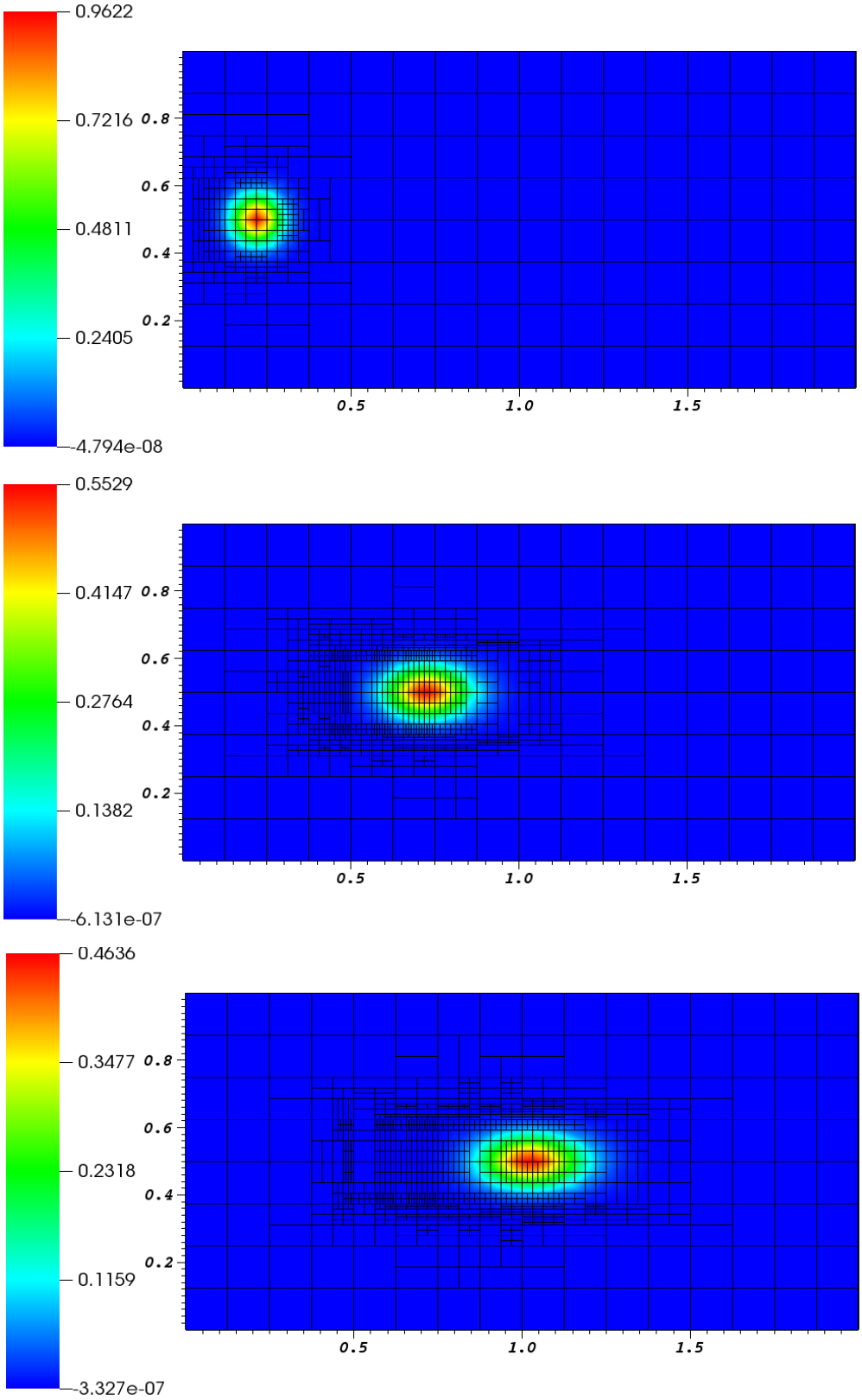

Figure 7: Anisotropic meshes and solution. Top to bottom: $t = 0.02$, $t = 0.53$ and $t = 0.83$.

8. Conclusions

We presented, for the first time, an anisotropic *hp* a posteriori error estimator for a discontinuous Galerkin method applied to non-stationary convection-diffusion problems with space-time adaptivity.

Through a set of numerical experiments we have shown that, using our error estimator to drive hp-adaptivity, the error can be controlled with fewer degrees of freedom for any given tolerance, hence reducing computational cost. To further improve efficiency, a computationally cheaper variation of the error estimator containing only face terms has been tested. The error estimator is capable to detect layers aligned with the mesh and refine accordingly to reduce the error more efficiently. In particular, this is true for non-stationary solution with moving layers, whereby the error estimator is able, also in the anisotropic setting, to follow the solution and align the elements shapes with the solution's layers.

We have shown that more sophisticated error estimators for non-stationary convection-diffusion problems lead to automatic adaptive algorithms able to produce truly layer-adapted meshes, thus opening the door to more highly-accurate but efficient simulations.

References

[1] M. Ainsworth, J.T. Oden, A posteriori error estimation in finite element analysis, Wiley-Interscience, 2000.

[2] P. Amestoy, I. Duff, J. Koster, J. L'Excellent, A fully asynchronous multifrontal solver using distributed dynamic scheduling, SIAM Journal on Matrix Analysis and Applications 23 (2001) 15–41.

[3] P. Amestoy, I. Duff, J. L'Excellent, Multifrontal parallel distributed symmetric and unsymmetric solvers, Computer Methods in Applied Mechanics and Engineering 184 (2000) 501–520.

[4] P. Amestoy, A. Guermouche, J. L'Excellent, S. Pralet, Hybrid scheduling for the parallel solution of linear systems, Parallel Computing 32 (2006) 136–156.

[5] R. Araya, E. Behrens, R. Rodríguez, An adaptive stabilized finite element scheme for the advection-reaction-diffusion equation, Applied Numerical Mathematics 54 (2005) 491–503.

[6] R. Araya, A.H. Poza, E.P. Stephan, A hierarchical a posteriori error estimate for an advection-diffusion-reaction problem, Mathematical Models & Methods in Applied Sciences 15 (2005) 1119–1139.

[7] E. Bänsch, F. Karakatsani, C. Makridakis, The effect of mesh modification in time on the error control of fully discrete approximations for parabolic equations, Applied Numerical Mathematics 67 (2013) 35–63.

[8] R. Becker, P. Hansbo, M.G. Larson, Energy norm a posteriori error estimation for discontinuous Galerkin methods, Computer Methods in Applied Mechanics and Engineering 192 (2003) 723–733.

[9] S. Berrone, C. Canuto, Multilevel a posteriori error analysis for reaction-convection-diffusion problems, Applied Numerical Mathematics 50 (2004) 371–394.

[10] A. Cangiani, Z. Dong, E. Georgoulis, P. Houston, hp-Version Discontinuous Galerkin Methods on Polygonal and Polyhedral Meshes, Springer Briefs in Mathematics, Springer, 2017.

[11] A. Cangiani, E.H. Georgoulis, S. Metcalfe, Adaptive discontinuous Galerkin methods for nonstationary convection-diffusion problems, IMA Journal of Numerical Analysis 34 (2014) 1578–1597.

[12] V. Dolejší, A. Ern, M. Vohralík, A framework for robust a posteriori error control in unsteady nonlinear advection-diffusion problems, SIAM Journal on Numerical Analysis 51 (2013) 773–793.

[13] T. Dupont, Mesh modification for evolution equations, Mathematics of Computation 39 (1982) 85–107.

[14] A. Ern, J. Proft, A posteriori discontinuous Galerkin error estimates for transient convection-diffusion equations, Applied Mathematics Letters 18 (2005) 833–841.

[15] A. Ern, M. Vohralík, A posteriori error estimation based on potential and flux reconstruction for the heat equation, SIAM Journal on Numerical Analysis 48 (2010) 198–223.

[16] E.H. Georgoulis, Discontinuous Galerkin methods on shape-regular and anisotropic meshes, PhD Thesis, University of Oxford (2003).

[17] E.H. Georgoulis, E. Hall, P. Houston, Discontinuous Galerkin methods for advection-diffusion-reaction problems on anisotropically refined meshes, SIAM Journal on Scientific Computing 30 (2008) 246–271.

[18] E.H. Georgoulis, O. Lakkis, J.M. Virtanen, A posteriori error control for discontinuous Galerkin methods for parabolic problems, SIAM Journal on Numerical Analysis 49 (2011) 427–458.

[19] S. Giani, D. Schötzau, L. Zhu, An a-posteriori error estimate for hp-adaptive DG methods for convection-diffusion problems on anisotropically refined meshes, Computers & Mathematics with Applications 67 (2014) 869–887.

[20] P. Houston, D. Schötzau, T.P. Wihler, Energy norm a posteriori error estimation of hp-adaptive discontinuous Galerkin methods for elliptic problems, Mathematical Models & Methods in Applied Sciences 17 (2007) 33–62.

[21] P. Houston, E. Süli, Adaptive Lagrange-Galerkin methods for unsteady convection-diffusion problems, Mathematics of Computation 70 (2001) 77–106.

[22] P. Houston, E. Süli, A note on the design of hp-adaptive finite element methods for elliptic partial differential equations, Computer Methods in Applied Mechanics and Engineering 194 (2005) 229–243.

[23] O.A. Karakashian, F. Pascal, A posteriori error estimates for a discontinuous Galerkin approximation of second-order elliptic problems, SIAM Journal on Numerical Analysis 41 (2003) 2374–2399.

[24] O.A. Karakashian, F. Pascal, Convergence of adaptive discontinuous Galerkin approximations of second-order elliptic problems, SIAM Journal on Numerical Analysis 45 (2007) 641–665.

[25] G. Kunert, A posteriori error estimation for convection dominated problems on anisotropic meshes, Mathematical Methods in the Applied Sciences 26 (2003) 589–617.

[26] O. Lakkis, C. Makridakis, Elliptic reconstruction and a posteriori error estimates for fully discrete linear parabolic problems, Mathematics of Computation 75 (2006) 1627–1658.

[27] A. Lozinski, M. Picasso, V. Prachittham, An anisotropic error estimator for the Crank-Nicolson method: application to a parabolic problem, SIAM Journal on Scientific Computing 31 (2009) 2757–2783.

[28] C. Makridakis, R.H. Nochetto, Elliptic reconstruction and a posteriori error estimates for parabolic problems, SIAM Journal on Numerical Analysis 41 (2003) 1585–1594.

[29] S. Metcalfe, Adaptive discontinuous Galerkin methods for nonlinear parabolic problems, Ph.D. thesis, PhD Thesis, University of Leicester, 2015.

[30] H.G. Roos, M. Stynes, L. Tobiska, Robust numerical methods for singularly perturbed differential equations, Springer-Verlag, 2008.

[31] G. Sangalli, Robust a-posteriori estimator for advection-diffusion-reaction problems, Mathematics of Computation 77 (2008) 41–70.

[32] D. Schötzau, L. Zhu, A robust a-posteriori error estimator for discontinuous Galerkin methods for convection-diffusion equations, Applied Numerical Mathematics 59 (2009) 2236–2255.

[33] D. Schötzau, L. Zhu, A r estimate for hp-adaptive DG methods for convection-diffusion equations, IMA Journal of Numerical Analysis 31 (2011) 971–1005.

[34] I. Šebestová, Two-sided a posteriori error estimates for the DGMs for the heat equation, in: Numerical Mathematics and Advanced Applications, Springer, 2013, pp. 379–387.

[35] S. Sun, M.F. Wheeler, A posteriori error estimation and dynamic adaptivity for symmetric discontinuous Galerkin approximations of reactive transport problems, Computer Methods in Applied Mechanics and Engineering 195 (2006) 632–652.

[36] R. Verfürth, A review of a posteriori error estimation and adaptive mesh-refinement techniques, Wiley-Teubner, 1996.

[37] R. Verfürth, A posteriori error estimators for convection-diffusion equations, Numerische Mathematik 80 (1998) 641–663.

[38] R. Verfürth, Robust a posteriori error estimates for nonstationary convection-diffusion equations, SIAM Journal on Numerical Analysis 43 (2005) 1783–1802.

[39] R. Verfürth, Robust a posteriori error estimates for stationary convection-diffusion equations, SIAM Journal on Numerical Analysis 43 (2005) 1766–1782.