A robust a-posteriori error estimate for $hp$-adaptive DG methods for convection-diffusion equations

LIANG ZHU AND DOMINIK SCHÖTZAU

Mathematics Department, University of British Columbia,
1984 Mathematics Road, Vancouver, BC, V6T 1Z2, Canada,
Email: {zhuliang,schoetzau}@math.ubc.ca


We derive a robust a-posteriori error estimate for $hp$-adaptive discontinuous Galerkin (DG) discretizations of stationary convection-diffusion equations. We consider 1-irregular meshes consisting of parallelograms. The estimate yields global upper and lower bounds of the errors measured in terms of the natural energy norm associated with the diffusion and a semi-norm associated with the convection. The ratio of the constants in the upper and lower bounds is independent of the local mesh sizes and weakly depending on the local polynomial degrees. Moreover, it is also independent of the magnitude of the Péclet number of the problem, and hence the estimate is fully robust for convection-dominated problems. We apply our estimator as an energy norm error indicator in an $hp$-adaptive refinement algorithm and illustrate its practical performance in a series of numerical examples.

Keywords: Discontinuous Galerkin (DG) methods, robust a-posteriori error estimation, $hp$-adaptivity, convection-diffusion equations

1. Introduction

It is well-known that solutions to convection-diffusion equations may have boundary or internal layers of small width where their gradients change extremely rapidly. One way to efficiently approximate convection-diffusion problems is to use adaptive finite element methods that are capable of locally refining the meshes in the vicinity of these layers. The decision when to refine an element is usually based on a-posteriori estimates of the errors (or functionals thereof). For excellent surveys on adaptive finite elements and a-posteriori error estimation, we refer to Ainsworth and Oden (2000) and Verfürth (1996).

The design of robust energy norm a-posteriori error estimates has attracted a lot of attention recently. Here, by robustness we mean that the estimates yield upper and lower bounds for the errors measured in suitable norms that differ by a factor that is independent of the Péclet number of the problem. Several robust and semi-robust a-posteriori estimates for energy norm error estimation can now be found in the literature. For conforming and mixed finite element methods, we refer to the recent papers by Sangalli (2005), Sangalli (2008), Verfürth (1998), Verfürth (2005), Vohralik (2007) and the references therein. The robustness of a-posteriori energy norm error estimates for non-conforming and discontinuous Galerkin (DG) finite element methods has been studied by Alaoui et al. (2007), Ern and Stephansen (2008), Ern et al. (2008), Schötzau and Zhu (2009) and the references therein. However, all the papers above are concerned with $h$-version finite element methods. These methods are based on employing a fixed, usually low polynomial degree. As a consequence, adaptive $h$-version methods yield at most algebraic rates of convergence. This is in contrast to $hp$-version finite element methods, where
the combination of $h$-refinement and $p$-refinement typically results in exponential rates of convergence, see, e.g., Schwab (1998) and the references therein.

Discontinuous Galerkin methods are naturally suited for realizing $hp$-adaptivity. Indeed, being based on discontinuous finite element spaces, these methods can easily deal with irregularly refined meshes and locally varying polynomial degrees. For recent accounts on the state-of-the-art of DG methods we refer to reader to Arnold et al. (2002), Cockburn et al. (2000), Cockburn and Shu (2001), Houston et al. (2002) and the references therein. Several approaches to energy norm error estimation for DG methods applied to elliptic problems can be found in the literature; see, e.g., Becker, Hansbo and Larson (2003), Becker, Hansbo and Stenberg (2003) and Karakashian and Pascal (2003). Extensions to $hp$-version DG methods have been recently developed by Houston et al. (2006, 2007, 2008).

In this paper, we extend the $h$-version technique proposed by Schötzau and Zhu (2009) to the $hp$-version of the DG method and derive a robust a-posteriori error estimate for convection-diffusion equations. Similarly to Verfürth (2005), we introduce as an error measure the natural energy norm and a dual norm associated with the convection. In this measure, we derive upper and lower bounds of the errors, which are explicit in the local mesh sizes and polynomial degrees. The ratio of the constants in the upper and lower bounds is independent of the local mesh size and weakly depending on the polynomial degrees. More importantly, it is independent of the Péclet number of the problem; hence, our estimate is robust. In our analysis, the error is decomposed into a conforming part and a remainder using an averaging operator as in the approaches of Houston et al. (2007) and Karakashian and Pascal (2003). The conforming contribution of the error can be dealt with using standard techniques, while the rest term is shown to be controlled by the jump. A major ingredient of our analysis is a new $L^2$-norm approximation property for the $hp$-version averaging operator. In Houston et al. (2007), an optimal $H^1$-seminorm approximation property was established on regular meshes without hanging nodes. It was then extended by Houston et al. (2008) to irregular meshes. Burman and Ern (2007) proved optimal $L^2$-norm (and $H^1$-seminorm) estimates on regular meshes and for fixed polynomial degrees. We extend the $L^2$-norm estimate to the case of 1-irregular meshes consisting of parallelograms and variable polynomial degrees. Similarly to Houston et al. (2008), we also use an auxiliary mesh underlying the possibly irregular computational mesh. However, to obtain the $L^2$-norm estimate, we also allow the auxiliary mesh to be 1-irregular.

We present a series of numerical tests where we use our a-posteriori error estimator as an error indicator in an $hp$-adaptive algorithm. To decide whether to apply $h$- or $p$-refinement of marked elements, we employ the smoothness estimation strategy developed by Houston et al. (2003) and Houston and Süli (2005). Our numerical examples indicate that our algorithm is effective in locating and resolving boundary layers. Moreover, we observe that both the energy error and the error indicator converge exponentially once the local mesh size is sufficiently small.

The outline of the rest of this article is as follows. In Section 2, we introduce $hp$-adaptive discontinuous Galerkin methods for a convection-diffusion model problem. In Section 3, we state and discuss our robust a-posteriori error estimate. The proof of this estimate is carried out in Sections 4 and 5. In Section 6, we present a series of numerical tests that illustrate the theoretical results. Finally, in Section 7, we end with some concluding remarks.

### 2. Interior penalty discretization of convection-diffusion problems

In this section, we introduce an $hp$-adaptive interior penalty discontinuous Galerkin finite element method for the discretization of convection-diffusion equations.
2.1 Model problem
We consider the convection-diffusion model problem:

\[-\varepsilon \Delta u + g(x) \cdot \nabla u = f(x) \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \Gamma.\]

(2.1)

Here, \( \Omega \) is a bounded Lipschitz polygon in \( \mathbb{R}^2 \) with boundary \( \Gamma = \partial \Omega \). The parameter \( \varepsilon > 0 \) is the (constant) diffusion coefficient, the vector-valued function \( g(x) \) a given flow field, and the function \( f(x) \) a generic right-hand side in \( L^2(\Omega) \). The coefficient \( g(x) \) is assumed to belong to \( W^{1,\infty}(\Omega)^2 \) and to satisfy

\[ \nabla \cdot g = 0 \quad \text{in } \Omega. \]

(2.2)

Without loss of generality, we shall assume that \( \|g\|_{L^\infty(\Omega)} \) and the length scale of \( \Omega \) are one so that \( \varepsilon^{-1} \) is the Péclet number of the problem. The standard weak form of the convection-diffusion equation in (2.1) is to find \( u \in H^1_0(\Omega) \) such that

\[ A(u, v) = \int_\Omega (\varepsilon \nabla u \cdot \nabla v + g \cdot \nabla v) \, dx = \int_\Omega f v \, dx \quad \forall v \in H^1_0(\Omega). \]

(2.3)

Under assumption (2.2), the variational problem (2.3) is uniquely solvable.

2.2 Discretization
Throughout, we assume that the computational domain \( \Omega \) can be partitioned into shape-regular (sequences of) meshes \( \mathcal{T} = \{K\} \) of parallelograms \( K \). Each element \( K \in \mathcal{T} \) is the image of the reference square \( \bar{K} = (-1,1)^2 \) under an affine elemental mapping \( F_K : \bar{K} \rightarrow K \). As usual, we denote by \( h_K \) the diameter of \( K \). We store the elemental diameters in the mesh size vector \( h = \{h_K : K \in \mathcal{T}\} \).

We will make use of the following sets of vertices and edges. For an element \( K \in \mathcal{T} \), we denote by \( \mathcal{N}(K) \) the set of its four vertices. A node \( v \) of a finite element mesh \( \mathcal{T} \) is the vertex of at least one element \( K \in \mathcal{T} \). The node \( v \) is called an interior node if \( v \notin \Gamma \); similarly, it is a boundary node if \( v \in \Gamma \). We denote by \( \mathcal{N}_I(\mathcal{T}) \), \( \mathcal{N}_B(\mathcal{T}) \) the sets of interior and boundary nodes, respectively, and set \( \mathcal{N}(\mathcal{T}) = \mathcal{N}_I(\mathcal{T}) \cup \mathcal{N}_B(\mathcal{T}) \). Further, we denote by \( \mathcal{E}(K) \) the set of its four elemental edges. If the intersection \( E = \partial K \cap \partial K' \) of two elements \( K, K' \in \mathcal{T} \) is a proper line segment (and not a single point), we call \( E \) an interior edge of \( \mathcal{T} \). The set of all interior edges is denoted by \( \mathcal{E}_I(\mathcal{T}) \). Analogously, if the intersection \( E = \partial K \cap \Gamma \) of an element \( K \in \mathcal{T} \) and \( \Gamma \) is a proper line segment, we call \( E \) 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 \( \mathcal{E}(\mathcal{T}) = \mathcal{E}_I(\mathcal{T}) \cup \mathcal{E}_B(\mathcal{T}) \). We denote by \( h_E \) the length of an edge \( E \in \mathcal{E}(K) \) or \( E \in \mathcal{E}(\mathcal{T}) \).

In our analysis, we allow for 1-irregularly refined meshes \( \mathcal{T} \) where each elemental edge \( E \in \mathcal{E}(K) \) may contain at most one hanging node located in the middle of it. That is, we either have \( E \in \mathcal{E}(\mathcal{T}) \) or \( E \) can be written as \( E = E_1 \cup E_2 \), for two edges \( E_1 \) and \( E_2 \) in \( \mathcal{E}(\mathcal{T}) \) that satisfy \( h_{E_1} = h_{E_2} = h_E/2 \).

Next, let us define the jumps and averages of piecewise smooth functions across edges of a mesh \( \mathcal{T} \). To that end, let the interior edge \( E \in \mathcal{E}_I(\mathcal{T}) \) be shared by two neighbouring elements \( K \) and \( K' \) where the superscript \( e \) stands for "exterior". For a piecewise smooth function \( v \), we denote by \( v|_E \) its trace on \( E \) taken from inside \( K \), and by \( v^e|_E \) the one taken from inside \( K' \). The average and jump of \( v \) across the edge \( E \) are then defined as

\[ \{v\} = \frac{1}{2} (v|_E + v^e|_E), \quad \llbracket v \rrbracket = v|_E n_K + v^e|_E n_{K'}. \]
Here, $n_K$ signifies the unit outward normal on the boundary of element $K$. Similarly, if $q$ is piecewise smooth vector field, its average and (normal) jump across $E$ are given by

$$\{\{q\}\} = \frac{1}{2} (q|_{E} + \text{q}^\varepsilon|_{E}), \quad [q] = q|_{E} \cdot n_K + q^\varepsilon|_{E} \cdot n_K.$$  

On a boundary edge $E \in \partial_\gamma(\mathcal{T})$ shared by $\Gamma$ and $\partial K$, we set accordingly $\{\{q\}\} = q$ and $[v] = v_n$, with $n$ denoting the unit outward normal vector on $\Gamma$.

With each element $K$ of a mesh $\mathcal{T}$, we associate a polynomial degree $p_K \geq 1$, introduce the degree 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 of bounded local variation. That is, there is a constant $\rho \geq 1$ independent of the particular mesh in a sequence of meshes such that, for any pair of neighbouring elements $K, K' \in \mathcal{T}$, we have $\rho^{-1} \leq p_K / p_{K'} \leq \rho$. For $E \in \partial_\gamma(\mathcal{T})$, we introduce the edge polynomial degree $p_E$ by

$$p_E = \begin{cases} \max\{ p_K, p_{K'} \}, & E = \partial K \cap \partial K' \in \partial_\gamma(\mathcal{T}), \\ p_K, & E = \partial K \cap \Gamma \in \partial_\gamma(\mathcal{T}). \end{cases} \quad (2.4)$$

For a partition $\mathcal{T}$ of $\Omega$ and a degree vector $\mathbf{p}$ on $\mathcal{T}$, we then define the $hp$-version discontinuous Galerkin finite element space by

$$S_p(\mathcal{T}) = \{ v \in L^2(\Omega) : v|_K \in \mathbb{P}_{p_K}(\hat{K}), K \in \mathcal{T} \}, \quad (2.5)$$

with $\mathbb{P}_p(\hat{K})$ denoting the set of all polynomials on the reference square $\hat{K}$ of degree less or equal than $p$ in each variable.

We now consider the following discontinuous Galerkin method for the numerical approximation of (2.1): Find $u_{hp} \in S_p(\mathcal{T})$ such that

$$A_{hp}(u_{hp}, v) = \int_\Omega f v \, dx \quad (2.6)$$

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

$$A_{hp}(u,v) = \sum_{K \in \mathcal{T}} \int_K (\varepsilon \nabla u \cdot \nabla v + a \cdot \nabla u v) \, dx - \sum_{E \in \partial_\gamma(\mathcal{T})} \int_E \{\{\varepsilon \nabla u\}\} \cdot [v] \, ds - \sum_{E \in \partial_\gamma(\mathcal{T})} \int_E \{\{\varepsilon \nabla v\}\} \cdot [u] \, ds + \sum_{E \in \partial_\gamma(\mathcal{T})} \int_E \varepsilon h_E \frac{\gamma p_E^2}{h_E} [u] \cdot [v] \, ds - \sum_{K \in \mathcal{T}} \int_{\partial K} a \cdot n_K uv \, ds + \sum_{K \in \mathcal{T}} \int_{\partial K} \frac{a \cdot n_K(u^\varepsilon - u) v}{h_K} \, ds.$$

Here, for a piecewise smooth function, the gradient operator $\nabla$ is taken element by element. The constant $\gamma > 0$ is the interior penalty parameter. To ensure the stability and well-posedness of the discontinuous Galerkin discretization, it is well-known that it has to be chosen sufficiently large, independently of $h$, $\mathbf{p}$ and $\varepsilon$. Finally, we denote by $\Gamma_{in}$ and $\partial K_{in}$ the inflow parts of $\Gamma$ and $K \in \mathcal{T}$, respectively:

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

**Remark 2.1** The discretization (2.6) is based on the original upwind discretization of Lesaint and Raviart (1974) and Reed and Hill (1973) for the convective term, and on the classical symmetric interior penalty discretization of Arnold (1980) and Nitsche (1971) for the diffusion term; see also Arnold et al. (2002).
3. Robust a-posteriori error estimates

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

3.1 Norms

We begin by introducing the norms in which the errors are measured. First, we introduce the following energy norm associated with the discontinuous Galerkin discretization of the diffusive term:

$$\|v\|_{E,T}^2 = \sum_{K \in T} \varepsilon \|\nabla v\|_{L^2(K)}^2 + \text{ejump}_{p,T}(v)^2,$$

(3.1)

Next, we define

$$|q|_* = \sup_{v \in H_0^1(\Omega) \setminus \{0\}} \frac{\int_{\Omega} q \cdot \nabla v \, dx}{\|v\|_{E,T}} \quad \forall q \in L^2(\Omega).$$

Analogously to Verfürth (2005) and Schötzau and Zhu (2009), we introduce the following semi-norm associated with the discretization of the convection term:

$$|v|_{O,T} = |av|_*^2 + \text{ojump}_{p,T}(v)^2,$$

$$\text{ojump}_{p,T}(v)^2 = \sum_{E \in \partial K \setminus \Gamma} \frac{h_E}{\varepsilon p} \|\nabla u\|_{L^2(E)}^2.$$

(3.2)

Notice that $h_E \varepsilon^{-1}$ is the local mesh Péclet number.

3.2 A robust a-posteriori error estimate

Let now $u_{hp} \in S_p(\mathcal{T})$ be the discontinuous Galerkin approximation obtained by (2.6). Moreover, let $f_{hp}$ and $a_{hp}$ denote piecewise polynomial approximations in $S_p(\mathcal{T})$ to the right-hand side $f$ and the flow field $a$, respectively. For each element $K \in \mathcal{T}$, we introduce the following local error indicator $\eta_K$ which is given by the sum of the three terms

$$\eta_K^2 = \eta_{R_K}^2 + \eta_{E_K}^2 + \eta_{J_K}^2.$$  

(3.3)

The first term $\eta_{R_K}$ is the interior residual defined by

$$\eta_{R_K}^2 = \varepsilon^{-1} p_K^{-2} h_K^2 \|f_{hp} + \varepsilon \Delta u_{hp} - a_{hp} \cdot \nabla u_{hp}\|_{L^2(K)}^2.$$

The second term $\eta_{E_K}$ is the edge residual given by

$$\eta_{E_K}^2 = \frac{1}{2} \sum_{E \in \partial K \setminus \Gamma} \varepsilon^{-1} p_E^{-1} h_E \|\nabla u_{hp}\|_{L^2(E)}^2.$$

The last residual $\eta_{J_K}$ measures the jumps of the approximate solution $u_{hp}$:

$$\eta_{J_K}^2 = \frac{1}{2} \sum_{E \in \partial K \setminus \Gamma} \left( \frac{\varepsilon p_E^3}{h_E} + \frac{h_E}{\varepsilon p_E} \right) \|u_{hp}\|_{L^2(E)}^2 + \sum_{E \in \partial K \setminus \Gamma} \left( \frac{\varepsilon p_E^3}{h_E} + \frac{h_E}{\varepsilon p_E} \right) \|u_{hp}\|_{L^2(E)}^2.$$
We also introduce the local data approximation term
\[
\Theta_K^2 = \varepsilon^{-1} p^2 h_K^2 \left( \| f - f_h \|_{L^2(K)}^2 + \| (a - a_h) \cdot \nabla u_h \|_{L^2(K)}^2 \right).
\]

We then introduce the global error estimator and data approximation error
\[
\eta^2 = \sum_{K \in \mathcal{T}} \eta_K^2, \quad \Theta^2 = \sum_{K \in \mathcal{T}} \Theta_K^2.
\]

In the following, we use the symbols \( \lesssim \) and \( \gtrsim \) to denote bounds that are valid up to positive constants, independently of \( h, p \) and \( \varepsilon \).

**Theorem 3.1 (Reliability)** Let \( u \) be the solution of (2.1) and \( u_h \in S_p(\mathcal{T}) \) its DG approximation obtained by (2.6). Let the error estimator \( \eta \) and the data approximation error \( \Theta \) be defined by (3.4). Then we have the a-posteriori error bound
\[
\| u - u_h \|_{E,\mathcal{T}} + | u - u_h |_{O,\mathcal{T}} \lesssim \eta + \Theta.
\]

**Remark 3.1** The power \( \gamma^2 p^3 \) in \( \eta_K \) is slightly suboptimal with respect to the one used in the jump terms of the energy norm (3.1). This suboptimality is due to the possible presence of hanging nodes in \( \mathcal{T} \). Indeed, for conforming meshes, the conforming \( hp \)-version Clément interpolant constructed by Melenk (2005) can be employed in our proof; see also Houston et al. (2007). As a consequence, Theorem 3.1 holds true with the following version of \( \eta_K \):
\[
\eta_K^2 = \frac{1}{2} \sum_{E \in \partial K \setminus \Gamma} \left( \frac{\gamma p^3}{h_E \varepsilon p} + \frac{h_E}{\varepsilon p} \right) \| u_h \|_{L^2(E)}^2 + \sum_{E \in \partial K \setminus \Gamma} \left( \frac{\gamma p^3}{h_E \varepsilon p} + \frac{h_E}{\varepsilon p} \right) \| u_h \|_{L^2(E)}^2.
\]

On the other hand, the numerical results in Section 6 indicate that the two versions of \( \eta_K \) yield practically identical results on 1-irregularly refined square meshes.

Our next theorem derives a lower bound for the error measured in terms of the energy norm and the semi-norm \( | \cdot |_{O,\mathcal{T}} \). For \( p \)-independence in both the upper and lower bounds, special weighting techniques seem to be necessary which we do not pursue in this article; see Braess et al. (2008). Here, we only present a weakly \( p \)-dependent lower bound for the a-posteriori error estimator \( \eta_K \) defined above.

**Theorem 3.2 (Efficiency)** Let \( u \) be the solution of (2.1) and \( u_h \in S_p(\mathcal{T}) \) its DG approximation obtained by (2.6). Let the error estimator \( \eta \) and the data approximation error \( \Theta \) be defined by (3.4). Then for any \( \delta \in (0, \frac{1}{2}) \) we have the bound
\[
\eta \lesssim | p |^{\delta + 1} \| u - u_h \|_{E,\mathcal{T}} + | p |^{2 \delta + 1} | u - u_h |_{O,\mathcal{T}} + | p |^{2 \delta + 1} \Theta.
\]

As the ratio of the constants in the upper and lower bounds in Theorem 3.1 and Theorem 3.2 is independent of the Péclet number of (2.1), the estimator \( \eta \) is robust.

**4. Proofs**

In this section, we present the proofs of Theorems 3.1 and 3.2.
4.1 Stability and auxiliary forms

The following inf-sup condition for the continuous form $A$ is the crucial stability result in our analysis. It holds with an absolute constant, which can be immediately inferred from Lemma 4.4 of Schötzau and Zhu (2009).

**Lemma 4.1** Assume (2.2). Then we have

$$\inf_{u \in H^1_0(\Omega) \setminus \{0\}} \sup_{v \in H^1_0(\Omega) \setminus \{0\}} \frac{A(u, v)}{|u||_{E, \mathcal{T}} + |au||_{E, \mathcal{T}}} \geq \frac{1}{3}.$$

Next, we split the discontinuous Galerkin form $A_{hp}$ into two parts, see Schötzau and Zhu (2009), and define

$$\tilde{A}_{hp}(u, v) = \sum_{K \in \mathcal{F}} \int_K (\varepsilon \nabla u \cdot \nabla v + a \cdot \nabla u v) \, dx + \sum_{E \in \delta(\mathcal{F})} \frac{\varepsilon y^2 h_E}{h_E} \int_E [u] \cdot [v] \, ds$$

$$- \sum_{K \in \mathcal{F}} \int_{\partial K \cap \Gamma_m} a \cdot n_K u v \, ds + \sum_{K \in \mathcal{F}} \int_{\partial K \cap \Gamma_p} a \cdot n_K (u^p - u) v \, ds,$$

$$K_{hp}(u, v) = - \sum_{E \in \delta(\mathcal{F})} \int_E \{\varepsilon \nabla u\} \cdot [v] \, ds - \sum_{E \in \delta(\mathcal{F})} \int_E \{\varepsilon \nabla v\} \cdot [u] \, ds.$$  

We shall use the above auxiliary forms to express both the continuous form $A$ in (2.3) and the discontinuous Galerkin form $A_{hp}$ in (2.6). Indeed, we have

$$A(u, v) = \tilde{A}_{hp}(u, v), \quad u, v \in H^1_0(\Omega), \quad (4.1)$$

$$A_{hp}(u, v) = \tilde{A}_{hp}(u, v) + K_{hp}(u, v), \quad u, v \in S_p(\mathcal{F}). \quad (4.2)$$

4.2 Auxiliary meshes

We shall make use of an auxiliary 1-irregular mesh $\tilde{\mathcal{F}}$ of parallelograms, similarly to the approach of Houston et al. (2008), which is obtained from $\mathcal{F}$ as follows. Let $K \in \mathcal{F}$. If all four elemental edges are edges of the mesh $\mathcal{F}$, that is, if $\delta(K) \subseteq \delta(\mathcal{F})$, we leave $K$ untouched. Otherwise, at least one of the elemental edges of $K$ contains a hanging node. In this case, we replace $K$ by the four parallelograms obtained from bisecting the elemental edges of $K$. This construction is illustrated in Figure 1. Clearly, the mesh $\tilde{\mathcal{F}}$ is a refinement of $\mathcal{F}$; it is also shape-regular and 1-irregular. We denote by $\delta_R(\tilde{\mathcal{F}})$ the set of edges in $\delta(\tilde{\mathcal{F}})$ that have been refined in the above process. We denote by $\mathcal{A}(\tilde{\mathcal{F}})$ respectively $\mathcal{A}_A(\tilde{\mathcal{F}})$ the vertices in $\mathcal{N}(\tilde{\mathcal{F}})$ respectively the edges in $\delta(\tilde{\mathcal{F}})$ which are inside an element $K$ of $\mathcal{F}$. Moreover, we write $\mathcal{A}(K)$ for the elements in $\tilde{\mathcal{F}}$ that are inside $K$. If $K$ is unrefined, $\mathcal{A}(K) = \{K\}$. Otherwise, the set $\mathcal{A}(K)$ consists of four newly created elements.

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

$$S_p(\tilde{\mathcal{F}}) = \{ v \in L^2(\Omega) : v|_K \circ F_K \in D_{p_K}(\hat{K}), \hat{K} \in \tilde{\mathcal{F}} \},$$

where the auxiliary polynomial degree vector $\hat{p}$ is defined by $p_K = p_K$, for $\hat{K} \in \mathcal{A}(K)$. Thus, we clearly have the inclusion $S_p(\mathcal{F}) \subseteq S_p(\tilde{\mathcal{F}})$. In complete analogy to (3.1) and (3.2), the energy and convective
norms associated with the auxiliary mesh $\mathcal{F}$ are given by

$$
\|v\|^2_{E, \mathcal{F}} = \sum_{K \in \mathcal{F}} \epsilon \|\nabla v\|^2_{L^2(K)} + \text{ejump}_{p, \mathcal{F}}(v)^2,
$$

$$
|v|^2_{O, \mathcal{F}} = \|v\|^2 + \text{ojump}_{p, \mathcal{F}}(v)^2,
$$

(4.3)

where the auxiliary edge polynomial degrees $p_e$ for the jump terms over $\mathcal{F}$ are defined as in (2.4), using the auxiliary degrees $p_K$. Obviously, $\|v\|_{E, \mathcal{F}} = \|v\|_{E, \mathcal{F}}$ and $|v|_{O, \mathcal{F}} = |v|_{O, \mathcal{F}}$ for all $v \in H^1_0(\Omega)$.

**Lemma 4.2** Let $v \in S_p(\mathcal{T}) + H^1_0(\Omega)$ and $w \in S_p(\mathcal{T}) + H^1_0(\Omega)$ be such that $\|v\|_E = \|w\|_E$, $E \in \mathcal{E}(\mathcal{T})$. Then we have

$$
\text{ejump}_{p, \mathcal{F}}(w) \lesssim \text{ejump}_{p, \mathcal{F}}(v) \lesssim \text{ejump}_{p, \mathcal{F}}(w),
$$

$$
\text{ojump}_{p, \mathcal{F}}(w) \lesssim \text{ojump}_{p, \mathcal{F}}(v) \lesssim \text{ojump}_{p, \mathcal{F}}(w).
$$

**Proof.** Since $w \in S_p(\mathcal{T}) + H^1_0(\Omega)$, we have that $\|v\|_E = \|w\|_E = 0$ over newly created edges in $\mathcal{E}_r(\mathcal{T})$. To look at the jumps over refined edges, let $E \in \mathcal{E}_r(\mathcal{T})$. We have $E = E_1 \cup E_2$ with $E_1$ and $E_2$ in $\mathcal{E}(\mathcal{T})$ and $h_{E_1} = h_{E_2} = h/2$. Thus

$$
\frac{h_{E_1}}{\varepsilon p_{E_1}} \|v\|^2_{L^2(E_1)} + \frac{h_{E_2}}{\varepsilon p_{E_2}} \|v\|^2_{L^2(E_2)} = \frac{1}{2} \frac{h_E}{\varepsilon p_E} \|v\|^2_{L^2(E)}.
$$

We conclude that

$$
\text{ojump}_{p, \mathcal{F}}(v)^2 = \sum_{E \in \mathcal{E}(\mathcal{T}) \setminus \mathcal{E}(\mathcal{F})} \frac{h_E}{\varepsilon p_E} \|v\|^2_{L^2(E)} + \sum_{E \in \mathcal{E}_r(\mathcal{T})} \frac{1}{2} \frac{h_E}{\varepsilon p_E} \|v\|^2_{L^2(E)}.
$$

This readily implies the desired conclusion for the convective jumps. The equivalence for the diffusive jumps follows completely analogously. □

As a consequence of this result, we also have the following estimate.

**Lemma 4.3** For $v \in S_p(\mathcal{T}) + H^1_0(\Omega)$, we have the bounds

$$
\|v\|_{E, \mathcal{F}} \lesssim \|v\|_{E, \mathcal{F}}, \quad |v|_{O, \mathcal{F}} \lesssim |v|_{O, \mathcal{F}}.
$$

**Proof.** Clearly, we have

$$
\sum_{K \in \mathcal{F}} \epsilon \|\nabla v\|^2_{L^2(K)} = \sum_{K \in \mathcal{F}} \epsilon \|\nabla v\|^2_{L^2(K)}.$$
be the conforming subspace of \( S \) Karakashian and Pascal (2003) and Ern and Stephansen (2008). To define this operator, we let \( S \) for (2008) or Burman and Ern (2007). For the \( h \) functions by continuous ones, analogously to the one used by Houston et al. (2007), Houston et al. Our analysis is based on an \( hp \)-version averaging operator that allows us to approximate discontinuous functions by continuous ones, analogously to the one used by Houston et al. (2007), Houston et al. (2008) and Schötzau and Zhu (2009), we decompose the DG solution into a variable polynomial degrees.

Following Houston et al. (2007) and Schötzau and Zhu (2009), we decompose the DG solution into a conforming part and a remainder:

\[
\bar{u}_{hp} = u_{hp}^c + u_{hp}^r,
\]

where \( u_{hp}^c = I_{hp}u_{hp} \in S_h^{e} (\mathcal{T}) \subset H_0^1 (\Omega) \), with \( I_{hp} \) the approximation operator from Theorem 4.1. The rest is then given by \( u_{hp}^r = u_{hp} - u_{hp}^c = u_{hp} - I_{hp}u_{hp} \in S_h^{e} (\mathcal{T}) \). By Lemma 4.3 and the triangle inequality, we obtain

\[
\| u - u_{hp} \|_{E,F} + | u - u_{hp} |_{0,F} \lesssim \| u - u_{hp} \|_{E,F} + \| u_{hp} - \| u_{hp}^c \|_{E,F} + \| u_{hp}^r \|_{0,F} \]

\[
\lesssim \| u - u_{hp}^c \|_{E,F} + | u - u_{hp}^c |_{0,F} + \| u_{hp}^r \|_{E,F} + \| u_{hp}^r \|_{0,F} \]

\[
= \| u - u_{hp}^c \|_{E,F} + | u - u_{hp}^c |_{0,F} + \| u_{hp}^r \|_{E,F} + \| u_{hp}^r \|_{0,F}.
\]

In a series of lemmas, we now bound both the continuous error \( u - u_{hp}^c \) and the rest term \( u_{hp}^r \) can be bounded by the estimator \( \eta \) and the data approximation term \( \Theta \).

**LEMMA 4.4** There holds

\[
\| u_{hp}^c \|_{E,F} + \| u_{hp}^r \|_{0,F} \lesssim \eta.
\]
Proof. Since \( \| u^p_h \|_E = \| u_h \|_E \) for all \( E \in \mathcal{E}( \mathcal{T} ) \) and \( u_h \in \mathcal{S}_p( \mathcal{T} ) \), Lemma 4.2 and the definition of the jump residual \( \eta_h \) yield
\[
\| u^p_h \|^2_{E, \mathcal{T}} + | u^p_h |^2_{\Omega, \mathcal{T}} = \sum_{K \in \mathcal{T}} \varepsilon \| \nabla u^p_h \|^2_{L^2(K)} + | aw^p_h |^2 + \text{ejump}_{p \mathcal{T}}( u^p_h )^2 + \text{ojump}_{p \mathcal{T}}( u^p_h )^2 \lesssim \sum_{K \in \mathcal{T}} \varepsilon \| \nabla u_h \|^2_{L^2(K)} + | aw_h |^2 + \sum_{K \in \mathcal{T}} \eta^2_h .
\]
Hence, only the volume terms and \( | aw^p_h | \) need to be bounded further. Since \( p_E \geq 1 \), Theorem 4.1 yields
\[
\varepsilon \sum_{K \in \mathcal{T}} \| \nabla u^p_h \|^2_{L^2(K)} \lesssim \gamma^{-1} \sum_{E \in \mathcal{E}( \mathcal{T} )} \frac{p_E^2}{h_E} \| u_h \|_{L^2(E)}^2 \lesssim \gamma^{-1} \sum_{K \in \mathcal{T}} \eta^2_h .
\]
To estimate \( | aw^p_h | \), we again use Theorem 4.1 and the fact that \( p_E \geq 1 \),
\[
| aw^p_h |^2 \lesssim \frac{1}{\varepsilon} \sum_{K \in \mathcal{T}} \left( | a |_{L^\infty(K)}^2 \| u^p_h \|^2_{L^2(K)} \right) \lesssim \sum_{E \in \mathcal{E}( \mathcal{T} )} \frac{h_E}{p_E^2} \| u_h \|_{L^2(E)}^2 \lesssim \sum_{K \in \mathcal{T}} \eta^2_h .
\]
This finishes the proof. \( \square \)

Next, we recall the following standard \( hp \)-version approximation result from, e.g., Lemma 3.7 of Houston et al. (2008): For any \( v \in H^1_0(\Omega) \), there exists a function \( v_h \in \mathcal{S}_p( \mathcal{T} ) \) such that
\[
\frac{p_K^2}{h_K} \| v - v_h \|^2_{L^2(K)} + \| \nabla (v - v_h) \|^2_{L^2(\partial K)} + \frac{p_K}{h_K} \| v - v_h \|^2_{L^2(\partial K)} \lesssim \| \nabla v \|^2_{L^2(K)},
\]
for any \( K \in \mathcal{T} \).

**Lemma 4.5** For any \( v \in H^1_0(\Omega) \), we have
\[
\int_\Omega f(v - v_h) \, dx - \tilde{A}_h(p(u_h, v - v_h)) + K_h(p(u_h, v_h)) \lesssim (\eta + \Theta) \| v \|_{E, \mathcal{T}} .
\]
Here, \( v_h \in \mathcal{S}_p( \mathcal{T} ) \) is the \( hp \)-interpolant of \( v \) in (4.8).

**Proof.** Integration by parts of the diffusive volume terms readily yields
\[
\int_\Omega f(v - v_h) \, dx - \tilde{A}_h(p(u_h, v - v_h)) + K_h(p(u_h, v_h)) = T_1 + T_2 + T_3 + T_4 + T_5,
\]
where
\[
T_1 = \sum_{K \in \mathcal{T}} \int_K (f + \varepsilon \Delta u_h - a \cdot \nabla u_h)(v - v_h) \, dx,
\]
\[
T_2 = - \sum_{E \in \mathcal{E}( \mathcal{T} )} \int_E [\varepsilon \nabla u_h] \cdot [v - v_h] \, ds,
\]
\[
T_3 = - \sum_{E \in \mathcal{E}( \mathcal{T} )} \int_E [\varepsilon \nabla v_h] \cdot [u_h] \, ds,
\]
\[
T_4 = \sum_{K \in \mathcal{T}} \int \alpha \cdot n_K(u_h - u_h^e)(v - v_h) \, ds + \sum_{K \in \mathcal{T}} \int \alpha \cdot n_K(u_h)(v - v_h) \, ds,
\]
\[
T_5 = - \sum_{E \in \mathcal{E}( \mathcal{T} )} \frac{\varepsilon \gamma p_E^2}{h_E} \int_E [u_h] \cdot [v - v_h] \, ds.
\]
To bound $T_1$, we first add and subtract the data approximations. From the weighted Cauchy-Schwarz inequality and the approximation properties in (4.8), we then readily obtain

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

Similarly, by the Cauchy-Schwarz inequality and (4.8), we have

$$T_2 \lesssim \left( \sum_{E \in \mathcal{E}(\mathcal{T})} e^{-1} p_E^{-1} h_E \left\| [e \nabla u_{hp}] \right\|_{L^2(E)}^2 \right)^{\frac{1}{2}} \left( \sum_{E \in \mathcal{E}(\mathcal{T})} e^{p_E} h_E^{-1} \left\| v - v_{hp} \right\|_{L^2(E)}^2 \right)^{\frac{1}{2}} \lesssim \left( \sum_{K \in \mathcal{T}} \eta_{K}^2 \right)^{\frac{1}{2}} \|v\|_{E, \mathcal{T}}.$$  

To estimate $T_3$, we employ the Cauchy-Schwarz inequality, the $hp$-version trace inequality and the $H^1$-stability of $v_{hp}$ from (4.8). This results in

$$T_3 \lesssim \left( \sum_{K \in \mathcal{T}} \epsilon p_E h_E^{-1} \left\| [u_{hp}] \right\|_{L^2(K)}^2 \right)^{\frac{1}{2}} \left( \sum_{K \in \mathcal{T}} \epsilon \left\| \nabla v_{hp} \right\|_{L^2(K)}^2 \right)^{\frac{1}{2}} \lesssim \left( \sum_{K \in \mathcal{T}} \eta_{K}^2 \right)^{\frac{1}{2}} \|v\|_{E, \mathcal{T}}.$$  

For $T_4$, we apply again the Cauchy-Schwarz inequality and (4.8) to get

$$T_4 \lesssim \left( \sum_{E \in \mathcal{E}(\mathcal{T})} e^{-1} p_E^{-1} h_E \left\| [u_{hp}] \right\|_{L^2(E)}^2 \right)^{\frac{1}{2}} \left( \sum_{E \in \mathcal{E}(\mathcal{T})} e^{p_E} h_E^{-1} \left\| v - v_{hp} \right\|_{L^2(E)}^2 \right)^{\frac{1}{2}} \lesssim \left( \sum_{K \in \mathcal{T}} \eta_{K}^2 \right)^{\frac{1}{2}} \|v\|_{E, \mathcal{T}}.$$  

Finally, we have

$$T_5 \lesssim \left( \sum_{E \in \mathcal{E}(\mathcal{T})} \epsilon^2 p_E^2 h_E^{-1} \left\| [u_{hp}] \right\|_{L^2(E)}^2 \right)^{\frac{1}{2}} \left( \sum_{E \in \mathcal{E}(\mathcal{T})} e^{p_E} h_E^{-1} \left\| v - v_{hp} \right\|_{L^2(E)}^2 \right)^{\frac{1}{2}} \lesssim \left( \sum_{K \in \mathcal{T}} \eta_{K}^2 \right)^{\frac{1}{2}} \|v\|_{E, \mathcal{T}}.$$  

The above estimates for the terms $T_1$ through $T_5$ imply the assertion.}

**Lemma 4.6** There holds:

$$\| u - u_{hp} \|_{E, \mathcal{T}} + \| u - u_{hp} \|_{0, \mathcal{T}} \lesssim \eta + \Theta.$$  

**Proof.** Since $u - u_{hp} \in H^1_0(\Omega)$, we have $\| u - u_{hp} \|_{0, \mathcal{T}} = |a(u - u_{hp})|$. Then the inf-sup condition in Lemma 4.1 yields:

$$\| u - u_{hp} \|_{E, \mathcal{T}} + \| u - u_{hp} \|_{0, \mathcal{T}} \lesssim \sup_{v \in H^1_0(\Omega) \setminus \{0\}} \frac{A(u - u_{hp}, v)}{\| v \|_{E, \mathcal{T}}}. \quad (4.9)$$  

To bound (4.9), let $v \in H^1_0(\Omega)$. Then, property (4.1) shows that

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

By employing the fact that $v \in H^1_0(\Omega)$ and integrating by parts the convection term, one can readily see that

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

where $R$ is a remainder term.
with
\[ R = \sum_{K \in \mathcal{T}} \int_K (-\varepsilon \nabla u_h \cdot \alpha u_h + u_h^2) \cdot \nabla v \, dx. \]

Furthermore, from the DG method in (2.6) and property (4.2), we have
\[ \int_{\Omega} f v_h \, dx = A_{hp}(u_h, v_h) + K_{hp}(u_h, v_h), \]
where \( v_h \in S_p(\mathcal{T}) \) is the \( hp \)-version. Combining the above results yields
\[ A(u - u_h^*, v) = \int_{\Omega} f(v - v_h) \, dx - \tilde{A}_{hp}(u_h, v - v_h) + K_{hp}(u_h, v_h) - R. \]

The estimate in Lemma 4.5 now shows that
\[ |A(u - u_h^*, v)| \lesssim (\eta + \Theta) \| v \|_{E, \mathcal{T}} + |R|. \]  \hspace{1cm} (4.10)

It remains to bound \( |R| \). From the Cauchy-Schwarz inequality, the definition of the norm \( |\cdot|_* \), the conformity of \( v \) and Lemma 4.4, one readily obtains
\[ |R| \lesssim \left( \| u_h^* \|_{E, \mathcal{T}} + |u_h^*|_{O, \mathcal{T}} \right) \| v \|_{E, \mathcal{T}} \lesssim \eta \| v \|_{E, \mathcal{T}}. \]  \hspace{1cm} (4.11)

Equations (4.9)–(4.11) imply the desired result. \( \square \)

The proof of Theorem 3.1 now immediately follows from (4.7), Lemma 4.4 and Lemma 4.6.

4.5 Proof of Theorem 3.2

We first introduce the following bubble functions; cf. Melenk and Wohlmuth (2001). On the reference element \( \tilde{K} = (-1, 1)^2 \), we define the weight function \( \Psi^*_K(\tilde{x}) = \text{dist}(\tilde{x}, \partial \tilde{K}) \). For an arbitrary element \( K \in \mathcal{T} \), we set \( \Psi_K = c_K \Psi^*_K \circ F_{\tilde{K}}^{-1} \), where \( c_K \) is a scaling factor chosen such that \( \int_K (\Psi_K - 1) \, dx = 0 \).

Similarly, on the reference interval \( \tilde{I} = (-1, 1) \), we define the weight function \( \Psi^*_I(\tilde{x}) = 1 - \tilde{x}^2 \). For an interior edge \( E \), we let \( \Psi_E = c_E \Psi^*_I \circ F_{\tilde{E}}^{-1} \), where \( F_{\tilde{E}} \) is the affine transformation that maps \( \tilde{I} \) onto \( E \) and \( c_E \) is chosen such that \( \int_E (\Psi_E - 1) \, ds = 0 \).

We now show the efficiency of \( \eta_{E_k}, \eta_{E_k} \) and \( \eta_{I_k} \), respectively.

**Lemma 4.7** Under the assumptions of Theorem 3.2, there holds
\[ \left( \sum_{K \in \mathcal{T}} \eta_{E_k}^2 \right)^{1/2} \lesssim \| p \|_{E, \mathcal{T}} + |p|_{O, \mathcal{T}} |a(u - u_h)| + |p|^{\delta + 1/2}. \]

**Proof.** For any element \( K \in \mathcal{T} \), we set
\[ v_K = \epsilon^{-1} (f_h + \epsilon \Delta u_h - a_h \cdot \nabla u_h)|_{\mathcal{K}}^\alpha, \]
where \( \alpha \in (1/2, 1] \). Applying the inverse inequality from Theorem 2.5 of Melenk and Wohlmuth (2001), we obtain
\[ \| f_h + \epsilon \Delta u_h - a_h \cdot \nabla u_h \|_{L^2(K)} \lesssim \rho_K^\alpha \| (f_h + \epsilon \Delta u_h - a_h \cdot \nabla u_h)\Psi^\alpha_K \|_{L^2(K)} = \epsilon \rho_K^\alpha \| v_K \Psi^\alpha_K \|_{L^2(K)}. \]
This leads to
\[
\sum_{K \in \mathcal{T}} \eta_{K}^2 \lesssim S^2 \quad \text{with} \quad S^2 = \sum_{K \in \mathcal{T}} p_K^{2-2\alpha} h_K^2 E \| v_K \psi_{K}^{-\alpha/2} \|^2_{L^2(K)}.
\]

Since the exact solution satisfies \((f + \varepsilon \Delta u - a \cdot \nabla u)|_K = 0\), we obtain, by integration by parts and insertion of the data \(a\) and \(f\),
\[
S^2 = \sum_{K \in \mathcal{T}} p_K^{2-2\alpha} h_K^2 \int_K (f_{hp} + \varepsilon \Delta u_{hp} - \omega_{hp} \cdot \nabla u_{hp}) v_K \, dx \\
= \sum_{K \in \mathcal{T}} p_K^{2-2\alpha} h_K^2 \int_K (\varepsilon \nabla (u - u_{hp}) - a(u - u_{hp})) \cdot \nabla v_K \, dx \\
+ \sum_{K \in \mathcal{T}} p_K^{2-2\alpha} h_K^2 \int_K (| f_{hp} - f | + | a - \omega_{hp} | \cdot \nabla u_{hp}) e^{-1/2} \psi_{K}^{\alpha/2} | e^{1/2} v_K \psi_{K}^{-\alpha/2} | \, dx.
\]

Here, we have also used that \(v_K|_{\partial K} = 0\). From Lemma 3.4 of Melenk and Wohlmuth (2001), we have
\[
\| \nabla v_K \|_{L^2(K)} \lesssim h_K^{-1} p_K^{2-\alpha} \| v_K \psi_{K}^{-\alpha/2} \|_{L^2(K)}.
\]

By the Cauchy-Schwarz inequality, the definition of the dual norm and the data approximation error \(\Theta\), we obtain
\[
S^2 \lesssim S (\| p \| u - u_{hp} \|_{E, \mathcal{T}} + \| a \| (u - u_{hp}) \|_\ast + | p |^\alpha \Theta).
\]

Therefore,
\[
\left( \sum_{K \in \mathcal{T}} \eta_{K}^2 \right)^{1/2} \lesssim | p | \| u - u_{hp} \|_{E, \mathcal{T}} + | p | | a (u - u_{hp}) | \|_\ast + | p |^\alpha \Theta.
\]

Choosing \(\delta = \alpha - 1/2\) finishes the proof.

For any edge \(E \in \mathcal{E}(\mathcal{T})\), we define the sets
\[
w_E = \{ K_1, K_2 \in \mathcal{T} : E = \partial K_1 \cap \partial K_2 \}, \quad \overline{w}_E = \{ \overline{K} \in \mathcal{T} \cup \mathcal{F} : E \in \mathcal{E}(\overline{K}) \}.
\]

For simplicity, we also use the notation \(w_E\) and \(\overline{w}_E\) to denote the domain formed by the elements in \(w_E\) and in \(\overline{w}_E\), respectively.

**Lemma 4.8** Under the assumptions of Theorem 3.2, there holds
\[
\left( \sum_{K \in \mathcal{T}} \eta_{K}^2 \right)^{1/2} \lesssim | p |^{\delta+1} \| u - u_{hp} \|_{E, \mathcal{T}} + | p |^{2\delta+1} | u - u_{hp} |_{O, \mathcal{T}} + | p |^{2\delta+1} \Theta.
\]

**Proof.** Let \(E = \partial K_1 \cap \partial K_2\) be an interior edge shared by two elements \(K_1, K_2 \in \mathcal{T}\). For \(\alpha \in (1/2, 1]\), we construct a bubble function \(\psi_E\) over \(w_E\).

**Case I:** Suppose that none of the end points of \(E\) is a hanging node. That is, \(E \in \mathcal{E}(K_1) \cap \mathcal{E}(K_2)\). Lemma 2.6 of Melenk and Wohlmuth (2001) then ensures the existence of a function \(\psi_E \in H^1_0(w_E)\) with \(\psi_E \mid E = \tau_E\). We construct a bubble function \(\psi_E\) over \(w_E\).

\[
\| \psi_E \|_{L^2(w_E)} \lesssim h_E^{1/2} p_E^{-1} \| \tau_E \psi_E^{-\alpha/2} \|_{L^2(E)},
\]

\[
\| \nabla \psi_E \|_{L^2(w_E)} \lesssim h_E^{-1/2} p_E^2 \| \tau_E \psi_E^{-\alpha/2} \|_{L^2(E)}.
\]
That Melenk and Wohlmuth (2001), we get

Now define the function

\[
\tilde{\psi}(\tilde{\theta}) = \begin{cases} 0, & \tilde{\theta} \lessdot 1 \\ \left(1-\tilde{\theta}\right)^{\alpha-1}, & 1 \lessdot \tilde{\theta} \lessdot 2 \end{cases}
\]

such that

\[
\tilde{\psi}(\tilde{\theta}) \in H^1_0((\tilde{\theta})) \quad \text{with} \quad \tilde{\psi}_{|\partial\tilde{\theta}} = 0.
\]

In both cases above, we now proceed as follows. From the inverse inequality in Lemma 2.4 of Melenk and Wohlmuth (2001), we get

\[
\left\| \mathbf{\nabla} u_{hp} \right\|_{L^2(E)} \lesssim p_E^\alpha \left\| \mathbf{\nabla} u_{hp} \right\|_{H^\alpha(E)} - \frac{1}{2} \left\| \mathbf{\nabla} u_{hp} \right\|_{L^2(E)}.
\]

Therefore,

\[
\sum_{E \in \mathcal{T}} \eta_E^2 \lesssim S^2 \quad \text{with} \quad S^2 = \sum_{E \in \mathcal{T}} p_E^\alpha h_E^{2-\alpha} \left\| \mathbf{\nabla} \tilde{u}_{hp} \right\|_{L^2(E)}^2.
\]

Since \( \| \mathbf{\epsilon} \mathbf{\nabla} u \| = 0 \) on interior edges, integration by parts over \( \tilde{\tilde{E}} \) yields

\[
\int_E \left( \mathbf{\epsilon} (u_{hp} - u) \right) \cdot \mathbf{\nabla} \psi_E \, dx = \int_{\tilde{\tilde{E}}} \mathbf{\epsilon} (\Delta u_{hp} - \Delta u) \psi_E + \mathbf{\epsilon} (\mathbf{\nabla} u_{hp} - \mathbf{\nabla} u) \cdot \mathbf{\nabla} \psi_E \, dx,
\]

where \( \Delta u_{hp} \) and \( \mathbf{\nabla} u_{hp} \) are understood piecewise. Using the differential equation, approximating the data and integrating by parts the convective term, we readily obtain

\[
S^2 = T_1 + T_2 + T_3 + T_4 + T_5,
\]

with

\[
T_1 = \sum_{E \in \mathcal{T}} p_E^\alpha h_E^{2-\alpha} \left( f_{hp} + \mathbf{\epsilon} \Delta u_{hp} - \mathbf{\epsilon} \mathbf{\nabla} u_{hp} \right) \cdot \mathbf{\nabla} \psi_E \, dx,
\]

\[
T_2 = \sum_{E \in \mathcal{T}} p_E^\alpha h_E^{2-\alpha} \left( \mathbf{\epsilon} \mathbf{\nabla} u_{hp} - \mathbf{\epsilon} \mathbf{\nabla} u \right) \cdot \mathbf{\nabla} \psi_E \, dx,
\]

\[
T_3 = \sum_{E \in \mathcal{T}} p_E^\alpha h_E^{2-\alpha} \mathbf{a}(u - u_{hp}) \cdot \mathbf{\nabla} \psi_E \, dx,
\]

\[
T_4 = \sum_{E \in \mathcal{T}} p_E^\alpha h_E^{2-\alpha} \mathbf{a} \cdot \left[ \mathbf{\nabla} u_{hp} \right] \psi_E \, ds,
\]

\[
T_5 = \sum_{E \in \mathcal{T}} p_E^\alpha h_E^{2-\alpha} \left( f - f_{hp} + \mathbf{a}_{hp} - \mathbf{a} \right) \cdot \mathbf{\nabla} u_{hp} \psi_E \, dx.
\]

The Cauchy-Schwarz inequality, Lemma 4.7 and inequality (4.12) yield

\[
T_1 \lesssim S \left| \mathbf{p} \right|^{\alpha-\frac{1}{2}} \left( \left| \mathbf{p} \right| \left[ \mathbf{\nabla} u_{hp} \right]_{E, \mathcal{T}} + \left| \mathbf{p} \right| \left[ \mathbf{a} \right]_{u_{hp}} \right) + \left| \mathbf{p} \right|^{\alpha} \Theta.
\]
Similarly, we obtain
\[ T_2 \lesssim S \| p \|^{1+\alpha} \| u - u_{hp} \|_{E,\mathcal{T}}, \]
as well as
\[ T_3 \lesssim S \| p \|^{1+\alpha} |a(u - u_{hp})|. \]
To bound \( T_4 \), we first notice that \( \| \Psi_E \|_{L^\infty(E)} = c_E = \frac{3}{2} \). By (4.14) and the definition of semi-norm \( \cdot |_{O,\mathcal{T}} \), we conclude that
\[ T_4 \lesssim \sum_{E \in \mathcal{E}(\mathcal{T})} \left( e^{-\frac{1}{2}} h_E p_E^{-\frac{1}{2}} \| \| u_{hp} \| \|_{L^2(E)} e^{\frac{1}{2}} h_E p_E^{3\alpha - \frac{1}{2}} \| \nabla \Psi_E \|_{L^2(E)} \right) \lesssim S \| p \|^{2\alpha} |u - u_{hp}|_{O,\mathcal{T}}. \]
Finally, the data error term \( T_5 \) can be bounded by
\[ T_5 \lesssim \sum_{E \in \mathcal{E}(\mathcal{T})} \left( e^{-\frac{1}{2}} h_E p_E^{-\frac{1}{2}} \| \| u_{hp} \| \|_{L^2(E)} e^{\frac{1}{2}} h_E p_E^{3\alpha - \frac{1}{2}} \| \nabla \Psi_E \|_{L^2(E)} \right) \lesssim S \| p \|^{2\alpha} |u - u_{hp}|_{O,\mathcal{T}}. \]
Combining the above bounds for \( T_1 \) through \( T_5 \), we obtain
\[ S^2 \lesssim S \left( \| p \|^{1+\frac{1}{2}} \| u - u_{hp} \|_{E,\mathcal{T}} + \| p \|^{2\alpha} |u - u_{hp}|_{O,\mathcal{T}} + \| p \|^{2\alpha - \frac{1}{2}} \Theta \right). \]
Thus,
\[ \left( \sum_{E \in \mathcal{E}(\mathcal{T})} \eta^2_E \right)^{\frac{1}{2}} \lesssim \| p \|^{1+\frac{1}{2}} \| u - u_{hp} \|_{E,\mathcal{T}} + \| p \|^{2\alpha} |u - u_{hp}|_{O,\mathcal{T}} + \| p \|^{2\alpha - \frac{1}{2}} \Theta. \]
Choosing \( \delta = \alpha - 1/2 \) implies the assertion.

Since the jumps of \( u \) vanish over the edges, we finally have the following result.

**Lemma 4.9** Under the assumptions of Theorem 3.2, there holds
\[ \left( \sum_{E \in \mathcal{E}(\mathcal{T})} \eta^2_E \right)^{1/2} \lesssim \| p \|^{1/2} \| u - u_{hp} \|_{E,\mathcal{T}} + |u - u_{hp}|_{O,\mathcal{T}}. \]

The proof of Theorem 3.2 now follows from Lemmas 4.7, 4.8 and 4.9.

### 5. Proof of Theorem 4.1

In this section, we prove the result of Theorem 4.1.

#### 5.1 Polynomial basis functions

We begin by introducing \( hp \)-version basis functions. To that end, let \( \tilde{I} = (-1, 1) \) be the reference interval. We denote by \( \mathcal{Z}^p = \{ z_0^p, \cdots, z_p^p \} \) the Gauss-Lobatto nodes of order \( p \geq 1 \) on \( \tilde{I} \). Recall that \( z_0^p = -1 \) and \( z_p^p = 1 \). We denote by \( \mathcal{Z}^p_{\text{int}} = \{ z_1^p, \cdots, z_{p-1}^p \} \) the interior Gauss-Lobatto nodes of order \( p \) on \( \tilde{I} \). Let now \( E \in \mathcal{E}(K) \) be an edge of an element \( K \). The nodes in \( \mathcal{Z}^p \) can be affinely mapped onto \( E \) and we denote by \( \mathcal{Z}^p_E = \{ z_0^E, \cdots, z_p^E \} \) the Gauss-Lobatto nodes of order \( p \) on \( E \). The points \( z_0^E \) and \( z_p^E \)
coincide with the two end points of $E$. The set \( \mathcal{P}_p^E (E) = \{ z_0^E, \cdots, z_{p-1}^E \} \) denotes the interior Gauss-Lobatto points of order $p$. We write $\mathcal{P}_p(E)$ for the space of all polynomials of degree less or equal than $p$ on $E$ and define

\[
\mathcal{P}_p^\text{int}(E) = \{ q \in \mathcal{P}_p(E) : q(z_0^E) = q(z_{p-1}^E) = 0 \},
\]

\[
\mathcal{P}_p^\text{mod}(E) = \{ q \in \mathcal{P}_p(E) : q(0) = 0, z \in \mathcal{P}_p^E (E) \}.
\]

By construction, we have $\mathcal{P}_p(E) = \mathcal{P}_p^\text{int}(E) \oplus \mathcal{P}_p^\text{mod}(E)$.

For an element $K$ and $p \geq 1$, we now define basis functions for polynomials of the form

\[
v \in \mathcal{P}_p(K), \quad v|_E \in \mathcal{P}_p^E (E), \quad E \in \mathcal{E}(K),
\]

where $1 \leq p_E \leq p$ is the edge polynomial degree associated with $E \in \mathcal{E}(K)$. As usual, we shall divide the basis functions into interior, edge and vertex basis functions.

We first consider the reference element $\hat{K} = (-1, 1)^2$. We denote its four edges by $\hat{E}_1, \cdots, \hat{E}_4$ and its four vertices by $\hat{v}_1, \cdots, \hat{v}_4$, numbered as in Figure 2. Let $\{ \hat{\Phi}_i^p \} _{0 \leq i \leq p}$ be the Lagrange basis functions associated with the nodes $\hat{\mathcal{P}}^p$. We denote by $\{ \hat{z}_{i,j}^p = (\hat{z}_i^p, \hat{z}_j^p) \}_{1 \leq i,j \leq p}$ the interior tensor-product Gauss-Lobatto nodes on $\hat{K}$. The interior basis functions are then given by

\[
\hat{\Phi}_{i,j}^{\text{int},p} (\hat{x}_1, \hat{x}_2) = \hat{\Phi}_i^p (\hat{x}_1) \hat{\Phi}_j^p (\hat{x}_2), \quad 1 \leq i, j \leq p - 1.
\]

Next, we consider exemplarily the edge $\hat{E}_1$ in Figure 2 with edge degree $p_{\hat{E}_1}$. The edge basis functions for $\hat{E}_1$ are

\[
\hat{\Phi}_{i}^{\hat{E}_1,p_{\hat{E}_1}} (\hat{x}_1, \hat{x}_2) = \hat{\Phi}_i^{p_{\hat{E}_1}} (\hat{x}_1) \hat{\Phi}_0^p (\hat{x}_2), \quad i = 1, \cdots, p_{\hat{E}_1} - 1.
\]

Note that $\hat{\Phi}_{i}^{\hat{E}_1,p_{\hat{E}_1}}$ vanishes on $\hat{E}_2$, $\hat{E}_3$ and $\hat{E}_4$. The other edge basis functions are defined analogously. Finally, we consider the vertex $\hat{v}_1$, which is shared by $\hat{E}_1$ and $\hat{E}_4$; see Figure 2. We then introduce the associated vertex basis function

\[
\hat{\Phi}_{\hat{v}_1}^p (\hat{x}_1, \hat{x}_2) = \hat{\Phi}_0^p (\hat{x}_1) \hat{\Phi}_0^p (\hat{x}_2).
\]
It vanishes on $\widehat{E}_2$ and $\widehat{E}_3$. The vertex basis functions associated with the other vertices of $\widehat{K}$ are defined analogously. This completes the definition of the shape functions on the reference element $\widehat{K}$.

For an arbitrary parallelogram $K$, shape functions $\Phi$ on $K$ can be defined from the analogous ones on $\widehat{K}$ by using the pull-back map $\Phi = \Phi \circ F_K^{-1}$, giving rise to shape functions $\Phi_{\hat{v}}^\nu$, $\Phi_{\hat{v}}^{E,p}E$ and $\Phi_{i,j}^{\text{int},p}$ on $K$. Therefore, a polynomial $v$ of the form (5.1) can be expanded into

$$v(x) = \sum_{v \in \mathcal{A}(K)} v(v) \Phi_{\hat{v}}^\nu(x) + \sum_{E \in \mathcal{E}(K)} \sum_{i=1}^{p_E-1} v(z_{i}^{E,p}) \Phi_{i}^{E,p}(x) + \sum_{1 \leq i,j \leq p-1} c_{ij} \Phi_{i,j}^{\text{int},p}(x),$$

with expansion coefficients $c_{ij}$.

We will make use of the following estimates, see Lemma 3.1 of Burman and Ern (2007).

**Lemma 5.1** There holds:

(i) For a function $\hat{v} \in \mathcal{P}_p(\widehat{\mathcal{E}}^d)$, $d = 1,2$, that vanishes at the interior Gauss-Lobatto nodes on $\widehat{\mathcal{E}}^d$, there holds $\|\hat{v}\|^2_{L^2(\partial p)} \lesssim p^{-2} \|\hat{v}\|^2_{L^2(\partial p)}$.

(ii) If the vertex $\hat{v}$ of the reference element $\widehat{K}$ is shared by two edges $\widehat{E}_n$ and $\widehat{E}_m$, the associated vertex basis function $\Phi_{\hat{v}}^\nu$ can be bounded by $\|\Phi_{\hat{v}}^\nu\|_{L^2(\widehat{K})} \lesssim P_E^{-1} P_n^{-1}$.

### 5.2 Extension operators

Next, we define extension operators over edges. Let $\widehat{E} \in \mathcal{E}(\widehat{K})$ be an elemental edge of the reference element $\widehat{K}$. We define $\widehat{L}_p^E$ by

$$\widehat{L}_p^E : \mathcal{P}_p^\text{int}(\widehat{\mathcal{E}}) \rightarrow \mathcal{P}_p(\widehat{\mathcal{K}}), \quad \hat{q}(x) \mapsto \sum_{i=1}^{p-1} \hat{q}(z_{i}^{E,p}) \Phi_{1}^{E,p}(x). \quad (5.2)$$

By construction, $\widehat{L}_p^E(\hat{q}) = \hat{q}$ on $\widehat{E}$, and $\widehat{L}_p^E(\hat{q})$ vanishes in all the interior tensor-product Gauss-Lobatto nodes $\{z_{i}^{E,p}\}_{1 \leq i \leq p-1}$ of $\widehat{K}$ and on the other three edges of $\widehat{K}$. From Lemma 3.1 of Burman and Ern (2007), we have the following inequality.

**Lemma 5.2** The linear extension operator $\widehat{L}_p^E$ introduced in (5.2) satisfies

$$\|\widehat{L}_p^E(\hat{q})\|_{L^2(\widehat{\mathcal{K}})} \lesssim p^{-1} \|\hat{q}\|_{L^2(\widehat{\mathcal{E}})}.$$

Now consider an arbitrary element $K \in \mathcal{T}$ and fix an edge $E \in \mathcal{E}(K)$. If $E$ contains no hanging node in $\mathcal{T}$ (i.e., $E \in \mathcal{E}(\mathcal{T})$), we define the extension operator $L_{p,K}^E(q) : \mathcal{P}_p^\text{int}(E) \rightarrow \mathcal{P}_p(K)$ by

$$L_{p,K}^E(q) = [L_{p}(q \circ F_K)] \circ F_K^{-1}, \quad q \in \mathcal{P}_p^\text{int}(E). \quad (5.3)$$

If $E$ contains a hanging node located in the middle of it, $E$ can be written as $E = E_1 \cup E_2$ for two edges $E_1$ and $E_2$ in $\mathcal{E}(\mathcal{T})$. We then partition $K$ into two parallelograms, $K = K_1 \cup K_2$, by connecting the hanging node on $E$ with the midpoint of the edge opposite to $E$, as illustrated in Figure 3 (left). For any $q_1 \in \mathcal{P}_p^\text{int}(E_1)$ and $q_2 \in \mathcal{P}_p^\text{int}(E_2)$, we define the extension operator $L_{p,K}^E(q_1, q_2)$ by

$$L_{p,K}^E(q_1, q_2) = L_{p,K_1}^E(q_1) + L_{p,K_2}^E(q_2), \quad (5.4)$$
with $L^E_{p,K_1}$ and $L^E_{p,K_2}$ given in (5.3). By definition, the extensions $L^E_{p,K}(q)$ and $L^E_{p,K}(q_1,q_2)$ are continuous in $K$ and satisfy $L^E_{p,K}(q)|_E = q$, $L^E_{p,K}(q_1,q_2)|_{E_1} = q_1$ and $L^E_{p,K}(q_1,q_2)|_{E_2} = q_2$. Moreover, the extensions $L^E_{p,K}(q)$ and $L^E_{p,K}(q_1,q_2)$ both vanish on the other edges of $\partial(K)$.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure3.png}
\caption{Left: Partition of $K$ into $K_1$ and $K_2$. Right: Element $K$ and $\tilde{K} \in \mathcal{R}(K)$.}
\end{figure}

5.3 Decomposition of functions in $S_p(\mathcal{T})$

We shall now decompose functions in $S_p(\mathcal{T})$, similarly to Proposition 5.4 of Houston et al. (2007). For $E \in \mathcal{E}(\mathcal{T}) \cup \mathcal{E}(\mathcal{K})$, we set

$$p_E = \min\{ p_\tilde{K} : \tilde{K} \in \tilde{w}_E \}. \quad (5.5)$$

Notice that an elemental edge $E$ in $\mathcal{E}(\tilde{K})$, $\tilde{K} \in \tilde{\mathcal{T}}$, belongs to $\mathcal{E}(\mathcal{T}) \cup \mathcal{E}(\mathcal{K})$. Hence, for any $\tilde{K} \in \tilde{\mathcal{T}}$, equation (5.5) defines the elemental edge polynomial degrees as used in (5.1). Furthermore, we denote by $v_K$ the restriction of a piecewise smooth function $v$ to an element $K \in \mathcal{T} \cup \tilde{\mathcal{T}}$.

Let $v \in S_p(\mathcal{T})$. Firstly, we introduce a (nodal) interpolant $v^\text{nod} \in S_p(\tilde{\mathcal{T}})$. For each element $K \in \mathcal{T}$ and $\tilde{K} \in \mathcal{R}(K)$, we will construct the restriction $v^\text{nod}_K$ of $v^\text{nod}$ to $\tilde{K}$ such that

$$v^\text{nod}_K \in P_{p_K}(\tilde{K}), \quad v^\text{nod}_K|_E \in P_{p_E}(E), \quad E \in \mathcal{E}(\tilde{K}), \quad (5.6)$$

with $p_E$ given in (5.5). To define $v^\text{nod}_K$, we distinguish two cases.

\textbf{Case 1:} If $\mathcal{R}(K) = \{K\}$ (i.e., if $K$ is unrefined), the interpolant $v^\text{nod}_K = v^\text{nod}_K$ is simply defined by

$$v^\text{nod}_K(x) = \sum_{v \in V(K)} v_K(v) \Phi^v_K(x). \quad (5.7)$$

\textbf{Case 2:} If $\mathcal{R}(K)$ consists of four newly created elements, we define $v^\text{nod}_K$ on each element $\tilde{K} \in \mathcal{R}(K)$ separately. To do so, fix $\tilde{K} \in \mathcal{R}(K)$. Without loss of generality, we may consider the situation shown in Figure 3 (right), where $\{E_i\}_{i=1}^4$ and $\{v_i\}_{i=1}^4$ denote the edges and vertices of $K$, $\{\tilde{E}_i\}_{i=1}^4$ and $\{\tilde{v}_i\}_{i=1}^4$ the ones of $\tilde{K}$. Notice that here we have $\tilde{v}_2 = v_2$ and $\tilde{v}_4 \in \mathcal{N}_A(\tilde{\mathcal{T}})$. Furthermore, $\tilde{E}_3$ and $\tilde{E}_4$ are in $\mathcal{E}_A(\tilde{\mathcal{T}})$ and $\tilde{p}_{E_3} = \tilde{p}_{E_4} = p_{\tilde{K}} = p_K$. Let us now define the value of $v^\text{nod}_K$ at the edge and vertex nodes of $\tilde{K}$. At the interior nodes of $\tilde{E}_3$ and $\tilde{E}_4$, we set

$$v^\text{nod}_K(z) = v_K(z), \quad z \in \mathcal{P}_{\text{int}}(\tilde{E}_3) \cup \mathcal{P}_{\text{int}}(\tilde{E}_4). \quad (5.8)$$
Similarly, we set $v^{\text{nod}}_K(v) = v_K(v)$ for the vertices $v = \tilde{v}_2$ and $v = \tilde{v}_4$.

It remains to define the values of $v^{\text{nod}}_K$ on the nodes of the edges $E_i$ and $E_2$, as well as on $\tilde{v}_1$ and $\tilde{v}_3$. We only consider $\tilde{v}_1$ and $E_1$ (the construction for $\tilde{v}_3$ and $E_2$ is completely analogous). If $\tilde{v}_1 \in \mathcal{N}(\mathcal{T})$ (i.e., $\tilde{v}_1$ is a hanging node in $\mathcal{T}$), then we define

$$v^{\text{nod}}_K(z) = 0, \quad z \in \mathcal{P}_1(E_1), \quad v^{\text{nod}}_K(\tilde{v}_1) = v_K(\tilde{v}_1). \quad (5.9)$$

On the other hand, if $\tilde{v}_1 \notin \mathcal{N}(\mathcal{T})$, then $E_1 \in \mathcal{E}(\mathcal{T})$ and $v_1 \in \mathcal{N}(\mathcal{T})$. In this case, we interpolate the values of the nodal interpolant over the long edge $E_1$ at the Gauss-Lobatto nodes on $E_1$. That is, we set

$$v^{\text{nod}}_K(z) = v_K(v_1) \Phi^{E_1}_K(z) + v_K(v_2) \Phi^{E_1}_K(z), \quad z \in \mathcal{P}_1(E_1) \cup \{\tilde{v}_1\}. \quad (5.10)$$

With the nodal values of $v^{\text{nod}}_K$ constructed (5.8)–(5.10), we have

$$v^{\text{nod}}_K(x) = \sum_{v \in \mathcal{N}(K)} v^{\text{nod}}_K(v) \Phi^v_K(x) + \sum_{E \in \mathcal{E}(K)} \sum_{i=1}^{p_e-1} \left( v^{\text{nod}}_K(z_i) \Phi^{E_i \mathcal{P}_E}_v(x) \right).$$

This finishes the construction of the interpolant of $v^{\text{nod}}$. Notice that $v^{\text{nod}} \in S_p(\mathcal{T})$ is continuous over edges $E \in \mathcal{E}_1(\mathcal{T})$ and satisfies

$$v_K(v) - v^{\text{nod}}_K(v) = 0, \quad v \in \mathcal{N}(\mathcal{T}) \text{ located on } \partial K, \quad (5.11)$$

as well as

$$v^{\text{nod}}_K|_E \in \mathcal{P}_E^{\text{nod}}(E), \quad E \in \mathcal{E}(\mathcal{T}), \quad K \in \mathcal{E}_E. \quad (5.12)$$

Secondly, we construct a function $v^{\text{edge}} \in S_p(\mathcal{T})$ related to the edge degrees of freedom. To do so, fix an element $K \in \mathcal{T}$. For any edge $E \in \mathcal{E}(K)$, we define $v^{E}_{K}$ by

$$v^{E}_{K} = \begin{cases} L^E_{p_{K,K}}([v_K - v^{\text{nod}}_K]_E), & E \in \mathcal{E}(\mathcal{T}), \\ L^E_{p_{K,K}}([v_K - v^{\text{nod}}_K]_{E_1},(v_K - v^{\text{nod}}_K)]_{E_2}), & E = E_1 \cup E_2, E_1, E_2 \in \mathcal{E}(\mathcal{T}), \end{cases}$$

with $L^E_{p_{K,K}}(\cdot, \cdot)$ in (5.3) and $L^E_{p_{K,K}}(\cdot, \cdot)$ in (5.4), respectively. We then define $v^{\text{edge}}$ elementwise as

$$v^{\text{edge}}_K(x) = \sum_{E \in \mathcal{E}(K)} v^{E}_{K}(x).$$

Thirdly, we construct a function $v^{\text{int}} \in S_p(\mathcal{T})$ simply by setting elementwise

$$v^{\text{int}}_K = v_K - v^{\text{nod}}_K - v^{\text{edge}}_K, \quad K \in \mathcal{T}.$$ 

Notice that $v^{\text{int}}_K$ belongs to $H^1_0(K)$. Hence, we have $v^{\text{int}} \in S_p(\mathcal{T})$.

In conclusion, any function $v \in S_p(\mathcal{T})$ can be decomposed into three parts:

$$v = v^{\text{nod}} + v^{\text{edge}} + v^{\text{int}}, \quad (5.13)$$

with $v^{\text{nod}}$, $v^{\text{edge}}$ and $v^{\text{int}}$ in $S_p(\mathcal{T})$ as defined above.
5.4 Proof of Theorem 4.1

In this section, we outline the proof of Theorem 4.1. Some of the auxiliary results are postponed to Sections 5.5 and 5.6. For $v \in S_h(\mathcal{F})$, we write $v = v^{\text{nod}} + v^{\text{edge}} + v^{\text{int}}$, according to (5.13). We will define the averaging operator $h_p v$ in three parts:

$$I_{hp} v = \vartheta^{\text{nod}} + \vartheta^{\text{edge}} + \vartheta^{\text{int}}, \quad (5.14)$$

with $\vartheta^{\text{nod}}$, $\vartheta^{\text{edge}}$, $\vartheta^{\text{int}} \in S_h^e(\mathcal{T})$. Since $v^{\text{int}} \in S_h^e(\mathcal{T})$, we simply take $\vartheta^{\text{int}} = v^{\text{int}}$. In Sections 5.5 and 5.6, we will further construct $\vartheta^{\text{nod}}$ and $\vartheta^{\text{edge}}$ such that the following two results hold true.

**Lemma 5.3** There is a conforming approximation $\vartheta^{\text{nod}} \in S_h^e(\mathcal{T})$ that satisfies

$$\sum_{K \in \mathcal{T}} \|v^{\text{nod}} - \vartheta^{\text{nod}}\|_{L^2(K)}^2 \lesssim \sum_{E \in \mathcal{E}(\mathcal{F})} \int_E p_E^{-2} h_E \|v^{\text{nod}}\|^2 ds,$$

$$\sum_{K \in \mathcal{T}} \|\nabla(v^{\text{nod}} - \vartheta^{\text{nod}})\|_{L^2(K)}^2 \lesssim \sum_{E \in \mathcal{E}(\mathcal{F})} \int_E p_E^{-2} h_E^{-1} \|v^{\text{nod}}\|^2 ds.$$

**Lemma 5.4** There is a conforming approximation $\vartheta^{\text{edge}} \in S_h^e(\mathcal{T})$ that satisfies

$$\sum_{K \in \mathcal{T}} \|v^{\text{edge}} - \vartheta^{\text{edge}}\|_{L^2(K)}^2 \lesssim \sum_{E \in \mathcal{E}(\mathcal{F})} \int_E p_E^{-2} h_E (\|v\|^2 + \|v^{\text{nod}}\|^2) ds,$$

$$\sum_{K \in \mathcal{T}} \|\nabla(v^{\text{edge}} - \vartheta^{\text{edge}})\|_{L^2(K)}^2 \lesssim \sum_{E \in \mathcal{E}(\mathcal{F})} \int_E p_E^{-2} h_E^{-1} (\|v\|^2 + \|v^{\text{nod}}\|^2) ds.$$

By the triangle inequality and Lemmas 5.3 and 5.4, we then obtain

$$\sum_{K \in \mathcal{T}} \|v - I_{hp} v\|_{L^2(K)}^2 \lesssim \sum_{E \in \mathcal{E}(\mathcal{F})} \int_E p_E^{-2} h_E \left(\|v\|^2 + \|v^{\text{nod}}\|^2\right) ds,$$

$$\sum_{K \in \mathcal{T}} \|\nabla(v - I_{hp} v)\|_{L^2(K)}^2 \lesssim \sum_{E \in \mathcal{E}(\mathcal{F})} \int_E p_E^{-2} h_E^{-1} \left(\|v\|^2 + \|v^{\text{nod}}\|^2\right) ds.$$

Theorem 4.1 now follows if we show that

$$\|v^{\text{nod}}\|_{L^2(E)}^2 \lesssim \|v\|_{L^2(E)}^2, \quad E \in \mathcal{E}(\mathcal{F}). \quad (5.15)$$

To prove (5.15), we denote by $v_1$ and $v_2$ the two end points of $E \in \mathcal{E}(\mathcal{F})$. By the construction of $v^{\text{nod}}$, the jump over $E$ satisfies

$$\|v^{\text{nod}}\|_{(v_1)} = \|v\|_{(v_1)}, \quad i = 1, 2.$$

Since $\|v^{\text{nod}}\|$ vanishes on all the interior Gauss-Lobatto nodes on $E$, item (i) in Lemma 5.1 and a scaling argument yield

$$\|v^{\text{nod}}\|_{L^2(E)}^2 \lesssim p_E^{-2} h_E (\|v^{\text{nod}}\|_{(v_1)}^2 + \|v^{\text{nod}}\|_{(v_2)}^2),$$

$$= p_E^{-2} h_E (\|v\|_{(v_1)}^2 + \|v\|_{(v_2)}^2) \lesssim p_E^{-2} h_E \|v\|_{L^2(E)}^2.$$
From Theorem 3.92 of Schwab (1998), we further have the inverse estimate

\[ \|v\|_{L^2(E)}^2 \lesssim p_E^2 h_E^{-1} \|v\|_{L^2(E)}^2. \]

This shows (5.15) and finishes the proof of Theorem 4.1, up to the proofs of Lemmas 5.3 and 5.4 which we present next.

5.5 Proof of Lemma 5.3

Let \( \nu_{\text{nod}} \in S_p(\mathcal{T}) \) be the nodal interpolant in the decomposition (5.13). We now shall construct the conforming approximation \( \vartheta_{\text{nod}} \) in \( S_p(\mathcal{T}) \). For simplicity, we shall omit the superscript "nod" and, in the sequel, write \( v \) for \( \nu_{\text{nod}} \) and \( \vartheta \) for \( \vartheta_{\text{nod}} \). For a node \( v \), we introduce the sets:

\[ \tilde{w}(v) = \{ K \in \mathcal{T} : v \in \mathcal{N}(K) \}, \quad w_E(v) = \{ E \in \mathcal{E}(\mathcal{T}) : v \in E \}. \]

Fix \( K \in \mathcal{T} \) and \( \tilde{K} \in \mathcal{A}(K) \). We proceed by distinguishing the same two cases as in Subsection 5.3.

Case 1: If \( \mathcal{A}(K) = \{ K \} \), we have \( K = \tilde{K} \). Then any elemental edge \( E \in \mathcal{E}(\tilde{K}) \) belongs to \( \mathcal{E}(\mathcal{T}) \) and we have \( \nu_{\tilde{K}|E} \in S_{p_E}(\tilde{E}) \). For any Gauss-Lobatto node \( v \) located on \( \partial \tilde{K} \), we define the value of \( \vartheta(v) \) by

\[
\vartheta(v) = \begin{cases} 
[w(v)]^{-1} \sum_{K \in \tilde{w}(v)} v_{\tilde{K}}(v), & v \in \mathcal{N}(\mathcal{T}), \\
0, & \text{otherwise}.
\end{cases} 
\]

Here, \( |\tilde{w}(v)| \) denotes the cardinality of the set \( \tilde{w}(v) \). (We have \( |\tilde{w}(v)| = 4 \) for \( v \in \mathcal{N}(\mathcal{T}) \).) Then we define \( \vartheta \) on \( \tilde{K} \) by

\[
\vartheta(x) = \sum_{v \in \mathcal{N}(\tilde{K})} \vartheta(v) \Phi_{\tilde{K}}^v(x). 
\]

From (5.7) and (5.17), we have

\[
\|v_{\tilde{K}} - \vartheta\|_{L^2(\tilde{K})} \lesssim \sum_{v \in \mathcal{N}(\tilde{K})} |v_{\tilde{K}}(v) - \vartheta(v)| \|\Phi_{\tilde{K}}^v\|_{L^2(\tilde{K})}. 
\]

Analogously to the argument on Page 1125 of Burman and Ern (2007), we have

\[
|v_{\tilde{K}}(v) - \vartheta(v)| \lesssim \sum_{E \in w_E(v)} p_E h_E^{-1/2} \|v\|_{L^2(E)}. 
\]

Hence, by combining (5.18), (5.19) and item (ii) in Lemma 5.1 (with scaling), we obtain

\[
\|v_{\tilde{K}} - \vartheta\|_{L^2(\tilde{K})}^2 \lesssim \sum_{E \in w_E(v)} \int_E p_E^2 h_E^2 \|v\|^2 ds.
\]

Case 2: If \( \mathcal{A}(K) \) consists of four elements, we define \( \vartheta \) on each element \( \tilde{K} \in \mathcal{A}(K) \) separately, analogously to the construction of the nodal interpolant in Subsection 5.3. Without loss of generality, we may again consider the case illustrated in Figure 3 (right). Since \( E_3, E_4 \in \mathcal{E}(\mathcal{T}) \), the function \( v \)
Furthermore, if the node \( \nu \) an argument as in (5.19). We obtain
\[
\theta(z) = v_{K}(z), \quad z \in \mathcal{Z}_{1}^{p}(E_{3}) \cup \mathcal{Z}_{1}^{p}(E_{4}), \quad \tilde{\theta}(\tilde{v}_{3}) = v_{K}(\tilde{v}_{4}). \tag{5.21}
\]
We further define the value of \( \theta \) on the vertex \( \tilde{v}_{2} \) by (5.16).

It remains to define the values of \( \theta \) for the nodes on the edges \( \tilde{E}_{1} \) and \( \tilde{E}_{2} \), as well as for \( \tilde{v}_{1} \) and \( \tilde{v}_{3} \). We only consider \( \tilde{v}_{1} \) and \( E_{1} \) (the construction for \( v_{3} \) and \( E_{2} \) is completely analogous). If \( \tilde{v}_{1} \notin \mathcal{N}(\mathcal{T}) \), then \( \tilde{v}_{1} \) is a hanging node of \( \mathcal{T} \) and \( \tilde{E}_{1} \in \mathcal{E}(\mathcal{T}) \). Thus, \( v_{K}|_{\tilde{E}_{1}} \in \mathcal{Z}_{1}^{P}(\tilde{E}_{1}) \). For any \( z \in \mathcal{Z}_{1}^{p}(\tilde{E}_{1}) \cup \{ \tilde{v}_{1} \} \), the value of \( \tilde{\theta}(z) \) is taken as in (5.16). On the other hand, if \( \tilde{v}_{1} \notin \mathcal{N}(\mathcal{T}) \), then \( E_{1} \in \mathcal{E}(\mathcal{T}) \) and we have that \( v_{K}|_{E_{1}} \in \mathcal{Z}_{1}^{p}(\tilde{E}_{1}) \). We define \( \tilde{\theta}(v_{1}) \) again by (5.16). Recall that \( \tilde{\theta}(v_{2}) = \tilde{\theta}(v_{3}) \) has already been defined. Then we set
\[
\tilde{\theta}(z) = \tilde{\theta}(v_{1}) \Phi_{K}^{v_{1}}(z) + \tilde{\theta}(v_{2}) \Phi_{K}^{v_{2}}(z), \quad z \in \mathcal{Z}_{1}^{p}(\tilde{E}_{1}) \cup \{ \tilde{v}_{1} \}. \tag{5.22}
\]
Now we construct \( \theta \) on \( K \) by setting
\[
\theta(x) = \sum_{v \in \mathcal{N}(K)} \tilde{\theta}(v) \Phi_{K}^{v}(x) + \sum_{E \in \mathcal{E}(K)} \sum_{i=1}^{p_{E}^{-1}} \left( \tilde{\theta}(\tilde{z}_{i}^{E}) \Phi_{E}^{\tilde{z}_{i}^{E}}(x) \right). \tag{5.23}
\]
This completes the construction of \( \theta \). Clearly, \( \theta \in \mathcal{S}_{0}^{2}(\mathcal{K}) \).

We shall now derive an estimate analogous to (5.20). To do so, we bound the difference between \( v_{K} \) and \( \tilde{\theta} \) on \( K \) as follows:
\[
\|v_{K} - \tilde{\theta}\|_{L^{2}(K)} \lesssim \sum_{v \in \mathcal{N}(K)} \|v_{K}(v) - \tilde{\theta}(v)\|_{L^{2}(K)} + \sum_{E \in \mathcal{E}(K)} \|s_{E}\|_{L^{2}(E)} \leq T_{1} + T_{2}, \tag{5.24}
\]
with \( s_{E}(x) = \sum_{i=1}^{p_{E}^{-1}} \left( v_{K}(\tilde{z}_{i}^{E}) - \tilde{\theta}(\tilde{z}_{i}^{E}) \right) \Phi_{E}^{\tilde{z}_{i}^{E}}(x) \). For the second inequality in (5.23), we have used estimate (i) in Lemma 5.1 and a scaling argument noticing that the function \( s_{E}(x) \) vanishes at all the interior tensor-product Gauss-Lobatto nodes in \( K \) and on the edges of \( K \) that are different from \( E \).

Let us first bound the term \( T_{1} \) in (5.23). If the node \( v \in \mathcal{N}(\mathcal{T}) \), then, by (5.21),
\[
(v_{K}(v) - \tilde{\theta}(v)) \Phi_{K}^{v}(x) = 0, \quad x \in K. \tag{5.24}
\]
Furthermore, if the node \( v \) belongs to \( \mathcal{N}(\mathcal{T}) \), we apply estimate (ii) in Lemma 5.1 (with scaling) and an argument as in (5.19). We obtain
\[
\|v_{K}(v) - \tilde{\theta}(v)\|_{L^{2}(K)} \lesssim \sum_{E \in \mathcal{E}(v)} p_{E}^{-1} h_{E}^{1/2} \|v\|_{L^{2}(E)}. \tag{5.25}
\]
Finally, if \( v \notin \mathcal{N}(\mathcal{T}) \cup \mathcal{N}(\mathcal{T}) \), then \( v \) is midpoint of the edge \( E \), \( E \in \mathcal{E}(K) \cap \mathcal{E}(\mathcal{T}) \). Denote the two end points of this edge \( E \) by \( v_{1} \) and \( v_{2} \). In view of (5.10) and (5.22), we have
\[
|v_{K}(v) - \tilde{\theta}(v)| \leq |v_{K}(v_{1}) - \tilde{\theta}(v_{1})| + |v_{K}(v_{2}) - \tilde{\theta}(v_{2})|. \]
Thus, as before, we obtain

$$
\| (v_E - \vartheta) \Phi_k \|_{L^2(E)} \lesssim \sum_{E \in w^F(v_1),w^F(v_2)} \tilde{p}_E^{-1} h_E^2 \| v \|_{L^2(E)}.
$$

To combine the results in (5.24)-(5.26), we define the set $\mathcal{N}^*(\tilde{K})$ as follows. We start from the set $\mathcal{N}(\tilde{K})$ and first remove all the vertices belonging to $\mathcal{N}_A(\tilde{T})$. Then, we replace any vertex $\tilde{v} \in \mathcal{N}(\tilde{K})$ with $\tilde{v} \notin \mathcal{N}(\tilde{T}) \cup \mathcal{N}_A(\tilde{T})$ by the vertex $v \in \mathcal{N}(K)$ which is on the same elemental edge of $K$ as $\tilde{v}$. For example, in the case shown in Figure 3 (right), we have $\mathcal{N}^*(\tilde{K}) = \{ v_1, v_2, v_3 \}$, provided that we have $\tilde{v}_1 \notin \mathcal{N}(\tilde{T})$ and $v_3 \in \mathcal{N}(T)$. In conclusion, if we also set

$$
\mathcal{N}^*(\tilde{K}) = \{ E \in w^F(v) : v \in \mathcal{N}^*(\tilde{K}) \}
$$

then the term $T_1$ can be bounded by

$$
T_1 \lesssim \sum_{E \in \mathcal{N}^*(\tilde{K})} \tilde{p}_E^{-1} h_E^2 \| v \|_{L^2(E)}.
$$

Next, let us estimate the term $T_2$ in (5.23). If $\tilde{E} \in \mathcal{N}(\tilde{T})$ or $\tilde{E} \in \mathcal{N}_A(\tilde{T})$, by the constructions of $v$ and $\vartheta$, we clearly have $\| \xi_E \|_{L^2(\tilde{E})} = 0$. Otherwise, one of the two end points of $\tilde{E}$, say $v_1$, is a newly created node in $\tilde{T}$ and the other one, $v_2$, belongs to $\mathcal{N}(\tilde{T})$. Thus, we have

$$
\tilde{p}_k^{-1} h_k^2 \xi_E \|_{L^2(E)} \lesssim \tilde{p}_k^{-1} h_k^2 \sum_{i=1}^2 \| (v_k(v_i) - \vartheta(v_i)) \Phi_k \|_{L^2(\tilde{E})} + \tilde{p}_k^{-1} h_k^2 \| v_k - \vartheta \|_{L^2(\tilde{E})} = T_{21} + T_{22}.
$$

Then there exists an elemental edge $E \in \mathcal{N}(K)$ such that $v_1$ is the midpoint of $E$. Denote the end points of $E$ by $v_1$ and $v_2$. Similarly to (5.25) and (5.26), we have

$$
T_{21} \lesssim \sum_{E \in w^F(v_1),w^F(v_2)} \tilde{p}_E^{-1} h_E^2 \| v \|_{L^2(E)}.
$$

In view of (5.22), we proceed as in (5.19) and obtain

$$
T_{22} \lesssim \tilde{p}_k^{-1} h_k^2 \sum_{i=1}^2 \| (v_k(v_i) - \vartheta(v_i)) \Phi_k \|_{L^2(E)} \lesssim \sum_{E \in w^F(v_1),w^F(v_2)} \tilde{p}_E^{-1} h_E^2 \| v \|_{L^2(E)}.
$$

Combining the above results shows that

$$
T_2 \lesssim \sum_{E \in \mathcal{N}^*(\tilde{K})} \tilde{p}_E^{-1} h_E^2 \| v \|_{L^2(E)}.
$$

The bounds for $T_1$ and $T_2$ in (5.27) and (5.28) yield

$$
\| v_k - \vartheta \|_{L^2(\tilde{K})} \lesssim \sum_{E \in \mathcal{N}^*(\tilde{K})} \tilde{p}_E^{-2} h_E \| v \|_{L^2(\tilde{E})}^2 ds.
$$

This finishes the discussion of Case 2.
We are now ready to complete the proof of Lemma 5.3. Notice that, by the key estimates in (5.20) for Case 1 and (5.29) for Case 2, we have
\[
\|v_{\tilde{K}} - \vartheta\|^2_{L^2(\tilde{K})} \lesssim \sum_{E \in \mathcal{E}(\tilde{K})} \int_E p_{E}^{-2} h_E \|v\|^2 \, ds, \quad \tilde{K} \in \tilde{\mathcal{T}}.
\] (5.30)
This proves the first inequality in Lemma 5.3. Moreover, by the inverse inequality
\[
\|
abla v\|_{L^2(\tilde{K})} \lesssim p_{E}^{-2} h_E^{-1} \|v\|_{L^2(\tilde{K})}, \quad v \in S_p(\tilde{\mathcal{T}}), \quad \tilde{K} \in \tilde{\mathcal{T}},
\] (5.31)
see Schwab (1998), we obtain from (5.30)
\[
\|
abla (v_{\tilde{K}} - \vartheta)\|_{L^2(\tilde{K})} \lesssim \sum_{E \in \mathcal{E}(\tilde{K})} \int_E p_{E}^{-2} h_E^{-1} \|v\|^2 \, ds, \quad \tilde{K} \in \tilde{\mathcal{T}},
\] (5.32)
which shows the second assertion in Lemma 5.3.

5.6 Proof of Lemma 5.4
Fix an element \( K \in \mathcal{T} \) and let \( E \) be an elemental edge in \( \mathcal{E}(K) \). We define the function \( W^E_K \) as follows: If \( E \in \mathcal{E}_B(\mathcal{T}) \), we set
\[
W^E_K = L^E_{p, K}((v_K - v_K^{\text{nod}})|_E),
\]
with the extension operator \( L^E_{p, K}(\cdot) \) in (5.3). If \( E \in \mathcal{E}(\mathcal{T}) \), let \( K' \) be the element such that \( E \) is also an elemental edge of \( K' \), that is, \( E \in \mathcal{E}(K') \). Denote by \( K'' \) the element which has the lower polynomial degree of the elements \( K \) and \( K' \), i.e., \( K'' = K \) if \( p_K \leq p_{K'} \) and \( K'' = K' \) otherwise. We define \( W^E_K \) by
\[
W^E_K = L^E_{p, K}((v_K' - v_K'^{\text{nod}})|_E),
\]
with \( L^E_{p, K}(\cdot, \cdot) \) in (5.4). In the case where \( E \) contains a hanging node, \( E \) is partitioned into \( E = E_1 \cup E_2 \) with \( E_1, E_2 \in \mathcal{E}_I(\mathcal{T}) \). There exist two elements \( K', K'' \in \mathcal{T} \) such that \( E_1 \in \mathcal{E}(K') \) and \( E_2 \in \mathcal{E}(K'') \). Denote by \( K' \) the element that has the lower polynomial degree of \( K \) and \( K' \), and by \( K'' \) the element that has the lower degree of \( K' \) and \( K'' \). We now define \( W^E_K \) by
\[
W^E_K = L^E_{p, K}((v_K' - v_K'^{\text{nod}})|_{E_1}, (v_K'' - v_K''^{\text{nod}})|_{E_2}),
\]
with \( L^E_{p, K}(\cdot, \cdot) \) in (5.4).

Then we define \( \vartheta^{\text{edge}} \) elementwise by setting \( \vartheta^{\text{edge}}|_{K} = \sum_{E \in \mathcal{E}(K)} W^E_K \), with \( W^E_K \) defined above. Clearly, the function \( \vartheta^{\text{edge}} \) belongs to \( S^p_{p}(\tilde{\mathcal{T}}) \). We now prove the approximation properties of Lemma 5.4.

By Lemma 5.2 (with a scaling argument), we have
\[
\sum_{K \in \mathcal{T}} \|v - \vartheta^{\text{edge}}\|^2_{L^2(\tilde{K})} \lesssim \sum_{K \in \mathcal{T}} \sum_{E \in \mathcal{E}(K)} \|v - \vartheta^{\text{edge}}\|^2_{L^2(\tilde{K})} \lesssim \sum_{K \in \mathcal{T}} \sum_{E \in \mathcal{E}(K)} \|L^E_{p, K}((v_K - v_K^{\text{nod}})|_E) - W^E_K\|^2_{L^2(\tilde{K})} \lesssim \sum_{K \in \mathcal{T}} \sum_{E \in \mathcal{E}(K)} \int_E p_{E}^{-2} h_E \|(v_K - v_K^{\text{nod}})|_E - W^E_E\|^2_{L^2(\tilde{K})} ds \lesssim \sum_{K \in \mathcal{T}} \sum_{E \in \mathcal{E}(K)} \int_E p_{E}^{-2} h_E (\|v\|^2 + \|v^{\text{nod}}\|^2) \, ds.
\]
This completes the proof of the first assertion of Lemma 5.4; the second one follows again from the first one by using the inverse inequality in (5.31).

6. Numerical experiments

In this section, we present a series of numerical examples where we use $\eta$ in (3.4) as an error indicator in an $h$-$p$-adaptive refinement strategy. Our implementation of the DG method (2.6) is based on the Deal.II finite element library, see Bangerth et al. (2009) and Bangerth et al. (2007). The non-symmetric sparse linear systems of equations are solved by using the UMFPACK package; cf. Davis (2004a,b). In all the examples, the $h$-$p$-adaptive meshes are constructed by first marking the elements for refinement and derefinement according to the size of the local error indicator $\eta_K$ in (3.3), with refinement and derefinement fractions set to 25% and 10%, respectively. To decide whether to $h$- or $p$-refine/derefine a marked element, we employ the smoothness estimation technique developed by Houston and Süli (2005) and Houston et al. (2003). If the approximate solution is smooth enough on element $K$ according to the criterion employed, we increase the local polynomial degree $p_K$ by one; otherwise, we refine the element isotropically into four elements by bisecting the elemental edges of $K$. Additional refinement might be performed to ensure that the meshes are 1-irregular.

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 $S_p(\mathcal{T})$. Moreover, since the flow field $g$ is either constant or linear, we simply choose $a_{hp} = a$ in $\eta_{R_e}$. We numerically reproduce solutions that are analytic over the computational domain, although they have steep gradients along boundary and internal layers. In all our examples, we observe $p$-refinement to be dominating once the local mesh size is sufficiently resolved.

Based on the a-priori error analysis for $p$-version methods in Schwab (1998), we thus plot all computed quantities against $N^{1/2}$ in a logarithmic scale, with $N = \dim(S_p(\mathcal{T}))$.

6.1 Example 1

In this example, we take $\Omega = (0, 1)^2$, choose $g = (1, 1)$ and select the right-hand side $f$ so that the analytical solution to the convection-diffusion problem (2.1) is given by

$$u(x_1, x_2) = \left(\frac{e^{x_1 - 1}}{e^{-\tau} - 1} + x_1 - 1\right)\left(\frac{e^{x_2 - 1}}{e^{-\tau} - 1} + x_2 - 1\right).$$
The solution is smooth, but has boundary layers at $x_1 = 1$ and $x_2 = 1$; their widths are both of order $O(\varepsilon)$. This problem is well-suited to test whether the indicator $\eta$ is able to pick up the steep gradients near these boundaries.

We begin this test with a uniform mesh of $16 \times 16$ elements and uniform polynomial degree $p_K = 1$. In Figure 4(a), we show the performance of our $hp$-adaptive algorithm for $\varepsilon = 10^{-3}$. In the curves labelled "Error Indicator" and "Energy Error", we see that the indicator $\eta$ always overestimates the true energy error $\|u - u_{hp}\|_{E,\mathcal{T}}$, in agreement with Theorem 3.1. Additionally, the convergence lines using $hp$-refinement are (roughly) straight on a linear-log scale, which indicates that exponential convergence is attained for this problem. In the curve "$L^2$ Error", we calculate the error $\varepsilon^{-\frac{1}{2}}\|u - u_{hp}\|_{L^2(\Omega)}$, which is an upper bound for $|g (u - u_h)|_\ast$. We see that this error is at least of the same order as the energy error. The same behaviour is observed for the error $\left(\sum_{K \in \mathcal{T}} (\varepsilon^{-1} p_K^{-1} h_E \|u_h\|_{L^2(E)}^2)\right)^{1/2}$ shown in "Jump Error". Finally, in the curve labelled "Theta", we calculate an approximation to the data error $\Theta$ in (3.4), by using a Gauss-Legendre quadrature rule of order $p_K + 3$ to approximate $\Theta^2_K = \varepsilon^{-1} p_K^{-2} h_E^2 \|f - f_{hp}\|_{L^2(K)}^2$ on each element $K$. The data approximation error $\Theta$ is of almost three orders of magnitude smaller than $\eta$. In Figure 4(b), we compare the true energy error and the error indicator generated by our $hp$-adaptive algorithm using the indicator $\eta$ in (3.4) (denoted by $p^3$ in the figure) and the corresponding one outlined in Remark 3.1 (denoted by $p^2$ in the figure). As in Houston et al. (2008), we observe that the two error indicators give rise to almost identical results. In Figure 4(c), we plot the ratios of the indicator and the true energy error. It stays around 8, uniformly in $N^2$.

In Figure 5, we show the same plots for $\varepsilon = 2 \cdot 10^{-4}$. Qualitatively, we observe the same behaviour as before. Together with Figure 4(c), we see that the ratio of the indicator and the true energy error oscillates around 8 for both $\varepsilon = 10^{-3}$ and $\varepsilon = 2 \cdot 10^{-4}$, in agreement with Theorems 3.1 and 3.2. In Figure 6, we show the meshes and polynomial degree distribution after $7 hp$-adaptive refinement steps. We observe that the $p$-refinement is dominating once the local mesh size is of order $O(\varepsilon)$, the order of the width of the boundary layer. The $p$-refinement is concentrated around the boundary layers. Away from the layers, the solution is almost linear and can be approximated efficiently with low-order polynomials.

### 6.2 Example 2

Next, we consider an example with an internal layer and with variable coefficients. In the domain $\Omega = (-1,1)^2$, we take $g(x_1,x_2) = (-x_1,x_2)$. We choose $f$ and the inhomogeneous Dirichlet boundary
conditions such that the solution to (2.1) is given by

\[ u(x_1,x_2) = \text{erf} \left( \frac{x_1}{\sqrt{2\varepsilon}} \right) (1 - x_2^2), \quad \text{with} \quad \text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt. \]

For small values of \( \varepsilon \), the solution \( u \) has an internal layer around \( x_1 = 0 \), whose width is of order \( O(\sqrt{\varepsilon}) \).

We use standard DG terms to incorporate the inhomogeneous boundary conditions, and modify the error indicator \( \eta \) correspondingly. For details we refer to Houston et al. (2007). We begin this test with a uniform mesh of \( 8 \times 8 \) elements and the uniform polynomial degree \( p_K = 2 \). In Figure 7 and Figure 8, the numerical results for this example are shown for the values \( \varepsilon = 10^{-3} \) and \( \varepsilon = 5 \cdot 10^{-6} \), respectively. We plot the same quantities as in Example 1.

For \( \varepsilon = 10^{-3} \), we observe exponential convergence rates for both the energy error and the indicator. The curves associated with the convection and approximation errors are of the same order as the energy error. In particular, the jump error related to convection in the curve "Jump Error" and \( \Theta \) in "Theta" are clearly below the energy error. If we now decrease the value of \( \varepsilon \) to \( \varepsilon = 5 \cdot 10^{-6} \), the jump related to
reciprocal of the local Péclet number depicted in “Jump Error” is dominating the estimator $\eta$. (Recall that the error plotted in the curve “$L^2$ Error” is only an upper bound for the error $|g(u - u_h)|_*$ and can overestimate $\eta$.) Nevertheless, exponential convergence rates are observed for all quantities. This illustrates the fact that the estimator $\eta$ is not robust in estimating the energy error alone; the inclusion of the dual norm in the error measure is essential. This is further reflected in the plots at the right-hand sides of Figures 7 and 8 where we show the ratio of the indicator and the energy error. While for $\epsilon = 10^{-3}$ the values are between 8 and 12, they clearly increase for $\epsilon = 5 \cdot 10^{-5}$. Initially, they also strongly oscillate. Again, this is due to the fact that we do not include the dual norm in the error measure.

Figure 9 shows the $hp$-adaptive meshes and polynomial degree distributions after 9 refinement and 16 refinement steps, both for $\epsilon = 10^{-3}$ and $\epsilon = 5 \cdot 10^{-6}$. We observe that the mesh refinement stops once the local mesh size is of order $O(\sqrt{\epsilon})$ and $p$-refinement starts to take over in the vicinity of the layer, which is much more pronounced for $\epsilon = 5 \cdot 10^{-6}$.

### 6.3 Example 3

Finally, we test the performance of our method for a problem with a convection field that is not aligned with the mesh. We take $\Omega = (-1, 1)^2$, $a = (-\sin \frac{\pi}{6}, \cos \frac{\pi}{6})$, $f = 0$ and consider the boundary conditions $u = 0$ on $x_1 = -1$ and $x_2 = 1$, as well as

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

The boundary data is almost discontinuous near the point $(0, -1)$ and causes $u$ to have an internal layer of width $O(\sqrt{\epsilon})$ 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 outflow boundary. There is no exact solution available to this problem. We test this problem with $\epsilon = 10^{-3}$ and start the algorithm for $p_K = 2$ on a uniform mesh of $16 \times 16$ elements.

In Figures 10 (left), we plot the values of the indicator $\eta$ for $\epsilon = 10^{-3}$. We observe almost immediately exponential convergence for $\eta$. Figure 10 (right) depicts the adaptive meshes after 7 refinement steps. Since the solution is almost constant away from the layers, $p$-refinement is again concentrated along the layers.
7. Conclusions

We have derived a robust a-posteriori error estimate for \( hp \)-adaptive discontinuous Galerkin methods for convection-diffusion equations on 1-irregular parallelogram meshes. The ratio of the constants in the reliability and efficiency bounds is independent of the Péclet number \( \varepsilon \) of the equation, and hence the estimate is fully robust. We have applied our estimate as an error indicator for energy norm error estimation in an \( hp \)-adaptive refinement algorithm. Our numerical results indicate that the indicator is effective in locating and resolving interior and boundary layers. Once the local mesh size is of the same order as the width of the boundary or interior layer, both the energy error and the error indicator are observed to converge exponentially.

The results of this paper can be extended to problems with zero-order terms; see Schötzau and Zhu (2009) for the \( h \)-version of the DG method. In this case, the flow field \( \psi \) does not necessarily need to be divergence-free. Instead, an assumption on the coefficient functions as in Schötzau and Zhu (2009) or Verfürth (2005) is sufficient.

The a-posteriori analysis presented in this paper is based on the availability of an averaging operator as in Theorem 4.1. The difficulties in extending our results to three-dimensional problems are only related to the technicalities of the construction and analysis of such an operator on three-dimensional irregular meshes. This will be discussed in a forthcoming paper.
Finally, we remark that our analysis only holds for isotropically refined meshes. In view of the powerful $hp$-version approximation results on anisotropic meshes, see Schwab (1998), it would be desirable to allow for anisotropic $h$- and $p$- refinement as well, similarly to the approach for functional error estimation developed by Georgoulis et al. (2007). This is the subject of ongoing research.

Acknowledgments
The authors are thankful to Professor Paul Houston (University of Nottingham) for helpful discussions on the local smoothness estimation algorithms in Houston et al. (2003) and Houston and Süli (2005).

References


REFERENCES


