A posteriori error analysis of hp-version discontinuous Galerkin finite-element methods for second-order quasi-linear elliptic PDEs

Paul Houston†
School of Mathematical Sciences, University of Nottingham,
University Park, Nottingham NG7 2RD, UK

Endre Süli‡
Computing Laboratory, University of Oxford, Parks Road,
Oxford OX1 3QD, UK

And

Thomas P. Wihler§
Department of Mathematics and Statistics, McGill University,
805 Sherbrooke West, Montréal, Québec H3A 2K6, Canada

[Received on 2 November 2006; revised on 5 April 2007]

We develop the a posteriori error analysis of hp-version interior-penalty discontinuous Galerkin finite-element methods for a class of second-order quasi-linear elliptic partial differential equations. Computable upper and lower bounds on the error are derived in terms of a natural (mesh dependent) energy norm. The bounds are explicit in the local mesh size and the local polynomial degree of the approximating finite element function. The performance of the proposed error indicators within an automatic hp-adaptive refinement procedure is studied through numerical experiments.

Keywords: hp-adaptivity; a posteriori error analysis; discontinuous Galerkin finite-element methods; quasi-linear elliptic PDEs.

1. Introduction

In this article, we consider the a posteriori error analysis, in a natural mesh-dependent energy norm, for a class of interior-penalty hp-version discontinuous Galerkin finite-element methods (DGFEMs) for the numerical solution of the following quasi-linear elliptic boundary-value problem:

\[-\nabla \cdot (\mu(x, |\nabla u|)\nabla u) = f \quad \text{in } \Omega, \quad \quad (1.1)\]
\[u = 0 \quad \text{on } \Gamma. \quad \quad (1.2)\]

Here, \(\Omega\) is a bounded polygonal domain in \(\mathbb{R}^2\) with boundary \(\Gamma\) and \(f \in L^2(\Omega)\). Additionally, we assume that the non-linearity \(\mu\) satisfies the following assumptions:

\[(A1) \quad \mu \in C^0(\overline{\Omega} \times [0, \infty)) \text{ and} \]

† Email: paul.houston@nottingham.ac.uk
‡ Email: endre.suli@comlab.ox.ac.uk
§ Email: wihler@math.mcgill.ca
(A2) there exist positive constants $m_\mu$ and $M_\mu$ such that

$$m_\mu(t-s) \leq \mu(x, t)t - \mu(x, s)s \leq M_\mu(t-s), \quad t \geq s \geq 0, \quad x \in \overline{\Omega}.$$  \hfill (1.3)

We remark that if $\mu$ satisfies (1.3), there exist constants $C_1$ and $C_2$, $C_1 \geq C_2 > 0$, such that for all vectors $v, w \in \mathbb{R}^2$ and all $x \in \overline{\Omega}$,

$$|\mu(x, |v|)v - \mu(x, |w|)w| \leq C_1|v - w|,$$  \hfill (1.4)

$$C_2|v - w|^2 \leq (\mu(x, |v|)v - \mu(x, |w|)w) \cdot (v - w),$$  \hfill (1.5)

see Liu & Barrett, 1994, Lemma 2.1.

Non-linearities of this kind appear in numerous problems in continuum mechanics. In particular, they arise in mathematical models for non-Newtonian fluids, such as the following generalized power-law model: given $f \in L^2(\Omega)^2$, find $(u, p) \in H^1(\Omega)^2 \times L^2(\Omega)/\mathbb{R}$ such that

$$-\nabla \cdot (\mu(x, |u|)e(u)) + \nabla p = f \quad \text{in } \Omega,$$

$$\text{div} u = 0 \quad \text{in } \Omega,$$

$$u = 0 \quad \text{on } \Gamma,$$

where $u = (u_1, u_2)^T$ is the velocity vector, $p$ is the pressure, $f = (f_1, f_2)^T$ is the applied force, $e(u)$ is the symmetric $2 \times 2$ strain tensor defined by

$$e_{ij}(u) = \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \quad i, j = 1, 2,$$

and $|e(u)|$ is the Frobenius norm of $e(u)$.

For the sake of notational simplicity, we shall suppress the dependence of $\mu$ on $x$ and write $\mu(t)$ instead of $\mu(x, t)$. Indeed, in many physical applications, $\mu$ is in fact independent of $x$; e.g. in the Carreau law for a non-Newtonian fluid, we have $\mu(t) = \mu_\infty + (\mu_0 - \mu_\infty)(1 + \lambda t^2)^{\frac{r-2}{2}}$, where $\lambda > 0$, $1 < r \leq 2$ and $0 < \mu_\infty < \mu_0$.

In recent years, there has been considerable interest in DGFEMs for the numerical solution of a wide range of partial differential equations (PDEs). We shall not attempt to give an extensive survey of this area of research: the reader is referred to Cockburn et al. (2000) for a detailed review. DGFEMs were introduced in the early 1970s for the numerical solution of first-order hyperbolic problems (see Cockburn & Shu, 1989, 1991, 1998b; Johnson & Pitkäranta, 1986; LeSaint & Raviart, 1974; Reed & Hill, 1973). Simultaneously, but quite independently, they were proposed as non-standard schemes for the approximation of second-order elliptic equations (Arnold, 1982; Nitsche, 1971; Wheeler, 1978). The recent upsurge of interest in this class of methods has been stimulated by the computational convenience of DGFEMs due to a high degree of locality, the need to approximate advection-dominated diffusion problems without excessive numerical stabilization, the necessity to accommodate high-order $hp$- and spectral-element discretizations for first-order hyperbolic equations and advection–diffusion problems (Flaherty et al., 2000; Karniadakis & Sherwin, 1999) and the desire to handle non-linear hyperbolic problems in a locally conservative manner and without auxiliary numerical stabilisation (Cockburn et al., 1990; Cockburn & Shu, 1998a), see also Castillo et al. (2000) and Cockburn et al. (2001) for the error analysis of the local version of the DGFEM in the elliptic case, as well as Arnold et al. (2001) and Oden et al. (1998).

In the recent article Houston et al. (2005), a family of interior-penalty $hp$-DGFEMs was formulated for the numerical approximation of the scalar quasi-linear boundary-value problem (1.1)–(1.2), and $a$ priori bounds were derived on the error, measured in terms of a mesh-dependent energy norm.
These error bounds are optimal with respect to the mesh size $h$ and mildly suboptimal (by $p^{1/2}$) in the polynomial degree $p$; more precisely, for $u \in C^1(\Omega) \cap H^k(\Omega)$, $k \geq 2$, it was shown that for any member of the family of methods considered, the error tends to zero at the rate $O(h^{s-1/p}k^{3/2})$, where $1 \leq s \leq \min\{p+1, k\}$, as $h$ tends to zero and $p$ tends to infinity. For related work on $h$-version local DGFEMs for quasi-linear PDEs, we refer to the articles by Bustinza & Gatica (2004) and Gatica et al. (2004), for example.

In the present work, we shall extend the above-mentioned results by considering the a posteriori error analysis of the interior-penalty $hp$-DGFEMs from Houston et al. (2005). In particular, we shall derive computable upper and lower bounds on the error, measured in terms of the underlying DG energy norm, which are explicit in terms of their dependence on the local element size $h$ and the local polynomial degree $p$. Our error analysis is based upon the following key ideas:

1. The ‘upper error bound’ is obtained by applying similar techniques as presented in Houston et al. (2004); Houston et al. (2005, 2007a,b). In contrast to the $hp$-version a posteriori error analysis in Houston et al. (2007b), however, we present a suitably extended approach that permits the use of one-irregular meshes. In particular, this means that elements can be divided into smaller elements without the need of connecting the resulting hanging nodes. This feature clearly improves both the feasibility and the flexibility of an $hp$-adaptive process. In addition, we note that the use of irregular meshes is very natural and quite easily realisable in the context of DG schemes due to the discontinuous character of the corresponding finite-element spaces. In order to derive a posteriori error indicators for the DG methods, in this paper we consider a given $hp$-DG finite-element space (based on irregular meshes) as a subspace of a slightly larger space that is based on conforming meshes only (with an equivalent distribution of local element sizes and local polynomial degrees). We will then follow the ideas in Houston et al. (2005, 2007a,b) to decompose the latter space into a direct sum of two subspaces consisting of continuous and purely discontinuous finite-element functions, respectively. The error of the numerical approximation can then be split into two corresponding components: a conforming part and a discontinuous (but polynomial) part. In order to deal with the former part of the error (under minimal regularity assumptions on the analytical solution), the approach in Houston et al. (2007b) suggests the introduction of an auxiliary formulation for the DG method through the use of lifting operators. In the present paper, we propose an alternative technique. More precisely, we split the integrals appearing in the norm of the error in a suitable manner so that each term remains well-defined even under the given minimal regularity conditions. Furthermore, in order to deal with the discontinuous part of the error, we proceed as in Houston et al. (2007b) and apply an $hp$-norm equivalence property on the discontinuous part of the $hp$-DG space.

2. The proof of the local ‘lower error bounds’ (efficiency) is based on the techniques presented in Melenk & Wohlmuth (2001), and follows, subject to the treatment of the non-linearity, in a similar manner to the proof of the efficiency estimates derived in Theorem 3.2 of Houston et al. (2005).

As in the case of the conforming $hp$-version finite-element methods considered in Melenk & Wohlmuth (2001), reliability and efficiency of our error bounds cannot be established uniformly with respect to the polynomial degree since the proof of efficiency relies on employing inverse estimates which are suboptimal in the spectral order.

The outline of this article is as follows: In Section 2, we revisit the $hp$-DGFEM introduced in Houston et al. (2005) for the numerical approximation of the boundary-value problem (1.1)–(1.2). In Section 3, our a posteriori error bounds are presented and discussed; both upper and lower energy norm
bounds will be derived. In Section 4, we present a series of numerical experiments to illustrate the performance of the proposed \(hp\)-error indicators within an automatic \(hp\)-refinement algorithm. Finally, in Section 5, we summarize the main results of this article and draw some conclusions.

Throughout the paper, we use the following standard function spaces. For a bounded Lipschitz domain \(D \subset \mathbb{R}^d, d \geq 1\), we write \(H^t(D)\) to denote the usual (real) Sobolev space of order \(t \geq 0\) with norm \(|| \cdot ||_t, D\). In the case \(t = 0\), we set \(L^2(D) = H^0(D)\). We define \(H^1_0(D)\) to be the subspace of functions in \(H^1(D)\) with zero trace on \(\partial D\). For a function space \(X(D)\), we write \(X(D)^d\) to denote the space of \(d\)-component vector fields whose components belong to \(X(D)\); this space is equipped with the usual product norm which, for simplicity, is denoted in the same way as the norm in \(X(D)\).

2. \(hp\)-Version DGFEM

Let \(\mathcal{T}_h\) be a subdivision of \(\Omega\) into disjoint open-element domains \(\kappa\) such that \(\overline{\Omega} = \bigcup_{\kappa \in \mathcal{T}_h} \overline{\kappa}\). We assume that the family of subdivisions \(\{\mathcal{T}_h\}_{h > 0}\) is shape regular (see, e.g. Braess, 1997, pp. 61, 113, and Remark 2.2, p. 114) and each \(\kappa \in \mathcal{T}_h\) is an affine image of a fixed master element \(\widehat{\kappa}\); i.e. for each \(\kappa \in \mathcal{T}_h\), there exists an affine mapping \(F_\kappa: \widehat{\kappa} \rightarrow \kappa\) such that \(\kappa = F_\kappa(\widehat{\kappa})\), where \(\widehat{\kappa}\) is either the open unit triangle \(\{(x, y): -1 < x < 1, -1 < y < -x\}\) or the open unit square \((-1, 1)^2\) in \(\mathbb{R}^2\). By \(h_\kappa\) we denote the element diameter of \(\kappa \in \mathcal{T}_h\), \(h = \max_{\kappa \in \mathcal{T}_h} h_\kappa\), and \(n_\kappa\) signifies the unit outward normal vector to \(\kappa\). We allow the meshes \(\mathcal{T}_h\) to be ‘one irregular’, i.e. each edge of any one element \(\kappa \in \mathcal{T}_h\) contains at most one hanging node (which, for simplicity, we assume to be the midpoint of the corresponding edge).

Here, we suppose that \(\mathcal{T}_h\) is ‘regularly reducible’ (cf. Ortner & Süli, 2006, Section 7.1), i.e. there exists a shape-regular conforming (regular) mesh \(\widehat{\mathcal{T}}_h\) (consisting of triangles and parallelograms) such that the closure of each element in \(\mathcal{T}_h\) is a union of closures of elements of \(\widehat{\mathcal{T}}_h\), and that there exists a constant \(C > 0\), independent of the element sizes, such that for any two elements \(\kappa \in \mathcal{T}_h\) and \(\widehat{\kappa} \in \widehat{\mathcal{T}}_h\) with \(\widehat{\kappa} \subseteq \kappa\), we have \(h_\kappa / h_{\widehat{\kappa}} \leq C\). Note that these assumptions imply that the family \(\{\mathcal{T}_h\}_{h > 0}\) is of ‘bounded local variation’, i.e. there exists a constant \(\rho_1 \geq 1\), independent of the element sizes, such that

\[
\rho_1^{-1} \leq h_\kappa / h_{\kappa'} \leq \rho_1
\]  

(2.1)

for any pair of elements \(\kappa, \kappa' \in \mathcal{T}_h\) which share a common edge \(e = \partial \kappa \cap \partial \kappa'\).

For a non-negative integer \(k\), we denote by \(\mathcal{P}_k(\widehat{\kappa})\) the set of polynomials of total degree \(k\) on \(\widehat{\kappa}\). When \(\widehat{\kappa}\) is the unit square, we also consider \(\mathcal{P}_k(\widehat{\kappa})\) the set of all tensor-product polynomials on \(\widehat{\kappa}\) of degree \(k\) in each co-ordinate direction. To each \(\kappa \in \mathcal{T}_h\), we assign a polynomial degree \(p_\kappa\) (local approximation order).

We store the \(h_\kappa, p_\kappa\) and \(F_\kappa\) in the vectors \(\mathbf{h} = \{h_\kappa: \kappa \in \mathcal{T}_h\}, \mathbf{p} = \{p_\kappa: \kappa \in \mathcal{T}_h\}\) and \(\mathbf{F} = \{F_\kappa: \kappa \in \mathcal{T}_h\}\), respectively, and consider the finite-element space

\[
\mathcal{S}_\mathcal{P}(\Omega, \mathcal{T}_h, \mathbf{F}) = \left\{ v \in L^2(\Omega): v|_{\kappa} \circ F_\kappa \in \mathcal{P}_{p_\kappa}(\widehat{\kappa}) \forall \kappa \in \mathcal{T}_h \right\},
\]

where \(\mathcal{P}\) is either \(\mathcal{P}\) or \(\mathcal{Q}\). We shall suppose that the polynomial degree vector \(\mathbf{p}\), with \(p_\kappa \geq 1\) for each \(\kappa \in \mathcal{T}_h\), has bounded local variation, i.e. there exists a constant \(\rho_2 \geq 1\) independent of \(\mathbf{h}\) and \(\mathbf{p}\) such that for any pair of neighbouring elements \(\kappa, \kappa' \in \mathcal{T}_h\),

\[
\rho_2^{-1} \leq p_\kappa / p_{\kappa'} \leq \rho_2
\]  

(2.2)

Let us consider the set \(\mathcal{E}\) of all 1D open edges or, simply, ‘edges’ of all elements \(\kappa \in \mathcal{T}_h\). Further, we denote by \(\mathcal{E}_{\text{int}}\) the set of all edges \(e\) in \(\mathcal{E}\) that are contained in \(\Omega\) (interior edges). Additionally, let \(\Gamma_{\text{int}} = \{x \in \Omega: x \in e \text{ for some } e \in \mathcal{E}_{\text{int}}\}\), and introduce \(\mathcal{E}_{\text{B}}\) to be the set of boundary edges consisting of all \(e \in \mathcal{E}\) that are contained in \(\partial \Omega\). Moreover, let \(\Gamma_{\text{int, B}} = \Gamma_{\text{int}} \cup \Gamma\).
Suppose that \( e \) is an edge of an element \( \kappa \in \mathcal{T}_h \). Then, by \( h_e \) we denote the length of \( e \). Due to our assumptions on the subdivision \( \mathcal{T}_h \), we have that if \( e \subset \partial \kappa \), then \( h_e \) is commensurate with \( h_\kappa \), the diameter of \( \kappa \).

Given that \( e \in \mathcal{E}_{\text{int}} \), there exist indices \( i \) and \( j \) such that \( i > j \) and \( \kappa_i, \kappa_j \in \mathcal{T}_h \) share the edge \( e \). We define the (element-numbering dependent) jump of an (element-wise smooth) function \( v \) across \( e \) and the mean value of \( v \) on \( e \) by

\[
[v]_e = v|_{\partial \kappa_i \cap e} - v|_{\partial \kappa_j \cap e} \quad \text{and} \quad \langle v \rangle_e = \frac{1}{2} \left( v|_{\partial \kappa_i \cap e} + v|_{\partial \kappa_j \cap e} \right),
\]

respectively. For a boundary edge \( e \subset \Gamma \) and thereby \( e \subset \partial \kappa \cap \Gamma \) for some \( \kappa \in \mathcal{T}_h \), we define

\[
[v]_e = \langle v \rangle_e = v|_{\partial \kappa \cap e}.
\]

When there is no danger of confusion, the subscript \( \cdot_e \) will be suppressed. Additionally, we associate with each edge \( e \subset \Gamma_{\text{int}} \) the unit normal vector \( \nu \) which points from \( \kappa_i \) to \( \kappa_j \) (\( i > j \)); if \( e \subset \Gamma \), then \( \nu \) is defined as the outward unit normal vector on \( \Gamma \).

With these notations and a parameter \( \theta \in [-1, 1] \), we introduce the semi-linear form

\[
B(w, v) = \int_{\Omega} \mu(|\nabla_h w|) \nabla_h w \cdot \nabla_h v \, dx
- \int_{\Gamma_{\text{int}} \cup \Gamma_{\text{bnd}}} \mu(|\nabla_h w|) \nabla_h w \cdot \nu \|v\|ds + \theta \int_{\Gamma_{\text{int}}} \mu(h^{-1}|\nabla w|) \nabla_h v \cdot \nu \|\nabla w\|ds
+ \int_{\Gamma_{\text{int}}} \sigma \|\nabla w\| \|v\|ds
\tag{2.3}
\]

and the linear functional

\[
\ell(v) = \int_{\Omega} f v \, dx. \tag{2.4}
\]

Here, \( \nabla_h \) denotes the element-wise gradient operator defined, for \( v \in H^1(\Omega, \mathcal{T}_h) \), by \( (\nabla_h v)|_\kappa = \nabla(v|_\kappa) \). For an edge \( e \in \mathcal{E} \), the discontinuity penalization parameter \( \sigma \), featuring in \( B(\cdot, \cdot) \) above, is defined by

\[
\sigma|_e = \sigma_e = \gamma \frac{\|P^2_h|_e\}}{h_e}, \tag{2.5}
\]

where \( \gamma \geq 1 \) is a (sufficiently large) constant, cf. Theorem 2.2 below.

The \( hp \)-DGFEM approximation of problem (1.1)–(1.2) reads as follows: find \( u_{\text{DG}} \in S^p(\Omega, \mathcal{T}_h, F) \) such that

\[
B(u_{\text{DG}}, v) = \ell(v) \quad \forall v \in S^p(\Omega, \mathcal{T}_h, F). \tag{2.6}
\]

Remark 2.1 In the case of an inhomogeneous Dirichlet boundary condition, \( u = g \) on \( \Gamma \), the third term in the semi-linear form \( B_{\text{DG}}(\cdot, \cdot) \) must be replaced by

\[
\theta \int_{\Gamma_{\text{int}}} \mu(h^{-1}|\nabla w|) \nabla_h v \cdot \nu \|w\|ds + \theta \int_{\Gamma} \mu(h^{-1}|w - g|) \nabla_h v \cdot n(w - g)ds,
\]
while the linear functional $\ell(\cdot)$ defined in (2.4) must be substituted by

$$\ell(v) = \int_{\Omega} fv \, dx + \int_{\Gamma} \sigma gv \, ds;$$

we refer to Houston et al. (2005) for further details.

The existence and uniqueness of the DG solution $u_{DG}$ satisfying (2.6) is guaranteed by the following result proved in Theorem 2.5 of Houston et al. (2005).

**Theorem 2.2** Suppose that $\gamma$ in (2.5) is chosen sufficiently large. Then, there exists a unique element $u_{DG}$ in $S^p(\Omega, T_h, F)$ such that (2.6) holds.

We conclude this section by equipping the DG space $S^p(\Omega, T_h, F)$ with the DG energy norm $\|\cdot\|_{DG}$ defined by

$$\|v\|_{DG} = \left( \sum_{K \in T_h} \int_K |\nabla v|^2 \, dx + \int_{\Gamma_{int,B}} \sigma \|v\|^2 \, ds \right)^{1/2} \quad (2.7)$$

induced by the energy inner product

$$(v, w)_{DG} = \sum_{K \in T_h} \int_K \nabla v \cdot \nabla w \, dx + \int_{\Gamma_{int,B}} \sigma \|v\| \|w\| \, ds.$$

The *a priori* error analysis of the DGFEM (2.6) has been developed in Houston et al. (2005); here, we shall be concerned with its *a posteriori* error analysis.

### 3. hp-Version a posteriori error analysis

Under the structural hypotheses (1.4) and (1.5) on the coefficient $\mu$, the existence of a unique solution $u \in H^1_0(\Omega)$ to (1.1)–(1.2) follows from the following result from the theory of monotone operators (see Nečas, 1986, Theorem 3.3.23) with $H = H^1_0(\Omega)$, $A = C_1$ and $\lambda = C_2$. Henceforth, we shall therefore assume that $u \in H^1_0(\Omega)$.

**Proposition 3.1** Let $H$ be a real Hilbert space with inner product $(\cdot, \cdot)_H$ and norm $\|\cdot\|_H$, and let $T$ be an operator from $H$ into itself. Suppose that $T$ is Lipschitz continuous on $H$, i.e. there exists $A > 0$ such that

$$\|T(w_1) - T(w_2)\|_H \leq A\|w_1 - w_2\|_H \quad \forall w_1, w_2 \in H,$$

and strongly monotone on $H$, i.e. there exists $\lambda > 0$ such that

$$(T(w_1) - T(w_2), w_1 - w_2)_H \geq \lambda\|w_1 - w_2\|_H^2 \quad \forall w_1, w_2 \in H.$$

Then, $T$ is a bijection of $H$ onto itself, and the inverse $T^{-1}$ of $T$ is Lipschitz continuous on $H$:

$$\|T^{-1}f - T^{-1}g\|_H \leq \lambda^{-1}\|f - g\|_H \quad \forall f, g \in H.$$
3.1 Upper bound

In this section, we will formulate the main result of this paper, Theorem 3.2. To this end, we first define, for an element $\kappa \in \mathcal{T}_h$ and an edge $e \in \mathcal{E}_{\text{int}}$, the data-oscillation terms

$$\mathcal{O}_k^{(1)} = h_k^2 p_k^{-2} \left\| (1 - \Pi_{\mathcal{T}_h}) (f + \nabla \cdot (\mu (|\nabla u_{DG}|) \nabla u_{DG})) \right\|_{0,\kappa}^2$$

and

$$\mathcal{O}_e^{(2)} = h_e p_e^{-1} \left\| (1 - \Pi_{\mathcal{E}_e}) (\mu (|\nabla u_{DG}|) \nabla u_{DG}) \cdot \psi \right\|_{0,e}^2,$$  

(3.1)

which depend on the right-hand side $f$ in (1.1) and the numerical solution $u_{DG}$ from (2.6). Here, $I$ is a generic identity operator and $\Pi_{\mathcal{T}_h}$ denotes the element-wise $L^2$-projector onto the space $S^{p-1}(\Omega, \mathcal{T}_h, \mathbf{F})$, where $p = 1 = (p_\kappa - 1)_{\kappa \in \mathcal{T}_h}$. Furthermore, $\Pi_{\mathcal{E}_e}$ is defined as the $L^2$-projector onto $\mathcal{P}_{p_e-1}(e)$; here, $p_e = \max\{p_\kappa, p_{\kappa'}\}$ with $\kappa, \kappa' \in \mathcal{T}_h$, $e = e_\kappa \cap e_{\kappa'}$ (we note that due to our assumptions on the polynomial degree vector $p$, the quantities $p_e, p_\kappa$ and $p_{\kappa'}$ are all commensurate with one another).

THEOREM 3.2 Let the analytical solution $u$ of (1.1)–(1.2) belong to $H^1_0(\Omega)$. Furthermore, let $u_{DG} \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ be its discontinuous Galerkin (DG) approximation, i.e. the solution of (2.6). Then, the following $hp$-version a posteriori error bound holds:

$$\|u - u_{DG}\|_{DG} \leq C \left( \sum_{\kappa \in \mathcal{T}_h} \eta_{\kappa}^2 + \mathcal{O}(f, u_{DG}) \right)^{\frac{1}{2}},$$

(3.3)

where the local error indicators $\eta_{\kappa}, \kappa \in \mathcal{T}_h$, are defined by

$$\eta_{\kappa}^2 = h_k^2 p_k^{-2} \left\| \Pi_{\mathcal{T}_h} (f + \nabla \cdot (\mu (|\nabla u_{DG}|) \nabla u_{DG})) \right\|_{0,\kappa}^2$$

$$+ h_k p_k^{-1} \left\| \Pi_{\mathcal{E}_e} (\mu (|\nabla u_{DG}|) \nabla u_{DG} \cdot \psi) \right\|_{0,\partial \kappa \cap \Gamma}^2 + \gamma^2 h_k^{-1} p_k^3 \|u_{DG}\|_{0,\partial \kappa}^2,$$

(3.4)

and

$$\mathcal{O}(f, u_{DG}) = \sum_{\kappa \in \mathcal{T}_h} \mathcal{O}_k^{(1)} + \sum_{e \in \mathcal{E}_{\text{int}}} \mathcal{O}_e^{(2)}.$$

Here, $C > 0$ is a constant that is independent of $h$, the polynomial degree vector $p$ and the parameters $\gamma$ and $\theta$ and only depends on the shape regularity of the mesh and the constants $\rho_1$ and $\rho_2$ from (2.1) and (2.2), respectively.

REMARK 3.3 We observe a slight suboptimality with respect to the polynomial degree in the last term of the local error indicator $\eta_{\kappa}$ in (3.4). This results from the fact that due to the possible presence of hanging nodes in $\mathcal{T}_h$, a non-conforming interpolant is used in the proof of the above Theorem 3.2, cf. Section 3.1.2. Indeed, for conforming (regular) meshes, i.e. meshes without any hanging nodes, a conforming ($hp$-version) Clément interpolant, as constructed in Melenk (2005), can be employed. This then results in an a posteriori error bound of the form (3.3) with the term $\gamma^2 h_k^{-1} p_k^3 \|u_{DG}\|_{0,\partial \kappa}^2$ in (3.4) replaced by the improved expression $\gamma h_k^{-1} p_k^3 \|u_{DG}\|_{0,\partial \kappa}^2$, cf. Houston et al. (2007b).

REMARK 3.4 In order to incorporate the inhomogeneous boundary condition $u = g$ on $\Gamma$, the error indicators $\eta_{\kappa}$ are simply adjusted by modifying the jump indicators $\mathcal{O}_e^{(2)} = \frac{1}{2} \|u_{DG}\|_{0,\partial \kappa}^2$ on $\partial \kappa \cap \Gamma$, with the inclusion of an additional data-oscillation term, see Houston et al. (2007b) for details.
3.1.1 DG decompositions. The hp-version a posteriori error analysis for the DGFEM (2.6) will be based on an approach similar to the one discussed in Houston et al. (2007b) (see also Houston et al. (2004); Houston et al. (2005, 2007a) and Wihler (2006), for related work). In contrast to the analysis presented in Houston et al. (2007b) though, here we shall also admit one-irregular meshes containing hanging nodes. To this end, consider a given subdivision \( \mathcal{H}_h \) of \( \Omega \) that is regularly reducible, i.e. \( \mathcal{H}_h \) can be refined to a shape-regular conforming mesh \( \mathcal{H}_h \) as described in Section 2. Furthermore, denote by \( S^p(\Omega, \mathcal{H}_h, \tilde{\mathbf{F}}) \) the corresponding DG space with a suitable affine element mapping vector \( \tilde{\mathbf{F}} \) and a polynomial degree vector \( \tilde{\mathbf{p}} \) that is defined by \( p_{\tilde{k}} = p_k \), for any \( \tilde{k} \in \mathcal{H}_h \) with \( \tilde{k} \subseteq k \) and for some \( k \in \mathcal{H}_h \).

We note that \( S^p(\Omega, \mathcal{H}_h, \tilde{\mathbf{F}}) \subseteq S^p(\Omega, \mathcal{H}, \tilde{\mathbf{F}}) \) and that due to our assumptions in Section 2 (specifically, the commensurability of the local element sizes and the local polynomial degrees in \( \mathcal{H}_h \) and \( \mathcal{H} \) due to our bounded local variation assumptions), the DG energy norms \( \| \cdot \|_{DG} \) and \( \| \cdot \|_{DG} \) corresponding to the DG spaces \( S^p(\Omega, \mathcal{H}_h, \tilde{\mathbf{F}}) \) and \( S^p(\Omega, \mathcal{H}, \tilde{\mathbf{F}}) \), respectively, are equivalent on \( S^p(\Omega, \mathcal{H}_h, \tilde{\mathbf{F}}) \); in particular, there exist positive constants \( N_1 \) and \( N_2 \), independent of \( \mathbf{h} \) and \( \mathbf{p} \), such that

\[
N_1 \int_{I_{int,B}} \sigma \|v\|^2 \, ds \leq \int_{I_{int,B}} \tilde{\sigma} \|v\|^2 \, ds \leq N_2 \int_{I_{int,B}} \sigma \|v\|^2 \, ds \quad \forall v \in S^p(\Omega, \mathcal{H}_h, \tilde{\mathbf{F}}). \tag{3.5}
\]

Here, \( I_{int,B} \) denotes the union of all element edges of \( \mathcal{H}_h \) and \( \tilde{\sigma} \) is the discontinuity penalization parameter on \( S^p(\Omega, \mathcal{H}_h, \tilde{\mathbf{F}}) \) which is defined analogously as for \( S^p(\Omega, \mathcal{H}_h, \tilde{\mathbf{F}}) \) in (2.5); note that for \( v \in S^p(\Omega, \mathcal{H}_h, \tilde{\mathbf{F}}) \), the jump \( \|v\| \) vanishes on \( I_{int,B} \). An important step in our analysis is the decomposition of the DG space \( S^p(\Omega, \mathcal{H}_h, \tilde{\mathbf{F}}) \) into two orthogonal subspaces: a conforming part \( [S^p(\Omega, \mathcal{H}_h, \tilde{\mathbf{F}})]^\| = S^p(\Omega, \mathcal{H}_h, \tilde{\mathbf{F}}) \cap H_0^1(\Omega) \) and a non-conforming part \([S^p(\Omega, \mathcal{H}_h, \tilde{\mathbf{F}})]^\perp \) defined as the orthogonal complement of \([S^p(\Omega, \mathcal{H}_h, \tilde{\mathbf{F}})]^\| \) in \( S^p(\Omega, \mathcal{H}_h, \tilde{\mathbf{F}}) \) with respect to the DG energy inner product \( (\cdot, \cdot)_{DG} \) (inducing the DG energy norm \( \| \cdot \|_{DG} \)), i.e.

\[
S^p(\Omega, \mathcal{H}_h, \tilde{\mathbf{F}}) = [S^p(\Omega, \mathcal{H}_h, \tilde{\mathbf{F}})]^\| \oplus \| \cdot \|_{DG} [S^p(\Omega, \mathcal{H}_h, \tilde{\mathbf{F}})]^\perp.
\]

Based on this setting, the DG solution \( u_{DG} \) obtained by (2.6) may be split accordingly,

\[
u_{DG} = u^\|_{DG} + u^\perp_{DG}, \tag{3.6}
\]

where \( u^\|_{DG} \in [S^p(\Omega, \mathcal{H}_h, \tilde{\mathbf{F}})]^\| \) and \( u^\perp_{DG} \in [S^p(\Omega, \mathcal{H}_h, \tilde{\mathbf{F}})]^\perp \). Furthermore, we define the error of the \( hp \)-DGFEM by

\[
e_{DG} = u - u_{DG}, \tag{3.7}
\]

and let

\[
e_{DG} = u - u_{DG} \in H_0^1(\Omega). \tag{3.8}
\]

3.1.2 Auxiliary results. For the proof of the above Theorem 3.2, we shall require some auxiliary results.

**Proposition 3.5** Under the assumptions in Section 2 on the (regularly reduced) subdivision \( \mathcal{H}_h \), the following norm equivalence holds over the space \([S^p(\Omega, \mathcal{H}_h, \tilde{\mathbf{F}})]^\perp\):

\[
\tilde{C}_1 \|v\|^2_{DG} \leq \int_{I_{int,B}} \tilde{\sigma} \|v\|^2 \, ds \leq \tilde{C}_2 \|v\|^2_{DG} \quad \forall v \in [S^p(\Omega, \mathcal{H}_h, \tilde{\mathbf{F}})]^\perp, \tag{3.9}
\]
where the constants $\tilde{C}_1, \tilde{C}_2 > 0$ depend only on the shape regularity of $\mathcal{T}_h$ and the constants $\rho_1$ and $\rho_2$ in (2.1) and (2.2), respectively.

**Proof.** See Houston et al. (2007b, Proposition 4.5).

**Corollary 3.6** With $u_{DG}^\perp$ and $e_{DG}^\parallel$ defined by (3.6) and (3.8), respectively, the following bounds hold:

$$
\|u_{DG}^\perp\|_{DG} \leq D_1 \left( \int_{\Gamma_{int,B}} \sigma \|u_{DG}\|^2 \, ds \right)^{\frac{1}{2}}, \quad \|e_{DG}^\parallel\|_{DG} \leq D_2 \|e_{DG}\|_{DG},
$$

where the constants $D_1, D_2 > 0$ are independent of $\gamma, h$ and $p$ and only depend on the shape regularity of $\mathcal{T}_h$ and the constants $\rho_1$ and $\rho_2$ in (2.1) and (2.2), respectively.

**Proof.** In order to prove the first of the above bounds, we recall that $u_{DG}^\parallel \in H_0^1(\Omega)$. This implies that $\|u_{DG}\| = 0$ on $\Gamma_{int,B}$, and hence,

$$
\|u_{DG}^\perp\| = \|u_{DG}\| + \|u_{DG}^\parallel\| = \|u_{DG}^\parallel\| = \|u_{DG}\|.
$$

Then, due to Proposition 3.5, we obtain

$$
\|u_{DG}^\perp\|_{DG}^2 \leq C \int_{\Gamma_{int,B}} \tilde{\sigma} \|u_{DG}^\parallel\|^2 \, ds = C \int_{\Gamma_{int,B}} \tilde{\sigma} \|u_{DG}\|^2 \, ds. \tag{3.10}
$$

Furthermore, since $u_{DG} \in S^p(\Omega, \mathcal{T}_h, F)$ and because of (3.5), we conclude that

$$
\|u_{DG}^\parallel\|_{DG}^2 \leq C \int_{\Gamma_{int,B}} \sigma \|u_{DG}\|^2 \, ds.
$$

For the second bound, we use the triangle inequality, the bound (3.10) and the fact that since the analytical solution $u$ of (1.1)–(1.2) and $e_{DG}$ belong to $H_0^1(\Omega)$, we have

$$
\|u\| = \|e_{DG}^\parallel\| = 0 \quad \text{and} \quad \|e_{DG}\| = \|u\| - \|u_{DG}\| = -\|u_{DG}\| \tag{3.11}
$$
on $\Gamma_{int,B}$ (and thereby also on $\Gamma_{int,B}$). Thus,

$$
\|e_{DG}\|_{DG} = \|e_{DG}\|_{DG} \leq \|e_{DG}\|_{DG} + \|u_{DG}\|_{DG} \leq \|e_{DG}\|_{DG} + C \left( \int_{\Gamma_{int,B}} \tilde{\sigma} \|u_{DG}\|^2 \, ds \right)^{\frac{1}{2}}
$$

$$
\|e_{DG}\|_{DG} \leq \|e_{DG}\|_{DG} + C \left( \int_{\Gamma_{int,B}} \tilde{\sigma} \|e_{DG}\|^2 \, ds \right)^{\frac{1}{2}} \leq C \|e_{DG}\|_{DG}. \tag{3.12}
$$

In a similar way, we obtain

$$
\|e_{DG}\|_{DG}^2 = \sum_{\kappa \in \mathcal{T}_h} \|\nabla e_{DG}\|_{0,\kappa}^2 + \int_{\Gamma_{int,B}} \tilde{\sigma} \|e_{DG}\|^2 \, ds = \sum_{\kappa \in \mathcal{T}_h} \|\nabla e_{DG}\|_{0,\kappa}^2 + \int_{\Gamma_{int,B}} \tilde{\sigma} \|u_{DG}\|^2 \, ds.
$$
Moreover, observing that \( u_{DG} \in S^p(\Omega, \mathcal{H}, F) \) and applying (3.5) lead to

\[
\|e_{DG}\|_{DG}^2 \leq \sum_{\kappa \in \mathcal{H}} \|\nabla e_{DG}\|_{0,K}^2 + C \int_{\Gamma_{\text{int},B}} \sigma \|u_{DG}\|^2 \, ds
\]

\[
= \sum_{\kappa \in \mathcal{H}} \|\nabla e_{DG}\|_{0,K}^2 + C \int_{\Gamma_{\text{int},B}} \sigma \|e_{DG}\|^2 \, ds \leq C \|e_{DG}\|_{DG}^2,
\]

which, referring to (3.12), yields the second bound.

Next, we state the following approximation property.

**Lemma 3.7** For any \( \varphi \in H^1_0(\Omega) \), there exists a function \( \varphi_{hp} \in S^p(\Omega, \mathcal{H}, F) \) such that

\[
h^{-2}_\kappa p_{\kappa}^2 \|\varphi - \varphi_{hp}\|_{0,K}^2 + \|\nabla (\varphi - \varphi_{hp})\|_{0,K}^2 + h^{-1}_\kappa p_{\kappa} \|\varphi - \varphi_{hp}\|_{0,\partial K}^2 \leq C_1 \|\nabla \varphi\|_{0,K}^2,
\]

for any \( \kappa \in \mathcal{H} \), with an interpolation constant \( C_1 > 0 \), which is independent of \( h \) and \( p \) and only depends on the shape regularity of the mesh and the constants \( p_1 \) and \( p_2 \) in (2.1) and (2.2), respectively.

**Proof.** We first consider the proof of the upper bounds on the \( L^2(\kappa) \)-norms of \( \varphi - \varphi_{hp} \) and \( \nabla (\varphi - \varphi_{hp}) \). In this case, on quadrilateral elements, the above approximation property follows from the tensorization of the corresponding 1D approximation results for an \( H^1 \)-projector, see Houston et al. (2000), for details. For triangular elements, we employ a reflection technique. More precisely, writing \( \hat{\kappa} \) to denote the canonical triangle with vertices \((-1,-1), (1,-1) \) and \((-1,1)\), we define \( \hat{\kappa}' \) to be the triangle with vertices \((1,-1), (1,1) \) and \((-1,1)\) obtained by reflecting \( \hat{\kappa} \) about its longest edge. Analogously, given \( \hat{\nu} \in H^1(\hat{\kappa}) \), we write \( \hat{\nu}' \in H^1(\hat{\kappa}') \) to denote the reflection of \( \nu \) in the line \( \xi_2 = -\xi_1 \), where \((\xi_1, \xi_2)\) denotes the local coordinate system for the reference element \( \hat{\kappa} \). With this notation, we define the function \( \nu \in H^1(\hat{S}) \) by \( \nu|_{\hat{\kappa}} = \hat{\nu} \) and \( \nu|_{\hat{\kappa}'} = \hat{\nu}' \), where \( \hat{S} \) is the unit square \((-1,1)^2 \). Due to symmetry, we deduce that there exists a positive constant \( C \) such that

\[
\sqrt{2} \|\hat{\nu}\|_{0,\hat{\kappa}} \leq \|\hat{\nu}\|_{0,\hat{S}} \leq C \|\hat{\nu}\|_{0,\hat{\kappa}} \quad \text{and} \quad \sqrt{2} \|\nabla \hat{\nu}\|_{0,\hat{\kappa}} \leq \|\nabla \hat{\nu}\|_{0,\hat{S}} \leq C \|\nabla \hat{\nu}\|_{0,\hat{\kappa}}.
\]

Thereby, the approximation properties on the reference element \( \hat{\kappa} \) now follow from the corresponding results on the unit square \( \hat{S} \); the proof is then completed by employing a standard scaling argument.

The upper bound on the approximation error measured in terms of the \( L^2(\hat{\kappa}) \)-norm now follows from the above results together with the trace inequality

\[
\|\nu\|_{0,\hat{\kappa}}^2 \leq C(\|\nabla \nu\|_{0,K} \|\nu\|_{0,K} + h^{-1}_\kappa \|\nu\|_{0,\hat{\kappa}}^2),
\]

where \( \nu \in H^1(\kappa) \) and \( C \) is a positive constant which depends only on the shape regularity of \( \kappa \).
Proof of Theorem 3.2. We commence the proof of our main theorem by recalling the definition of the \(\|\cdot\|_{DG}\)-norm (2.7) and by applying (1.5). This yields

\[
C_2\|e_{DG}\|_{DG}^2 = C_2 \left( \sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} |\nabla u - \nabla u_{DG}|^2 \, dx + \int_{\Gamma_{int,B}} \sigma \|e_{DG}\|^2 \, ds \right)
\]

\[
= C_2 \left( \sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} |\nabla u - \nabla u_{DG}|^2 \, dx + \int_{\Gamma_{int,B}} \sigma \|e_{DG}\|^2 \, ds \right)
\]

\[
\leq \sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} (\mu(|\nabla u|) \nabla u - \mu(|\nabla u_{DG}|) \nabla u_{DG}) \cdot \nabla e_{DG} \, dx + C_2 \int_{\Gamma_{int,B}} \sigma \|e_{DG}\|^2 \, ds.
\]

Next, we split the above integral expressions into three parts, i.e.

\[
C_2\|e_{DG}\|_{DG}^2 \leq T_1 + T_2 + T_3,
\]

with

\[
T_1 = \sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} (\mu(|\nabla u|) \nabla u - \mu(|\nabla u_{DG}|) \nabla u_{DG}) \cdot \nabla e_{DG}^\parallel \, dx,
\]

\[
T_2 = - \sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} (\mu(|\nabla u|) \nabla u - \mu(|\nabla u_{DG}|) \nabla u_{DG}) \cdot u_{DG}^\perp \, dx,
\]

\[
T_3 = C_2 \int_{\Gamma_{int,B}} \sigma \|e_{DG}\|^2 \, ds.
\]

Here, \(e_{DG}^\parallel \in H^1_0(\Omega)\) and \(u_{DG}^\perp \in [S^\parallel(\Omega, \mathcal{T}_h), \Gamma]\) are defined as in (3.6) and (3.8), respectively.

We will now analyse the three terms \(T_1\), \(T_2\) and \(T_3\) separately. Note that due to the possible low regularity of the analytical solution \(u \in H^1_0(\Omega)\), the analysis in Houston et al. (2007b) considers the introduction of a generalized DG formulation through the use of lifting operators. In contrast, in the current article, we propose the splitting (3.14) of the error, which is chosen in such a way that the analysis can proceed under the same regularity conditions, however, without the need of an auxiliary formulation.

Term \(T_1\): We first note that

\[
T_1 = \sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} (\mu(|\nabla u|) \nabla u - \mu(|\nabla u_{DG}|) \nabla u_{DG}) \cdot \nabla e_{DG}^\parallel \, dx.
\]

Then, using integration by parts, we obtain

\[
T_1 = - \sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} \nabla \cdot (\mu(|\nabla u|) \nabla u) e_{DG}^\parallel \, dx - \sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} \mu(|\nabla u_{DG}|) \nabla u_{DG} \cdot \nabla e_{DG}^\parallel \, dx
\]

\[
= \sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} f e_{DG}^\parallel \, dx - \sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} \mu(|\nabla u_{DG}|) \nabla u_{DG} \cdot \nabla e_{DG}^\parallel \, dx.
\]
Therefore,

\[
T_1 = \sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} f(e^\parallel_{DG} - \varphi_{hp}) \, dx - \sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} \mu(|\nabla u_{DG}|) \nabla u_{DG} \cdot \nabla (e^\parallel_{DG} - \varphi_{hp}) \, dx
\]

\[
- \int_{\Gamma_{int,B}} \langle \mu(|\nabla h u_{DG}|) \nabla h u_{DG} \cdot \nu \rangle \parallel \varphi_{hp} \parallel ds + \int_{\Gamma_{int,B}} \langle \mu(h^{-1}||u_{DG}||) \nabla h \varphi_{hp} \cdot \nu \rangle \parallel u_{DG} \parallel ds
\]

\[
+ \int_{\Gamma_{int,B}} \sigma \parallel u_{DG} \parallel \parallel \varphi_{hp} \parallel ds.
\]

Hence, integrating the second term on the right-hand side of the above equality by parts leads to

\[
T_1 = \sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} (f + \nabla \cdot (\mu(|\nabla u_{DG}|) \nabla u_{DG}))(e^\parallel_{DG} - \varphi_{hp}) \, dx
\]

\[
- \sum_{\kappa \in \mathcal{T}_h} \int_{\partial \kappa} \langle \mu(|\nabla u_{DG}|) \nabla u_{DG} \cdot n \rangle (e^\parallel_{DG} - \varphi_{hp}) \, ds - \int_{\Gamma_{int,B}} \langle \mu(|\nabla h u_{DG}|) \nabla h u_{DG} \cdot \nu \rangle \parallel \varphi_{hp} \parallel ds
\]

\[
+ \int_{\Gamma_{int,B}} \langle \mu(h^{-1}||u_{DG}||) \nabla h \varphi_{hp} \cdot \nu \rangle \parallel u_{DG} \parallel ds + \int_{\Gamma_{int,B}} \sigma \parallel u_{DG} \parallel \parallel \varphi_{hp} \parallel ds.
\]

Using the fact that \(e^\parallel_{DG} = 0\) on \(\Gamma_{int,B}\), since \(e^\parallel_{DG} \in H^1_0(\Omega)\), and a few elementary calculations, we have that

\[
- \sum_{\kappa \in \mathcal{T}_h} \int_{\partial \kappa} \langle \mu(|\nabla u_{DG}|) \nabla u_{DG} \cdot n \rangle (e^\parallel_{DG} - \varphi_{hp}) \, ds = - \int_{\Gamma_{int}} \langle \mu(|\nabla h u_{DG}|) \nabla h u_{DG} \cdot \nu \rangle \parallel e^\parallel_{DG} - \varphi_{hp} \parallel ds
\]

\[
+ \int_{\Gamma_{int,B}} \langle \mu(|\nabla h u_{DG}|) \nabla h u_{DG} \cdot \nu \rangle \parallel \varphi_{hp} \parallel ds.
\]

Therefore,

\[
T_1 = \sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} (f + \nabla \cdot (\mu(|\nabla u_{DG}|) \nabla u_{DG}))(e^\parallel_{DG} - \varphi_{hp}) \, dx
\]

\[
- \int_{\Gamma_{int}} \parallel \mu(|\nabla h u_{DG}|) \nabla h u_{DG} \cdot \nu \parallel \parallel e^\parallel_{DG} - \varphi_{hp} \parallel ds
\]

\[
+ \theta \int_{\Gamma_{int,B}} \parallel \mu(h^{-1}||u_{DG}||) \nabla h \varphi_{hp} \cdot \nu \parallel \parallel u_{DG} \parallel ds + \int_{\Gamma_{int,B}} \sigma \parallel u_{DG} \parallel \parallel \varphi_{hp} \parallel ds,
\]
and thus,

\[ T_1 \leq \sum_{\kappa \in T_h} \int_{\kappa} |f + \nabla \cdot (\mu(|\nabla u_{DG}|) \nabla u_{DG})| e_{DG}^{\parallel} - \varphi_{hp} | dx \\
+ \sum_{\kappa \in T_h} \int_{\partial \kappa \setminus \Gamma} \left[ \| \mu(|\nabla u_{DG}|) \nabla u_{DG} \cdot \nu \| \right] e_{DG}^{\parallel} - \varphi_{hp} | ds \\
+ |\theta| \int_{\Gamma_{\text{int},B}} h \mu(h^{-1}||u_{DG}||)(h^{-1}||u_{DG}||)|\nabla h_{\varphi_{hp}} \cdot \nu || ds + \int_{\Gamma_{\text{int},B}} \sigma \|u_{DG}\||\varphi_{hp}|| ds \\
\leq \sum_{\kappa \in T_h} \| f + \nabla \cdot (\mu(|\nabla u_{DG}|) \nabla u_{DG}) \|_{0,\kappa} e_{DG}^{\parallel} - \varphi_{hp} \|_{0,\kappa} \\
+ C \sum_{\kappa \in T_h} \| \mu(|\nabla u_{DG}|) \nabla u_{DG} \cdot \nu \|_{0,\partial \kappa \setminus \Gamma} e_{DG}^{\parallel} - \varphi_{hp} \|_{0,\partial \kappa} \\
+ M_\mu |\theta| \int_{\Gamma_{\text{int},B}} \| \nabla h_{\varphi_{hp}} \| ds + \int_{\Gamma_{\text{int},B}} \sigma \|u_{DG}\||\varphi_{hp}|| ds, \]

where we have applied (1.3) (with \( s = 0 \) and \( t = h^{-1}||u_{DG}|| \)) to bound the second last of the above terms. Moreover, proceeding as in the proof of Lemma 2.2 of Houston et al. (2005) (cf. also Wihler et al., 2003, Lemma 3.5) and recalling that \( \gamma \geq 1 \), we obtain

\[
\int_{\Gamma_{\text{int},B}} \| u_{DG} \| \langle \nabla h_{\varphi_{hp}} \rangle \| ds \leq \left( \int_{\Gamma_{\text{int},B}} \sigma \| u_{DG} \|^2 \| ds \right)^{\frac{1}{2}} \left( \sum_{e \in \mathcal{E}} h^{-1} \| p_2 \| \| \nabla h_{\varphi_{hp}} \|^2 \| ds \right)^{\frac{1}{2}} \\
\leq C \left( \int_{\Gamma_{\text{int},B}} \sigma \| u_{DG} \|^2 \| ds \right)^{\frac{1}{2}} \left( \sum_{\kappa \in T_h} \| \nabla h_{\varphi_{hp}} \|^2_{0,\kappa} \right)^{\frac{1}{2}}.
\]

Furthermore, by (3.13), we have

\[
\sum_{\kappa \in T_h} \| \nabla h_{\varphi_{hp}} \|^2_{0,\kappa} \leq C \sum_{\kappa \in T_h} \| \nabla (e_{DG}^{\parallel} - \varphi_{hp}) \|^2_{0,\kappa} + C \sum_{\kappa \in T_h} \| \nabla e_{DG}^{\parallel} \|^2_{0,\kappa} \leq C \sum_{\kappa \in T_h} \| \nabla e_{DG}^{\parallel} \|^2_{0,\kappa},
\]

and hence,

\[
\int_{\Gamma_{\text{int},B}} \| u_{DG} \| \langle \nabla h_{\varphi_{hp}} \rangle \| ds \leq C \left( \int_{\Gamma_{\text{int},B}} \sigma \| u_{DG} \|^2 \| ds \right)^{\frac{1}{2}} \left( \sum_{\kappa \in T_h} \| \nabla e_{DG}^{\parallel} \|^2_{0,\kappa} \right)^{\frac{1}{2}}.
\]
Moreover, using again the fact that \( \| e_{DG}^\| \) = 0 on \( F_{\text{int},B} \) and recalling (2.1)–(2.2) imply

\[
\int_{F_{\text{int},B}} \sigma \| u_{DG} \| \| \varphi_{hp} \| ds = \int_{F_{\text{int},B}} \sigma \| u_{DG} \| \| e_{DG} - \varphi_{hp} \| ds
\]

\[
\leq C \left( \int_{F_{\text{int},B}} \sigma \langle p \rangle \| u_{DG} \|^2 ds \right)^{\frac{1}{2}} \left( \sum_{k \in \mathcal{T}_h} \int_{\partial_k} \sigma \langle p \rangle^{-1} \| e_{DG} - \varphi_{hp} \|^2 ds \right)^{\frac{1}{2}}
\]

\[
\leq C \gamma^{\frac{1}{2}} \left( \int_{F_{\text{int},B}} \sigma \langle p \rangle \| u_{DG} \|^2 ds \right)^{\frac{1}{2}} \left( \sum_{k \in \mathcal{T}_h} h_k^{-1} p_k \| e_{DG} - \varphi_{hp} \|^2_{0,\partial_k} \right)^{\frac{1}{2}}
\]

Thus, collecting the terms leads to

\[
T_1 \leq \sum_{k \in \mathcal{T}_h} h_k p_k^{-1} \| f + \nabla \cdot (\mu (| \nabla u_{DG} |) \nabla u_{DG}) \|_{0,k} h_k^{-1} p_k \| e_{DG}^\| - \varphi_{hp} \|_{0,k}
\]

\[
+ C \sum_{k \in \mathcal{T}_h} h_k \gamma \frac{1}{2} \| \mu (| \nabla u_{DG} |) \nabla u_{DG} \cdot v \|_{0,\partial_k} \Gamma h_k^{-\frac{1}{2}} p_k \gamma \frac{1}{2} \| e_{DG}^\| - \varphi_{hp} \|_{0,\partial_k}
\]

\[
+ C|\theta| \left( \int_{F_{\text{int},B}} \sigma \| u_{DG} \|^2 ds \right)^{\frac{1}{2}} \left( \sum_{k \in \mathcal{T}_h} \| \nabla e_{DG}^\|_{0,\partial_k} \right)^{\frac{1}{2}}
\]

\[
+ C \gamma \frac{1}{2} \left( \int_{F_{\text{int},B}} \sigma \langle p \rangle \| u_{DG} \|^2 ds \right)^{\frac{1}{2}} \left( \sum_{k \in \mathcal{T}_h} h_k^{-1} p_k \| e_{DG}^\| - \varphi_{hp} \|_{0,\partial_k} \right)^{\frac{1}{2}}
\]

Furthermore, applying again the approximation property (3.13), using that \( \gamma \geq 1 \geq |\theta| \geq 0 \) and incorporating (2.2) result in

\[
T_1 \leq \left( \sum_{k \in \mathcal{T}_h} h_k^2 p_k^{-2} \| f + \nabla \cdot (\mu (| \nabla u_{DG} |) \nabla u_{DG}) \|_{0,k} \right)^{\frac{1}{2}} \left( \sum_{k \in \mathcal{T}_h} h_k^{-2} p_k^2 \| e_{DG}^\| - \varphi_{hp} \|_{0,\partial_k} \right)^{\frac{1}{2}}
\]

\[
+ C \left( \sum_{k \in \mathcal{T}_h} h_k p_k^{-1} \| \mu (| \nabla u_{DG} |) \nabla u_{DG} \cdot v \|_{0,\partial_k} \right)^{\frac{1}{2}} \left( \sum_{k \in \mathcal{T}_h} h_k^{-1} p_k \| e_{DG}^\| - \varphi_{hp} \|_{0,\partial_k} \right)^{\frac{1}{2}}
\]

\[
+ C|\theta| \left( \sum_{k \in \mathcal{T}_h} \sigma \| u_{DG} \|^2 ds \right)^{\frac{1}{2}} \left( \sum_{k \in \mathcal{T}_h} \| \nabla e_{DG}^\|_{0,\partial_k} \right)^{\frac{1}{2}}
\]

\[
+ C \gamma \frac{1}{2} \left( \sum_{k \in \mathcal{T}_h} \sigma \langle p \rangle \| u_{DG} \|^2 ds \right)^{\frac{1}{2}} \left( \sum_{k \in \mathcal{T}_h} h_k^{-1} p_k \| e_{DG}^\| - \varphi_{hp} \|_{0,\partial_k} \right)^{\frac{1}{2}}
\]
\[
\begin{align*}
&\leq C \left( \sum_{\kappa \in \mathcal{T}_h} h_k^2 p_k^{-2} \| f + \nabla \cdot (\mu(|\nabla u_{DG}|) \nabla u_{DG}) \|_{0,\kappa}^2 \right)^{\frac{1}{2}} \left( \sum_{\kappa \in \mathcal{T}_h} \| e_{DG} \|_{0,\kappa}^2 \right)^{\frac{1}{2}} \\
&\quad + C \left( \sum_{\kappa \in \mathcal{T}_h} h_k p_k^{-1} \| \mu(|\nabla u_{DG}|) \nabla u_{DG} \cdot \mathbf{v} \|_{0,\partial \kappa \setminus \Gamma}^2 \right)^{\frac{1}{2}} \left( \sum_{\kappa \in \mathcal{T}_h} \| e_{DG} \|_{0,\kappa}^2 \right)^{\frac{1}{2}} \\
&\quad + C \left( \gamma^2 \sum_{\kappa \in \mathcal{T}_h} h_k^{-1} p_k^3 \| u_{DG} \|_{0,\partial \kappa}^2 \right)^{\frac{1}{2}} \left( \sum_{\kappa \in \mathcal{T}_h} \| e_{DG} \|_{0,\kappa}^2 \right)^{\frac{1}{2}}.
\end{align*}
\]

Therefore,

\[ T_1 \leq C \left( \sum_{\kappa \in \mathcal{T}_h} \tilde{\eta}_k^2 \right)^{\frac{1}{2}} \| e_{DG} \|_{DG}. \]

which, by Corollary 3.6, yields

\[ T_1 \leq C \left( \sum_{\kappa \in \mathcal{T}_h} \tilde{\eta}_k^2 \right)^{\frac{1}{2}} \| e_{DG} \|_{DG}. \]

Here, for \( \kappa \in \mathcal{T}_h \), the term \( \tilde{\eta}_k \) is defined by

\[ \tilde{\eta}_k^2 = h_k^2 p_k^{-2} \| f + \nabla \cdot (\mu(|\nabla u_{DG}|) \nabla u_{DG}) \|_{0,\kappa}^2 + h_k p_k^{-1} \| \mu(|\nabla u_{DG}|) \nabla u_{DG} \cdot \mathbf{v} \|_{0,\partial \kappa \setminus \Gamma}^2 \]

\[ + \gamma^2 h_k^{-1} p_k^3 \| u_{DG} \|_{0,\partial \kappa}^2. \]

Observing that

\[ \tilde{\eta}_k^2 \leq C \left( \eta_k^2 + \theta_k^{(1)} + \sum_{e \in \mathcal{E}_{\text{int}}} \theta_e^{(2)} \right), \]

we obtain

\[ T_1 \leq C \left( \sum_{\kappa \in \mathcal{T}_h} \left( \eta_k^2 + \theta_k^{(1)} + \sum_{e \in \mathcal{E}_{\text{int}}} \theta_e^{(2)} \right) \right)^{\frac{1}{2}} \| e_{DG} \|_{DG}. \quad (3.15) \]
Term $T_2$: In order to bound $T_2$, we recall (1.4). This yields

$$T_2 \leq \sum_{\kappa \in \mathcal{F}_h} \int_{\Gamma_{\kappa}} |\mu(|\nabla u|) \nabla u - \mu(|\nabla u_{DG}|) \nabla u_{DG}| \nabla u_{DG}^\perp \, dx$$

$$\leq C_1 \sum_{\kappa \in \mathcal{F}_h} \int_{\Gamma_{\kappa}} |\nabla e_{DG}| \nabla u_{DG}^\perp \, dx \leq C_1 \sum_{\kappa \in \mathcal{F}_h} \|\nabla e_{DG}\|_{0,\kappa} \|\nabla u_{DG}^\perp\|_{0,\kappa}$$

$$\leq C_1 \left( \sum_{\kappa \in \mathcal{F}_h} \|\nabla e_{DG}\|_{0,\kappa}^2 \right)^{\frac{1}{2}} \left( \sum_{\kappa \in \mathcal{F}_h} \|\nabla u_{DG}^\perp\|_{0,\kappa}^2 \right)^{\frac{1}{2}}.$$

Hence, we have

$$T_2 \leq C_1 \|e_{DG}\|_{DG} \|u_{DG}^\perp\|_{DG},$$

which, upon applying Corollary 3.6, gives

$$T_2 \leq C \|e_{DG}\|_{DG} \left( \int_{\Gamma_{\text{int},B}} \sigma \|u_{DG}\|^2 \, ds \right)^{\frac{1}{2}} \leq C \|e_{DG}\|_{DG} \left( \gamma \sum_{\kappa \in \mathcal{F}_h} h_{\kappa}^{-1} p_{\kappa}^2 \|u_{DG}\|_{0,\partial\kappa}^2 \right)^{\frac{1}{2}},$$

and thus since $\gamma \geq 1$,

$$T_2 \leq C \|e_{DG}\|_{DG} \left( \sum_{\kappa \in \mathcal{F}_h} \eta_{\kappa}^2 \right)^{\frac{1}{2}}. \quad (3.16)$$

Term $T_3$: A bound for $T_3$ is found by recalling (3.11). This gives

$$T_3 \leq C_2 \int_{\Gamma_{\text{int},B}} \sigma \|e_{DG}\| \|u_{DG}\| \, ds \leq C_2 \left( \int_{\Gamma_{\text{int},B}} \sigma \|e_{DG}\|^2 \, ds \right)^{\frac{1}{2}} \left( \int_{\Gamma_{\text{int},B}} \sigma \|u_{DG}\|^2 \, ds \right)^{\frac{1}{2}}$$

$$\leq C \|e_{DG}\|_{DG} \left( \gamma \sum_{\kappa \in \mathcal{F}_h} h_{\kappa}^{-1} p_{\kappa}^2 \|u_{DG}\|_{0,\partial\kappa}^2 \right)^{\frac{1}{2}}.$$

Thereby, we obtain

$$T_3 \leq C \|e_{DG}\|_{DG} \left( \sum_{\kappa \in \mathcal{F}_h} \eta_{\kappa}^2 \right)^{\frac{1}{2}}. \quad (3.17)$$

Finally, combining the bounds (3.14) and (3.15)–(3.17) leads to

$$\|e_{DG}\|_{DG}^2 \leq C \left( \sum_{\kappa \in \mathcal{F}_h} \eta_{\kappa}^2 + \epsilon(f, u_{DG}) \right)^{\frac{1}{2}} \|e_{DG}\|_{DG}.$$
3.2 Local lower bounds

In this section, we derive local lower bounds on the error measured in terms of the DG energy norm \( \| \cdot \|_{\text{DG}} \). As in the case of conforming \( hp \)-version finite-element methods, estimators which are both optimally reliable and efficient in the polynomial degree are not currently available in the literature, cf. Melenk & Wohlmuth (2001), for example. The key technical reason for this is that the proofs of the lower bounds exploit the use of inverse estimates which are suboptimal in the polynomial degree.

To minimize the deterioration of the efficiency bounds with respect to the polynomial degree, weighted versions of the local \textit{a posteriori} error indicators \( \eta_\kappa \) may be employed. This idea was first used in the context of conforming finite-element methods in Melenk & Wohlmuth (2001); subsequent extensions to DGFEMs have been undertaken in the article by Houston et al. (2007b), for example. For simplicity of exposition, we only present lower bounds for our \textit{unweighted} \( \Phi \) the weight function \( T \) reducible follows analogously, cf. Remark 3.9 below.

Here, \( \kappa \) is the generic constant \( C \) depends on \( \delta \), but is independent of \( h \) and \( p \).

**Proof.** We proceed similarly as in Melenk & Wohlmuth (2001), see also Houston et al. (2007b). To this end, we first introduce suitable cut off functions as follows: on the reference element \( \hat{T} \), we define a weight function \( \Phi_\kappa(x) = \min_{y \in \partial \hat{T}} |x - y| \). Then, for \( \kappa \in \mathcal{T}_h \), we let \( \Phi_\kappa = c_\kappa \Phi_\kappa \circ F_\kappa^{-1} \), where the factor \( c_\kappa \) is chosen so that \( \int_\kappa (\Phi_\kappa - 1) \, dx = 0 \). Furthermore, on the reference interval \( \hat{\tau} = (-1, 1) \), we define the weight function \( \Phi_{\hat{\tau}}(x) = 1 - x^2 \). Then, for an interior edge \( e \in \mathcal{E}_{\text{int}} \), we let \( \Phi_e = c_e \Phi_{\hat{\tau}} \circ F_e^{-1} \), where \( F_e \) is the affine mapping from \( \hat{\tau} \) to \( e \) and \( c_e \) is chosen so that \( \int_e (\Phi_e - 1) \, ds = 0 \).
Proof of (a): Let $\kappa \in \mathcal{T}_h$ and define $v_K = \Phi_k^{\alpha} \Pi_{\mathcal{T}_h} (f + \nabla \cdot (\mu(|\nabla u_{DG}|) \nabla u_{DG}))$, where $\alpha \in \left(\frac{1}{2}, 1\right)$. Then, using (1.1) and integrating by parts yield

$$\|\Phi_k^{-\frac{q}{2}} v_k\|^2_{0, \kappa} = \int_K v_k \Pi_{\mathcal{T}_h} (f + \nabla \cdot (\mu(|\nabla u_{DG}|) \nabla u_{DG})) \, dx$$

$$= \int_K v_k \cdot \nabla (\mu(|\nabla u_{DG}|) \nabla u_{DG} - \mu(|\nabla u|) \nabla u) \, dx$$

$$+ \int_K v_k (\Pi_{\mathcal{T}_h} - \mathbb{I}) (f + \nabla \cdot (\mu(|\nabla u_{DG}|) \nabla u_{DG})) \, dx$$

$$= - \int_K \nabla v_k \cdot (\mu(|\nabla u_{DG}|) \nabla u_{DG} - \mu(|\nabla u|) \nabla u) \, dx$$

$$+ \int_K v_k (\Pi_{\mathcal{T}_h} - \mathbb{I}) (f + \nabla \cdot (\mu(|\nabla u_{DG}|) \nabla u_{DG})) \, dx$$

$$\leq \int_K |\nabla v_k| \mu(|\nabla u_{DG}|) |\nabla u_{DG} - \mu(|\nabla u|) \nabla u| \, dx$$

$$+ \int_K |v_k| (\Pi_{\mathcal{T}_h} - \mathbb{I}) (f + \nabla \cdot (\mu(|\nabla u_{DG}|) \nabla u_{DG})) \, dx.$$ 

Recalling (1.4), this can be transformed into

$$\|\Phi_k^{-\frac{q}{2}} v_k\|^2_{0, \kappa} \leq C \int_K |\nabla v_k| |\nabla e_{DG}| \, dx + \int_K |v_k| (\Pi_{\mathcal{T}_h} - \mathbb{I}) (f + \nabla \cdot (\mu(|\nabla u_{DG}|) \nabla u_{DG})) \, dx$$

$$\leq C \|\nabla v_k\|_{0, \kappa} \|\nabla e_{DG}\|_{0, \kappa} + \|\Phi_k^{-\frac{q}{2}} v_k\|_{0, \kappa} \|\Phi_k^{-\frac{q}{2}} (\Pi_{\mathcal{T}_h} - \mathbb{I}) (f + \nabla \cdot (\mu(|\nabla u_{DG}|) \nabla u_{DG}))\|_{0, \kappa}$$

$$\leq C \|\nabla v_k\|_{0, \kappa} \|\nabla e_{DG}\|_{0, \kappa} + h_k^{-1} p_k \|\Phi_k^{-\frac{q}{2}} v_k\|_{0, \kappa} \sqrt{\theta_k^{(1)}}.$$

From the proof of Lemma 3.4 of Melenk & Wohlmuth (2001), we have

$$\|\nabla v_k\|_{0, \kappa} \leq Ch_k^{-1} p_k^{2-\alpha} \|\Phi_k^{-\frac{q}{2}} v_k\|_{0, \kappa};$$

thereby,

$$\|\Phi_k^{-\frac{q}{2}} v_k\|^2_{0, \kappa} \leq Ch_k^{-1} p_k \|\Phi_k^{-\frac{q}{2}} v_k\|_{0, \kappa} \left(p_k^{1-\alpha} \|\nabla e_{DG}\|_{0, \kappa} + \sqrt{\theta_k^{(1)}}\right).$$

Dividing both sides of the above inequality by $\|\Phi_k^{-\frac{q}{2}} v_k\|_{0, \kappa}$ and observing that (by applying the inverse inequality from Theorem 2.5 of Melenk & Wohlmuth (2001))

$$\|\Pi_{\mathcal{T}_h} (f + \nabla \cdot (\mu(|\nabla u_{DG}|) \nabla u_{DG}))\|_{0, \kappa} \leq C p_k^\alpha \|\Phi_k^{-\frac{q}{2}} \Pi_{\mathcal{T}_h} (f + \nabla \cdot (\mu(|\nabla u_{DG}|) \nabla u_{DG}))\|_{0, \kappa}$$

$$= C p_k^\alpha \|\Phi_k^{-\frac{q}{2}} v_k\|_{0, \kappa}.$$
lead to

$$\| \Pi_{\mathcal{H}} (f + \nabla \cdot (\mu (|\nabla u_DG|) \nabla u_DG)) \|_{0,K} \leq \frac{1}{\delta k} p_k^{1+a} \left( p_k^{1-a} \| \nabla e_{DG} \|_{0,K} + \sqrt{\delta k^b(1)} \right).$$  (3.18)

Choosing $\delta = a - \frac{1}{2}$ completes the proof of (a).

**Proof of (b):** Let $q_e = \phi_e^\alpha \Pi_{\mathcal{H}} (\| \mu (|\nabla u_DG|) \nabla u_DG \| \cdot v)|_e$, where again $\alpha \in (\frac{1}{2}, 1]$. Then, referring to Lemma 2.6 of Melenk & Wohlmuth (2001) with $\epsilon = p_k^{-2}$, there exists $\chi_e \in H^1_0(\omega_e)$ such that $\chi_e|_e = q_e$ and

$$\| \chi_e \|_{0,\omega_e} \leq \frac{1}{\delta k} p_k^{-1} \| \phi_e^{1/2} q_e \|_{0,e},$$

$$\| \nabla \chi_e \|_{0,\omega_e} \leq \frac{1}{\delta k} p_k^{-1} \| \phi_e^{1/2} q_e \|_{0,e}.$$  (3.19)

Noting that $-\nabla \cdot (\mu (|\nabla u|) \nabla u) = f \in L^2(\Omega)$, we conclude that $\| \mu (|\nabla u|) \nabla u \| \cdot v = 0$ on $e$. Hence, integrating by parts and assuming (without loss of generality) that the normal vector $v$ points from $\kappa$ to $\kappa'$ lead to

$$\| \phi_e^{1/2} q_e \|_{0,e}^2$$

$$= \int_e \Pi_{\mathcal{H}} (\| \mu (|\nabla u_DG|) \nabla u_DG \| \cdot v) \chi_e \, ds$$

$$= \int_e (\| \mu (|\nabla u_DG|) \nabla u_DG - \mu (|\nabla u|) \nabla u \| \cdot v) \chi_e \, ds + \int_e (\Pi_{\mathcal{H}} - \mathbb{I})(\| \mu (|\nabla u_DG|) \nabla u_DG \| \cdot v) \chi_e \, ds$$

$$= \int_{\partial \omega_e} ((\mu (|\nabla u_DG|) \nabla u_DG - \mu (|\nabla u|) \nabla u) \cdot \kappa) \chi_e \, ds$$

$$+ \int_{\partial \omega_e} ((\mu (|\nabla u_DG|) \nabla u_DG - \mu (|\nabla u|) \nabla u) \cdot \kappa) \chi_e \, ds$$

$$+ \int_e (\Pi_{\mathcal{H}} - \mathbb{I})(\| \mu (|\nabla u_DG|) \nabla u_DG \| \cdot v) \chi_e \, ds$$

$$= \int_{\partial \omega_e} (\mu (|\nabla u_DG|) \nabla u_DG - \mu (|\nabla u|) \nabla u) \cdot \nabla \chi_e \, dx + \int_{\partial \omega_e} (f + \nabla \cdot (\mu (|\nabla u_DG|) \nabla u_DG)) \chi_e \, dx$$

$$+ \int_{\partial \omega_e} (\Pi_{\mathcal{H}} - \mathbb{I})(\mu (|\nabla u_DG|) \nabla u_DG \cdot v) \chi_e \, ds$$

$$\equiv R_1 + R_2 + R_3.$$  (3.20)

Employing (1.4) and (3.19), $R_1$ can be bounded as follows:

$$R_1 \leq C \int_{\partial \omega_e} |\nabla e_{DG}| \cdot |\nabla \chi_e| \, dx \leq C \| \nabla e_{DG} \|_{0,\omega_e} \| \nabla \chi_e \|_{0,\omega_e} \leq C h_{\delta k}^{-1/2} p_k \| \nabla e_{DG} \|_{0,\omega_e} \| \phi_e^{1/2} q_e \|_{0,e}.$$  (3.21)
In order to obtain a bound for \( R_2 \), we use (3.18) and the definition of \( \Theta^{(1)}_k \) from (3.1); thereby,

\[
R_2 = \int_{\Omega_e} \Pi_{\mathcal{E}}(f + \nabla \cdot (\mu (|\nabla u_{DG}|) \nabla u_{DG})) \chi_e \, dx
\]

\[
- \int_{\Omega_e} (\Pi_{\mathcal{E}} - \mathbb{1}) (f + \nabla \cdot (\mu (|\nabla u_{DG}|) \nabla u_{DG})) \chi_e \, dx
\]

\[
\leq \left\| \Pi_{\mathcal{E}}(f + \nabla \cdot (\mu (|\nabla u_{DG}|) \nabla u_{DG})) \right\|_{0, \Omega_e} \| \chi_e \|_{0, \Omega_e}
\]

\[
+ \left\| (\Pi_{\mathcal{E}} - \mathbb{1}) (f + \nabla \cdot (\mu (|\nabla u_{DG}|) \nabla u_{DG})) \right\|_{0, \Omega_e} \| \chi_e \|_{0, \Omega_e}
\]

\[
\leq C h_{k}^{-\frac{1}{2}} p_{k}^{1+\alpha} \left( p_{k}^{1-\alpha} \| \nabla e_{DG} \|_{0, \Omega_e} + \sum_{\tau \in \{ \kappa, \kappa' \}} \sqrt{\Theta^{(1)}_\tau} \right) \| \chi_e \|_{0, \Omega_e},
\]

(3.22)

Recalling (3.19), this gives

\[
R_2 \leq C h_{k}^{-\frac{1}{2}} p_{k}^{\alpha} \left( p_{k}^{1-\alpha} \| \nabla e_{DG} \|_{0, \Omega_e} + \sum_{\tau \in \{ \kappa, \kappa' \}} \sqrt{\Theta^{(1)}_\tau} \right) \| \Phi_{e}^{-\frac{1}{2}} q_{e} \|_{0, \Omega_e}.
\]

The bound on \( R_3 \) is based on the definition of \( \Theta^{(2)}_e \) from (3.2) and on the fact that \( \chi_e = q_e \) on \( e \):

\[
R_3 \leq \left\| \Phi_{e}^{-\frac{1}{2}} (\Pi_{\mathcal{E}} - \mathbb{1}) (\| \mu (|\nabla u_{DG}|) \nabla u_{DG} \| \cdot \nu) \right\|_{0, \Omega_e} \| \Phi_{e}^{-\frac{1}{2}} \chi_e \|_{0, \Omega_e}
\]

\[
\leq C h_{k}^{-\frac{1}{2}} p_{k}^{\frac{1}{2}} \sqrt{\Theta^{(2)}_e} \left\| \Phi_{e}^{-\frac{1}{2}} q_{e} \right\|_{0, \Omega_e}.
\]

(3.23)

Combining (3.20)–(3.23) gives

\[
\left\| \Phi_{e}^{-\frac{1}{2}} q_{e} \right\|_{0, \Omega_e}^2 \leq C h_{k}^{-\frac{1}{2}} p_{k}^{\alpha} \left( \| \nabla e_{DG} \|_{0, \Omega_e} + p_{k}^{\alpha-1} \sum_{\tau \in \{ \kappa, \kappa' \}} \sqrt{\Theta^{(1)}_\tau} + p_{k}^{\frac{1}{2}} \sqrt{\Theta^{(2)}_e} \right) \| \Phi_{e}^{-\frac{1}{2}} q_{e} \|_{0, \Omega_e}.
\]

As in the proof of (a), we divide the above inequality by \( \left\| \Phi_{e}^{-\frac{1}{2}} q_{e} \right\|_{0, \Omega_e} \) and use the fact that \( \Phi_{e}^{-\frac{1}{2}} q_{e} = \Phi_{e}^{-\frac{1}{2}} \Pi_{\mathcal{E}} [\| \mu (|\nabla u_{DG}|) \nabla u_{DG} \| \cdot \nu]_{e} \). Then, applying the inverse inequality from Lemma 2.4 of Melenk & Wohlmuth (2001) (see also Bernardi & Maday, 1997; Bernardi et al., 2001), we get

\[
\| \Pi_{\mathcal{E}} (\| \mu (|\nabla u_{DG}|) \nabla u_{DG} \| \cdot \nu) \|_{0, \Omega_e} \leq C_{p_k} \left\| \Phi_{e}^{-\frac{1}{2}} \Pi_{\mathcal{E}} (\| \mu (|\nabla u_{DG}|) \nabla u_{DG} \| \cdot \nu) \right\|_{0, \Omega_e}
\]

\[
= C_{p_k} \left\| \Phi_{e}^{-\frac{1}{2}} q_{e} \right\|_{0, \Omega_e}.
\]

Thereby,

\[
\| \Pi_{\mathcal{E}} (\| \mu (|\nabla u_{DG}|) \nabla u_{DG} \| \cdot \nu) \|_{0, \Omega_e}
\]

\[
\leq C h_{k}^{-\frac{1}{2}} p_{k}^{1+\alpha} \left( \| \nabla e_{DG} \|_{0, \Omega_e} + p_{k}^{\alpha-1} \sum_{\tau \in \{ \kappa, \kappa' \}} \sqrt{\Theta^{(1)}_\tau} + p_{k}^{\frac{1}{2}} \sqrt{\Theta^{(2)}_e} \right).
\]

Again, selecting \( \delta = \alpha - \frac{1}{2} \) leads to estimate (b).
Proof of (c): This follows from (2.1), (2.2) and (3.11):

$$\|\|u_{DG}\|\|_{0,e} = \|\|e_{DG}\|\|_{0,e} \leq C \gamma^{-\frac{1}{2}} h_{\tilde{k}}^\frac{1}{2} p_{\tilde{k}}^{-1} \|\|\|e_{DG}\|\|_{0,e}^{\frac{1}{2}} \|\|\| e_{DG}\|\|_{0,e}.$$  

That completes the proof of the lower bounds.

**Remark 3.9** For the case when the mesh $\mathcal{T}_h$ is one-irregular (but assumed to be regularly reducible to a conforming mesh $\mathcal{T}_h$, cf. Section 2), analogous bounds to the ones derived in Theorem 3.8 still hold. Indeed, bounds (a) and (c) follow directly; for the proof of (b), employing the argument outlined in the proof of Theorem 3.8, we deduce that

$$\|\Pi_{\mathcal{E}}(\|\mu (|\nabla u_{DG}|) \nabla u_{DG} \cdot \mathbf{v})\|_{0,e} \leq C h_{\tilde{k}}^{-\frac{1}{2}} p_{\tilde{k}}^{\frac{\delta+1}{2}} \left( \|\nabla e_{DG}\|_{0,\tilde{o}_e} + p_{\tilde{k}}^{-\frac{1}{2}} \sum_{\tau \in [\tilde{k}, \tilde{k}']} \sqrt{\sigma_{\tau}^{(1)} + p_{\tilde{k}}^{-1} \sigma_{\tau}^{(2)}} \right).$$

(3.24)

where $\tilde{o}_e$ is defined so that the closure of $\tilde{o}_e$ is the union of the closure of the two elements $\tilde{k}, \tilde{k}' \in \mathcal{T}_h$ which share the common edge $e$. The right-hand side of (3.24) may now be bounded from above by a similar expression involving quantities measured over the (non-matching) elements $\kappa$ and $\kappa'$ which share the edge $e$; by this we mean that in the estimate (3.24), the element size $h_{\tilde{k}}$ and polynomial degree $p_{\tilde{k}}$ are commensurate with $h_{\kappa}$ and $p_{\kappa}$, respectively, and the error term $\|\nabla e_{DG}\|_{0,\tilde{o}_e}$ is bounded from above by $\|\nabla e_{DG}\|_{0,\tilde{o}_e}$.

We note that the data oscillation terms $\sigma_{\tau}^{(1)}$ appearing in (3.24) are, however, still measured over the elements $\tilde{k}, \tilde{k}' \in \mathcal{T}_h$ since they are in general not bounded by the corresponding oscillations on the elements $\kappa, \kappa' \in \mathcal{T}_h$.

4. Numerical experiments

In this section, we present a series of numerical examples to demonstrate the practical performance of the proposed *a posteriori* error estimator derived in Theorem 3.2 within an automatic $hp$-adaptive refinement procedure which is based on one-irregular quadrilateral elements. In each of the examples shown in this section, the DG solution $u_{DG}$ defined by (2.6) is computed with $\theta = 0$, i.e. we employ an incomplete interior-penalty-type DG method. Analogous results to those presented for $\theta = 0$ are also observed with $\theta = -1$ and $\theta = 1$; for brevity, these results have been omitted. Additionally, we set the constant $\gamma$ appearing in the definition of the interior-penalty parameter $\sigma$ defined in (2.5) equal to 10. The resulting system of non-linear equations is solved by employing a damped Newton method; within each inner (linear) iteration, we exploit a (left) preconditioned GMRES algorithm using a block symmetric Gauss–Seidel preconditioner.

The $hp$-adaptive meshes are constructed by first marking the elements for refinement/derefinement according to the size of the local error indicators $\eta_{\kappa}$. This is achieved by employing the fixed fraction strategy (see Houston & Süli, 2002) with refinement and derefinement fractions set to 25% and 10%, respectively. Once an element $\kappa \in \mathcal{T}_h$ has been flagged for refinement or derefinement, a decision must be made whether the local mesh size $h_{\kappa}$ or the local degree $p_{\kappa}$ of the approximating polynomial should be adjusted accordingly. The choice to perform either $h$-refinement/derefinement...
or $p$-refinement/derefinement is based on estimating the local smoothness of the (unknown) analytical solution. To this end, we employ the $hp$-adaptive strategy developed in Houston & Suli (2005), where the local regularity of the analytical solution is estimated from truncated local Legendre expansions of the computed numerical solution, see, also, Eibner & Melenk (2004) and Houston et al. (2003).

Here, the emphasis will be on investigating the asymptotic sharpness of the proposed a posteriori error bound on a sequence of non-uniform $hp$-adaptively refined one-irregular meshes. To this end, we shall compare the estimator derived in Theorem 3.2, which is slightly suboptimal (by a factor of $p^{1/2}$) in the spectral order $p$, with the corresponding optimal one (cf. Remark 3.3); we note that the derivation of the latter precludes the use of hanging nodes. Indeed, here we shall show that despite the loss of optimality in $p$, the former indicator performs extremely well on $hp$-refined meshes, in the sense that the ‘effectivity index’, which is defined as the ratio of the a posteriori error bound and the energy norm of the actual error, is roughly constant on all the meshes employed. Moreover, our numerical experiments indicate that both a posteriori error indicators give rise to very similar quantitative results. For simplicity, as in Becker et al. (2003), we set the constant $C$ arising in Theorem 3.2 equal to one; in general, to ensure the reliability of the error estimator, this constant must be determined numerically for the underlying problem at hand. In all our experiments, the data-approximation terms in the a posteriori bound stated in Theorem 3.2 will be neglected.

4.1 Example 1

In this example, we let $\Omega$ to be the unit square $(0, 1)^2$ in $\mathbb{R}^2$. The non-linear diffusion coefficient is defined as follows:

$$\mu(x, |\nabla u|) = 2 + \frac{1}{1 + |\nabla u|}.$$  

Further, we select $f$ so that the analytical solution to (1.1)–(1.2) is given by

$$u(x, y) = x(1 - x)y(1 - y)(1 - 2y)e^{-s(2x-1)^2},$$

where $s$ is a positive constant (cf. Houston et al., 2007b; Melenk & Wohlmuth, 2001); throughout this section, we set $s = 20$.

In Fig. 1(a), we present a comparison of the actual and estimated energy norm of the error versus the third root of the number of degrees of freedom in the finite-element space $S^p(\Omega, T_h, \mathbf{F})$ on a linear–log scale, for the sequence of meshes generated by our $hp$-adaptive algorithm using the suboptimal indicator stated in Theorem 3.2 (denoted by $p^3$ in the figure) and the corresponding optimal one outlined in Remark 3.3 (denoted by $p^2$ in the figure). We note that for both indicators meshes employing hanging nodes are employed, despite the fact that the derivation of the latter, $hp$-optimal, error indicator necessitates the use of conforming (regular) meshes. The third root of the number of degrees of freedom is chosen on the basis of the a priori error analysis performed in Wihler et al. (2003), for example. Here, we observe that the two error indicators perform in a very similar manner: in each case, the error bound overestimates the true error by a (reasonably) consistent factor. Indeed, from Fig. 1(b), we see that the computed effectivity indices oscillate around a value of approximately 13. Additionally, from Fig. 1(a), we observe that the convergence lines using $hp$-refinement are (roughly) straight on a linear–log scale, which indicates that exponential convergence is attained for this smooth problem, as we would expect. In Fig. 1(c, d), we present a comparison between the actual energy norm of the error employing both $h$- and $hp$-mesh refinement; here, the $hp$-refinement is based on employing the error indicator stated in
Theorem 3.2. In the former case, the DG solution $u_{DG}$ is computed using bilinear elements, i.e. $p = 1$; here, the adaptive algorithm is again based on employing the fixed fraction strategy with refinement and derefinement fractions set to 25% and 10%, respectively. From Fig. 1(c, d), we clearly observe the superiority of employing a grid adaptation strategy based on exploiting $hp$-adaptive refinement: on the final mesh, the energy norm of the error using $hp$-refinement is over two orders of magnitude smaller than the corresponding quantity computed when $h$-refinement is employed alone.

In Fig. 2, we show the mesh generated using the proposed $hp$-version a posteriori error indicator stated in Theorem 3.2 after 11 $hp$-adaptive refinement steps. For clarity, we show the $h$-mesh alone as well as the corresponding polynomial degree distribution on this mesh. Here, we observe that some $h$-refinement of the mesh has been performed in the vicinity of the base of the exponential ‘hills’ situated in the left- and the right-hand sides of the domain, where the gradient/curvature of the analytical solution is relative large. Once the $h$-mesh has adequately captured the structure of the solution, the $hp$-adaptive algorithm increased the degree of the approximating polynomial within the interior part of the domain containing these hills.
Fig. 2. Example 1. Finite-element mesh after 11 adaptive refinements with 1198 elements and 18443 degrees of freedom: (a) $h$-mesh alone and (b) $hp$-mesh.
4.2 Example 2

In this section, we let \( \Omega \) denote the L-shaped domain \((-1, 1)^2 \setminus [0, 1) \times (-1, 0)\), and select

\[
\mu(x, |\nabla u|) = 1 + e^{-|\nabla u|^2}.
\]

Then, writing \((r, \varphi)\) to denote the system of polar co-ordinates, we choose \( f \) and an appropriate inhomogeneous boundary condition for \( u \) so that

\[
u = r^{2/3} \sin(2\varphi/3),
\]

cf. Wihler et al. (2003), for example. We note that \( u \) is analytic in \( \overline{\Omega} \setminus \{0\} \), but \( \nabla u \) is singular at the origin; indeed, here \( u \notin H^2(\Omega) \).

Figure 3(a) shows the history of the actual and estimated energy norm of the error on each of the meshes generated by our \( hp \)-adaptive algorithm using both the indicator stated in Theorem 3.2 (denoted by \( p^3 \) in the figure) and the corresponding one outlined in Remark 3.3 (denoted by \( p^2 \) in the figure). As in Section 4.1, we observe that the two error indicators perform in a very similar manner, though for this non-smooth example, the loss in optimality in the jump indicator in the estimator stated in Theorem 3.2

![Figure 3(a)](image)

**Fig. 3.** Example 2. (a) Comparison of the actual and estimated energy norm of the error with respect to the (third root of the) number of degrees of freedom with \( hp \)-adaptive mesh refinement, (b) effectivity indices and (c, d) comparison of the actual error with \( h \)- and \( hp \)-adaptive mesh refinement.
does lead to a slight increase in the effectivity indices in comparison with the latter indicator. However, from Fig. 3(b), we observe that asymptotically both a posteriori bounds overestimate the true error by a consistent factor. Additionally, from Fig. 3(a), we observe exponential convergence of the energy norm of the error using both estimators with $hp$-refinement; indeed, on a linear–log scale, the convergence lines are, on average, straight. Figure 3(c, d) highlights the superiority of employing $hp$-adaptive refinement in comparison with $h$-refinement: on the final mesh, the energy norm of the error using the $hp$-refinement indicator stated in Theorem 3.2 is over two orders of magnitude smaller than the corresponding quantity when $h$-refinement is employed alone, based on using bilinear elements.

In Fig. 4, we show the mesh generated using the local error indicators $\eta_k$ stated in Theorem 3.2 after 13 $hp$-adaptive refinement steps. Here, we see that the $h$-mesh has been refined in the vicinity of the re-entrant corner located at the origin; from the zoom, we see that $h$-refinement is more pronounced in the direction $y = x$. In the normal direction, $y = -x$, $p$-refinement is employed instead as the solution is deemed to be smooth here. Additionally, we see that the polynomial degrees have been increased away from the re-entrant corner located at the origin since the underlying analytical solution is smooth in this region.

5. Concluding remarks

In this paper, we derived global upper and local lower residual-based a posteriori error bounds in the energy norm for the class of interior-penalty $hp$-DGFEMs developed in Houston et al. (2005) for the
numerical approximation of second-order quasi-linear elliptic PDEs. The analysis is based on employing a suitable DG space decomposition together with an *hp*-version projection operator. Numerical experiments presented in this article clearly demonstrate that the proposed *a posteriori* estimator converges to zero at the same asymptotic rate as the energy norm of the actual error on sequences of *hp*-adaptively refined meshes. Future work will be devoted to the extension of our analysis to *hp*-adaptive DG approximations of quasi-Newtonian incompressible flow models.

**Acknowledgement**

PH acknowledges the support of the UK Engineering and Physical Sciences Research Council, Grant number GR/R76615.

**REFERENCES**


