Report no. 00/15

Discontinuous *hp*–Finite Element Methods for Advection–Diffusion Problems

Paul Houston[a] Christoph Schwab[b], and Endre Süli[c]

[a]Department of Mathematics & Computer Science, University of Leicester, UK
[b]Seminar for Applied Mathematics, ETH Zürich, CH-8092 Zürich, Switzerland
[c]Oxford University Computing Laboratory, Wolfson Building, Parks Road, Oxford, UK

We consider the *hp*–version of the discontinuous Galerkin finite element method for second–order partial differential equations with nonnegative characteristic form. This class of equations includes second–order elliptic and parabolic equations, first–order hyperbolic equations, as well as problems of mixed hyperbolic–elliptic–parabolic type. Our main concern is the error analysis of the method in the absence of streamline–diffusion stabilization. In the hyperbolic case, an *hp*–optimal error bound is derived. In the self–adjoint elliptic case, an error bound that is h–optimal and p–suboptimal by $\frac{1}{2}$ a power of p is obtained. These estimates are then combined to deduce an error bound in the general case. For element-wise analytic solutions the method exhibits exponential rates of convergence under p–refinement. The theoretical results are illustrated by numerical experiments.

Subject classifications: AMS(MOS): 65N12, 65N15, 65N30

Key words and phrases: *hp*–finite element methods, discontinuous Galerkin methods, PDEs with nonnegative characteristic form

Oxford University Computing Laboratory
Numerical Analysis Group
Wolfson Building
Parks Road
Oxford, England OX1 3QD

June, 2000

1 Introduction

Discontinuous Galerkin Finite Element Methods (DGFEMs) were introduced in the early 1970s for the numerical solution of first–order hyperbolic problems (see [28, 20, 17, 18, 9, 10] and [12, 13, 25, 26]). Simultaneously, but quite independently, they were proposed as nonstandard schemes for the approximation of second–order elliptic equations [22, 33, 1]. In recent years there has been renewed interest in this class of techniques, 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 [14, 19], and the desire to handle nonlinear hyperbolic problems in a locally conservative manner and without auxiliary numerical stabilization [8, 11] (see also [6, 7] for the error analysis of the local version of the DGFEM in the elliptic case).

For first–order linear transport problems the use of stabilized *hp*–finite element methods, with a stabilization term of a streamline–diffusion type, was investigated recently in [15]. It was shown that a proper choice of the stabilization parameter leads to optimal convergence rates, in the mesh-width h and in the polynomial degree p, for the *hp*–versions of the discontinuous Galerkin finite element method and the (continuous) streamline–diffusion finite element method. Our purpose here is to extend that analysis to general advection–diffusion problems, without invoking streamline–diffusion stabilization so as to reduce the amount of numerical dissipation in the method.

The paper is structured as follows. After introducing, in Section 2, the requisite notation and our model boundary value problem for a partial differential equation with nonnegative characteristic form, we consider, in Section 3, the *hp*–DGFEM in the absence of streamline–diffusion stabilization, in the hyperbolic, purely advective case. Error bounds that are optimal in both h and p are derived by means of an analysis different from that in [15]. In the purely diffusive elliptic case, in Section 4, we analyze the *hp*–version of the DGFEM proposed in [24], which is closely related to mortar element methods with *a priori* determined multipliers (see Section 6.4 of [30]). We rigorously prove *hp*–convergence results similar to those announced in [3] which apply both in two and in three space dimensions. Our analysis exploits a stabilization device due to Nitsche [22], see also [1, 33], based on the penalization of discontinuities in the discrete counterpart of the diffusive normal flux at element interfaces. We establish an error bound that is optimal in h and suboptimal in p by $\frac{1}{2}$ a power of p. The results in Section 4 represent an extension of the recent analysis of Rivière, Wheeler, and Girault [27] to finite element spaces with locally varying polynomial degrees and highlight the nature of the dependence of the discontinuity–penalization parameter on the diffusion coefficient. This latter refinement, in particular, is crucial for our extension of the error analysis from the symmetric elliptic situation to the case of second–order partial differential equations with degenerate diffusion. Indeed, in Section 5, we address the analysis of the method for the class of second–order partial differential equations with nonnegative characteristic form (which includes advection–diffusion problems as well as partial differential equations of mixed elliptic–hyperbolic–parabolic type)

by combining the results of Sections 3 and 4 to deduce *hp*–error estimates for this general case. We also show that if the solution is element-wise analytic then the global error decays at an exponential rate. Section 6 presents numerical experiments which confirm the theoretical results. The analysis in this paper is a complete and improved account of our recent work announced in the conference papers [31, 32].

2 Preliminaries

2.1 Model problem

Let Ω be a bounded open polyhedral domain in $\mathbb{R}^d$, $d \geq 2$, and let Γ signify the union of its $(d-1)$–dimensional open faces. We consider the diffusion–advection–reaction equation

$$\mathcal{L}u \equiv -\sum_{i,j=1}^{d} \partial_j(a_{ij}(x)\,\partial_i u) + \sum_{i=1}^{d} b_i(x)\,\partial_i u + c(x)u = f(x)\ , \tag{2.1}$$

where $f \in L^2(\Omega)$ and $c \in L^\infty(\Omega)$ are real–valued, $b = \{b_i\}_{i=1}^d$ is a vector function whose entries b_i are Lipschitz continuous real–valued functions on $\bar{\Omega}$, and $a = \{a_{ij}\}_{i,j=1}^d$ is a *symmetric* matrix whose entries a_{ij} are bounded, piecewise continuous real–valued functions defined on $\bar{\Omega}$, with

$$\zeta^{\mathrm{T}} a(x)\zeta \geq 0 \quad \forall \zeta \in \mathbb{R}^d\ , \quad \text{a.e. } x \in \bar{\Omega}\ . \tag{2.2}$$

Under this hypothesis, (2.1) is termed a *partial differential equation with nonnegative characteristic form*. By $\mu(x) = \{\mu_i(x)\}_{i=1}^d$ we denote the unit outward normal vector to Γ at $x \in \Gamma$. On introducing the so called *Fichera function* $b \cdot \mu$, we define

$$\Gamma_0 = \left\{ x \in \Gamma :\ \mu(x)^{\mathrm{T}} a(x)\mu(x) > 0 \right\}\ , \tag{2.3}$$

$$\Gamma_- = \{x \in \Gamma\backslash\Gamma_0 :\ b(x)\cdot\mu(x) < 0\}\ , \quad \Gamma_+ = \{x \in \Gamma\backslash\Gamma_0 :\ b(x)\cdot\mu(x) \geq 0\}\ . \tag{2.4}$$

The sets Γ_- and Γ_+ will be referred to as the inflow and outflow boundary, respectively. Evidently, $\Gamma = \Gamma_0 \cup \Gamma_- \cup \Gamma_+$. If Γ_0 is nonempty, we shall further divide it into disjoint subsets Γ_{D} and Γ_{N} whose union is Γ_0, with Γ_{D} nonempty and relatively open in Γ. We supplement (2.1) with the boundary conditions

$$\begin{aligned} u &= g_{\mathrm{D}} \quad \text{on} \quad \Gamma_{\mathrm{D}} \cup \Gamma_-\ , \\ \mu\cdot(a\nabla u) &= g_{\mathrm{N}} \quad \text{on} \quad \Gamma_{\mathrm{N}}\ , \end{aligned} \tag{2.5}$$

and adopt the (physically reasonable) hypothesis that $b \cdot \mu \geq 0$ on Γ_{N}, whenever Γ_{N} is nonempty. The well-posedness of the boundary value problem (2.1), (2.5), in the case of homogeneous boundary conditions, is shown in the Appendix. Next we introduce the finite element spaces which the *hp*–DGFEM is based on.

2.2 Finite element spaces

Let $\mathcal{T}$ be a subdivision of Ω into disjoint open element domains κ such that $\bar{\Omega} = \cup_{\kappa\in\mathcal{T}}\bar{\kappa}$, where $\mathcal{T}$ is regular or 1-irregular, i.e., each face of κ in $\mathcal{T}$ has at most one hanging node. We assume that the family of subdivisions $\mathcal{T}$ is shape–regular (cf. pp. 61, 113, and Remark 2.2, p.114, in [5]) and each $\kappa \in \mathcal{T}$ is an affine image of a fixed master element $\hat{\kappa}$; i.e., $\kappa = F_\kappa(\hat{\kappa})$ for all $\kappa \in \mathcal{T}$, where $\hat{\kappa}$ is either the open unit simplex or the open unit hypercube in $\mathbb{R}^d$. For a nonnegative integer k, we denote by $\mathcal{P}_k(\hat{\kappa})$ the set of polynomials of total degree k on $\hat{\kappa}$. When $\hat{\kappa}$ is the unit hypercube, we also consider $\mathcal{Q}_k(\hat{\kappa})$, the set of all tensor–product polynomials on $\hat{\kappa}$ of degree k in each coordinate direction. To each $\kappa \in \mathcal{T}$ we assign a nonnegative integer p_κ (local polynomial degree) and a nonnegative integer s_κ (local Sobolev index), collect the p_κ, s_κ and F_κ in the vectors $\mathbf{p} = \{p_\kappa : \kappa \in \mathcal{T}\}$, $\mathbf{s} = \{s_\kappa : \kappa \in \mathcal{T}\}$ and $\mathbf{F} = \{F_\kappa : \kappa \in \mathcal{T}\}$, and consider the finite element space

$$S^{\mathbf{p}}(\Omega, \mathcal{T}, \mathbf{F}) = \{u \in L^2(\Omega) : \ u|_\kappa \circ F_\kappa \in \mathcal{R}_{p_\kappa}(\hat{\kappa})\} \ ,$$

where $\mathcal{R}$ is either $\mathcal{P}$ or $\mathcal{Q}$. Further, we assign to the subdivision $\mathcal{T}$ the broken Sobolev space of composite order $\mathbf{s}$,

$$H^{\mathbf{s}}(\Omega, \mathcal{T}) = \{u \in L^2(\Omega) : \ u|_\kappa \in H^{s_\kappa}(\kappa) \quad \forall \kappa \in \mathcal{T}\} \ ,$$

equipped with the broken Sobolev norm and corresponding seminorm, respectively,

$$\|u\|_{\mathbf{s},\mathcal{T}} = \Big(\sum_{\kappa\in\mathcal{T}} \|u\|^2_{H^{s_\kappa}(\kappa)}\Big)^{\frac{1}{2}}, \qquad |u|_{\mathbf{s},\mathcal{T}} = \Big(\sum_{\kappa\in\mathcal{T}} |u|^2_{H^{s_\kappa}(\kappa)}\Big)^{\frac{1}{2}} . \tag{2.6}$$

When $s_\kappa = s$ for all $\kappa \in \mathcal{T}$, we shall write $H^s(\Omega, \mathcal{T})$, $\|u\|_{s,\mathcal{T}}$ and $|u|_{s,\mathcal{T}}$. For $u \in H^1(\Omega, \mathcal{T})$ we define the broken gradient $\nabla_{\mathcal{T}} u$ of u by $(\nabla_{\mathcal{T}} u)|_\kappa = \nabla(u|_\kappa)$, $\kappa \in \mathcal{T}$.

Let us consider the set $\mathcal{E}$ of all open $(d-1)$–dimensional faces (open edges when $d = 2$ or open faces when $d = 3$) of all elements $\kappa \in \mathcal{T}$. Given that $\mathcal{T}$ may be irregular and hanging nodes are permitted in the DGFEM, $\mathcal{E}$ will be understood to contain the *smallest* common $(d-1)$-dimensional interfaces of neighboring elements (cf. Figure 1). Further, we denote by $\mathcal{E}_{\rm int}$ the set of all e in $\mathcal{E}$ that are contained in Ω, we let $\Gamma_{\rm int} = \{x \in \Omega : x \in e \text{ for some } e \in \mathcal{E}_{\rm int}\}$ and we introduce the set $\mathcal{E}_{\rm D}$ of $(d-1)$–dimensional boundary faces contained in the subset $\Gamma_{\rm D}$ of Γ. Implicit in these definitions is the assumption that $\mathcal{T}$ respects the decomposition of Γ in the sense that each $e \in \mathcal{E}$ that lies on Γ belongs to the interior of exactly one of Γ_-, Γ_+, $\Gamma_{\rm D}$, $\Gamma_{\rm N}$.

3 Pure advection

3.1 Formulation of the problem

In this section, we study the discontinuous Galerkin finite element approximation of the advective part $\mathcal{L}_0$ of $\mathcal{L}$; thus we set $a = 0$ and $\Gamma_0 = \emptyset$, and consider

$$\begin{aligned} \mathcal{L}_0 u \equiv b \cdot \nabla u + cu &= f \quad \text{in } \Omega \ , \\ u &= g_{\rm D} \ \text{ on } \Gamma_- \ . \end{aligned} \tag{3.1}$$

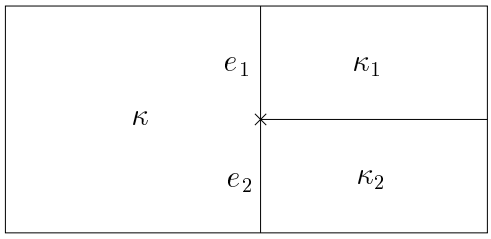


Figure 1: Hanging node $\times$ and faces $e_1, e_2 \in \mathcal{E}_{\rm int}$.

We adopt the following (standard) hypothesis: there exists a positive constant γ_0 such that

$$c(x) - \frac{1}{2}\,\nabla \cdot b(x) \geq \gamma_0 \ \text{ a.e. } \ x \in \Omega\ ; \tag{3.2}$$

we then define the function c_0 by

$$(c_0(x))^2 = c(x) - \frac{1}{2}\,\nabla \cdot b(x)\ . \tag{3.3}$$

For any element $\kappa \in \mathcal{T}$, we denote by $\partial\kappa$ the union of $(d-1)$-dimensional open faces of κ. Then, the inflow and outflow parts of $\partial\kappa$ are defined by

$$\partial_-\kappa = \{x \in \partial\kappa : b(x)\cdot\mu_\kappa(x) < 0\}\,,\ \ \partial_+\kappa = \{x \in \partial\kappa : b(x)\cdot\mu_\kappa(x) \geq 0\}\ ,$$

respectively, where $\mu_\kappa(x)$ denotes the unit outward normal vector to $\partial\kappa$ at $x \in \partial\kappa$.

For each $\kappa \in \mathcal{T}$ and $v \in H^1(\kappa)$, we denote by v^+_κ the interior trace of $v|_\kappa$ on $\partial\kappa$. Now consider $\kappa \in \mathcal{T}$ such that $\partial_-\kappa\backslash\Gamma$ is nonempty. Then, for almost every (with respect to the $(d-1)$–dimensional surface measure) $x \in \partial_-\kappa\backslash\Gamma$ there exists a unique $\kappa' \in \mathcal{T}$ (depending in general on the location of x on $\partial\kappa$) such that $x \in \partial_+\kappa'$. Assume now that $v \in H^1(\Omega,\mathcal{T})$. If $\partial_-\kappa\backslash\Gamma$ is nonempty for some $\kappa \in \mathcal{T}$, then the outer trace v^-_κ of v on $\partial_-\kappa\backslash\Gamma$ relative to κ is defined as the inner trace $v^+_{\kappa'}$ relative to the element(s) κ' such that the intersection of $\partial_+\kappa'$ with $\partial_-\kappa\backslash\Gamma$ has positive $(d-1)$–dimensional measure. We then define the jump of v across $\partial_-\kappa\backslash\Gamma$ by

$$\lfloor v \rfloor_\kappa := v^+_\kappa - v^-_\kappa\ .$$

Since below it will always be clear from the context which element κ in the subdivision $\mathcal{T}$ the quantities μ_κ, v^+_κ, v^-_κ and $\lfloor v\rfloor_\kappa$ correspond to, for the sake of simplicity we shall suppress the letter κ in the subscript and write, respectively, μ, v^+, v^- and $\lfloor v\rfloor$ instead.

For $v, w \in H^1(\Omega,\mathcal{T})$ we consider the bilinear form

$$A(w,v) = \sum_{\kappa\in\mathcal{T}}\left(\int_\kappa \mathcal{L}_0 w\cdot v\mathrm{d}x - \int_{\partial_-\kappa\cap\Gamma_-}(b\cdot\mu)\,w^+v^+\mathrm{d}s - \int_{\partial_-\kappa\backslash\Gamma}(b\cdot\mu)\lfloor w\rfloor\,v^+\mathrm{d}s\right)$$

and the linear functional

$$\ell_A(v) = \sum_{\kappa\in\mathcal{T}}\left(\int_\kappa f v\mathrm{d}x - \int_{\partial_-\kappa\cap\Gamma_-}(b\cdot\mu)\,g_{\rm D}\,v^+\mathrm{d}s\right).$$

The *hp*–DGFEM for (3.1) is defined as follows: find $u_{\rm DG} \in S^{\mathbf{p}}(\Omega, \mathcal{T}, \mathbf{F})$ such that

$$A(u_{\rm DG}, v) = \ell_A(v) \quad \forall v \in S^{\mathbf{p}}(\Omega, \mathcal{T}, \mathbf{F}) \; . \tag{3.4}$$

The next three sections are devoted to the error analysis of this method.

3.2 Stability analysis of the DGFEM

Let us define $\| \cdot \|_\tau$, $\tau \subset \partial\kappa$, as the (semi)norm associated with the (semi)inner–product

$$(v, w)_\tau = \int_\tau |b \cdot \mu| vw \mathrm{d}s \; .$$

The next result is a special case of Lemma 2.4 in [15] with $\delta = 0$.

Lemma 1 *The solution $u_{\rm DG}$ to (3.4) satisfies the following bound:*

$$\sum_{\kappa \in \mathcal{T}} \left(\|c_0 u_{\rm DG}\|^2_{L^2(\kappa)} + \frac{1}{2} \|u^+_{\rm DG}\|^2_{\partial_- \kappa \cap \Gamma_-} + \|u^+_{\rm DG} - u^-_{\rm DG}\|^2_{\partial_- \kappa \setminus \Gamma} + \|u^+_{\rm DG}\|^2_{\partial_+ \kappa \cap \Gamma} \right)$$
$$\leq \sum_{\kappa \in \mathcal{T}} \left(\|{c_0}^{-1} f\|^2_{L^2(\kappa)} + 2 \|g_{\rm D}\|^2_{\partial_- \kappa \cap \Gamma_-} \right) .$$

Lemma 1 implies the uniqueness of the solution to the *hp*–DGFEM (3.4); further, since (3.4) is a linear problem over the finite–dimensional space $S^{\mathbf{p}}(\Omega, \mathcal{T}, \mathbf{F})$, the existence of the solution $u_{\rm DG}$ follows from its uniqueness.

According to the classical theory of characteristics, a piecewise continuous solution u to the first–order linear hyperbolic equation (3.1) can only exhibit jump discontinuities across characteristic hypersurfaces. Thus, for any smooth, open, $(d-1)$–dimensional hypersurface $\mathcal{S} \subset \Omega$ with normal vector μ, the normal flux of the solution, $bu \cdot \mu$, is a continuous function across $\mathcal{S}$ even if u itself has a jump discontinuity across $\mathcal{S}$ (for in the latter case $b \cdot \mu = 0$ on $\mathcal{S}$). Thus, $(b \cdot \mu) \lfloor u \rfloor = \lfloor bu \rfloor \cdot \mu = 0$ on $\mathcal{S}$, and so the method (3.4) is fully consistent, i.e.,

$$A(u, v) = \ell_A(v) \quad \forall v \in S^{\mathbf{p}}(\Omega, \mathcal{T}, \mathbf{F}) \; .$$

Combining this with (3.4) yields the Galerkin orthogonality property

$$A(u - u_{\rm DG}, v) = 0 \quad \forall v \in S^{\mathbf{p}}(\Omega, \mathcal{T}, \mathbf{F}) \; . \tag{3.5}$$

It will be assumed in the proceeding error analysis that

$$b \cdot \nabla_{\mathcal{T}} v_h \in S^{\mathbf{p}}(\Omega, \mathcal{T}, \mathbf{F}) \qquad \forall v_h \in S^{\mathbf{p}}(\Omega, \mathcal{T}, \mathbf{F}) \; . \tag{3.6}$$

The condition (3.6) will be further commented on in Remark 4 below.

Let us denote by Π_p the orthogonal projector in $L^2(\Omega)$ onto the finite element space $S^{\mathbf{p}}(\Omega, \mathcal{T}, \mathbf{F})$; i.e., given that $u \in L^2(\Omega)$, we define $\Pi_p u$ by

$$(u - \Pi_p u, v) = 0 \quad \forall v \in S^{\mathbf{p}}(\Omega, \mathcal{T}, \mathbf{F}) \; . \tag{3.7}$$

We may then decompose the global error $u - u_{\rm DG}$ as

$$u - u_{\rm DG} = (u - \Pi_p u) + (\Pi_p u - u_{\rm DG}) \equiv \eta + \xi \; . \tag{3.8}$$

Lemma 2 *Assume that (3.2) and (3.6) hold and let* $\gamma_1 = \text{ess sup}_{x\in\Omega}\,|c_1(x)|$ *where* $c_1(x) = (c(x) - (\nabla\cdot b)(x))\,/(c_0(x))^2$*; then the functions* ξ *and* η *defined by (3.8) satisfy the inequality*

$$\sum_{\kappa\in\mathcal{T}}\left(\|c_0\xi\|^2_{L^2(\kappa)} + \|\xi^+\|^2_{\partial_-\kappa\cap\Gamma_-} + \frac{1}{2}\|\xi^+ - \xi^-\|^2_{\partial_-\kappa\setminus\Gamma} + \frac{1}{2}\|\xi^+\|^2_{\partial_+\kappa\cap\Gamma}\right)$$
$$\leq \sum_{\kappa\in\mathcal{T}}\left(\gamma_1^2\|c_0\eta\|^2_{L^2(\kappa)} + 2\|\eta^+\|^2_{\partial_+\kappa\cap\Gamma} + 2\|\eta^-\|^2_{\partial_-\kappa\setminus\Gamma}\right). \tag{3.9}$$

Proof From (3.5) and (3.8) we have that

$$A(\xi,\xi) = -A(\eta,\xi)\,. \tag{3.10}$$

Let us first consider the left–hand side of (3.10). After integrating by parts and using the definition of $c_0(x)$ given in (3.3), we find, analogously as in Lemma 1, that

$$\begin{aligned} A(\xi,\xi) = &\sum_{\kappa\in\mathcal{T}}\|c_0\,\xi\|^2_{L^2(\kappa)} + \frac{1}{2}\sum_{\kappa\in\mathcal{T}}\|\xi^+\|^2_{\partial_-\kappa\cap\Gamma_-} \\ &+\frac{1}{2}\sum_{\kappa\in\mathcal{T}}\|\xi^+ - \xi^-\|^2_{\partial_-\kappa\setminus\Gamma} + \frac{1}{2}\sum_{\kappa\in\mathcal{T}}\|\xi^+\|^2_{\partial_+\kappa\cap\Gamma}\,. \end{aligned} \tag{3.11}$$

Similarly, performing integration by parts we deduce that

$$\begin{aligned} A(\eta,\xi) = &\,2\sum_{\kappa\in\mathcal{T}}\int_\kappa (c_0(x))^2\,\xi\eta\mathrm{d}x - \sum_{\kappa\in\mathcal{T}}\int_\kappa \eta\mathcal{L}_0\,\xi\mathrm{d}x + \sum_{\kappa\in\mathcal{T}}\int_{\partial_+\kappa\cap\Gamma}(b\cdot\mu)\xi^+\eta^+\mathrm{d}s \\ &+\sum_{\kappa\in\mathcal{T}}\int_{\partial_+\kappa\setminus\Gamma}(b\cdot\mu)\,\xi^+\eta^+\mathrm{d}s + \sum_{\kappa\in\mathcal{T}}\int_{\partial_-\kappa\setminus\Gamma}(b\cdot\mu)\xi^+\eta^-\mathrm{d}s\,. \end{aligned} \tag{3.12}$$

Concerning the last two terms in (3.12) we note that

$$\begin{aligned} &\left|\sum_{\kappa\in\mathcal{T}}\int_{\partial_+\kappa\setminus\Gamma}(b\cdot\mu)\,\xi^+\eta^+\mathrm{d}s + \sum_{\kappa\in\mathcal{T}}\int_{\partial_-\kappa\setminus\Gamma}(b\cdot\mu)\,\xi^+\eta^-\,\mathrm{d}s\right| \\ &\qquad\leq \sum_{\kappa\in\mathcal{T}}\|\xi^+ - \xi^-\|_{\partial_-\kappa\setminus\Gamma}\,\|\eta^-\|_{\partial_-\kappa\setminus\Gamma} \\ &\qquad\leq \frac{1}{4}\sum_{\kappa\in\mathcal{T}}\|\xi^+ - \xi^-\|^2_{\partial_-\kappa\setminus\Gamma} + \sum_{\kappa\in\mathcal{T}}\|\eta^-\|^2_{\partial_-\kappa\setminus\Gamma}\,. \end{aligned} \tag{3.13}$$

In addition, by virtue of (3.6),

$$\int_\kappa \eta\,(b\cdot\nabla\xi)\mathrm{d}x = 0 \quad \forall\kappa\in\mathcal{T}\,, \tag{3.14}$$

given that $b\cdot\nabla_{\mathcal{T}}\xi \in S^{\mathbf{p}}(\Omega,\mathcal{T},\mathbf{F})$ and, by (3.8), $\eta = u - \Pi_p u$ where Π_p is the L^2–projector onto $S^{\mathbf{p}}(\Omega,\mathcal{T},\mathbf{F})$, cf. (3.7). Noting the definition of $c_1(x)$ in the statement of the Lemma, (3.14) yields

$$2\sum_{\kappa\in\mathcal{T}}\int_\kappa (c_0(x))^2\,\xi\eta\mathrm{d}x - \sum_{\kappa\in\mathcal{T}}\int_\kappa \eta\,\mathcal{L}_0\,\xi\mathrm{d}x = \sum_{\kappa\in\mathcal{T}}\int_\kappa c_1(x)(c_0(x))^2\xi\eta\mathrm{d}x\,. \tag{3.15}$$

Using (3.13) and (3.15) in (3.12) then gives

$$A(\eta,\xi) \le \frac{1}{2}\sum_{\kappa\in\mathcal{T}} \|c_0\,\xi\|^2_{L^2(\kappa)} + \frac{1}{2}\gamma_1^2 \sum_{\kappa\in\mathcal{T}} \|c_0\eta\|^2_{L^2(\kappa)} + \frac{1}{4}\sum_{\kappa\in\mathcal{T}} \|\xi^+\|^2_{\partial_+\kappa\cap\Gamma} + \sum_{\kappa\in\mathcal{T}} \|\eta^+\|^2_{\partial_+\kappa\cap\Gamma} + \frac{1}{4}\sum_{\kappa\in\mathcal{T}} \|\xi^+ - \xi^-\|^2_{\partial_-\kappa\setminus\Gamma} + \sum_{\kappa\in\mathcal{T}} \|\eta^-\|^2_{\partial_-\kappa\setminus\Gamma}. \tag{3.16}$$

Combining (3.10), (3.11) and (3.16) gives the desired result. ∎

Stimulated by the identity (3.11), we define the *DG–norm* $|||\cdot|||_{\mathrm{DG}}$ by

$$|||w|||^2_{\mathrm{DG}} = \sum_{\kappa\in\mathcal{T}} \left(\|c_0 w\|^2_{L^2(\kappa)} + \frac{1}{2}\|w^+\|^2_{\partial_-\kappa\cap\Gamma_-} + \frac{1}{2}\|w^+ - w^-\|^2_{\partial_-\kappa\setminus\Gamma} + \frac{1}{2}\|w^+\|^2_{\partial_+\kappa\cap\Gamma} \right).$$

By applying the triangle inequality to (3.8) and using (3.9) we obtain a bound on the global error $u - u_{\mathrm{DG}}$ in the DG–norm in terms of the projection error $\eta = u - \Pi_p u$. Next, we derive a bound on the DG–norm of η in terms of h and p.

3.3 *hp*–error estimates

To obtain bounds on the projection error η in (3.9), explicit in h and p, we shall assume here for convenience that $\mathcal{T} = \{\kappa\}$ is a subdivision of Ω into shape–regular d-parallelepipeds, i.e., the reference element is $\hat{\kappa} = (-1,1)^d$. Inequality (3.9) shows that in addition to $\|\eta\|_{L^2(\kappa)}$ we also need to estimate the norms $\|\eta^+\|_{\partial_+\kappa}$, $\|\eta^-\|_{\partial_-\kappa}$; these terms will be dealt with by bounding them above by $\|b\|^{1/2}_{L^\infty(\kappa)}\|\eta\|_{L^2(\partial\kappa)}$. We begin our analysis by recalling the following univariate bound from Theorem 3.11 in [29].

Lemma 3 *Let $I = (-1,1)$ and $\hat{u} \in H^k(I)$ for some integer $k \ge 1$. Let further $\hat{\Pi}_p\hat{u}$ be the $L^2(I)$–projection of $\hat{u}$ onto $\mathcal{P}_p(I)$, $p \ge 0$. Then the following error estimate holds for any integer s, $0 \le s \le \min(p+1,k)$, with $W = W(\hat{x}) = (1-\hat{x}^2)^{1/2}$:*

$$\|\hat{u} - \hat{\Pi}_p\hat{u}\|^2_{L^2(I)} \le \frac{\Gamma(p+2-s)}{\Gamma(p+2+s)}\,\|W^s\hat{u}^{(s)}\|^2_{L^2(I)} \le \frac{\Gamma(p+2-s)}{\Gamma(p+2+s)}\,|\hat{u}|^2_{H^s(I)}\,. \tag{3.17}$$

Error estimates in dimension $d > 1$ will now be deduced by tensor product construction. To this end, we denote by $\hat{\Pi}^{(i)}_p$, $1 \le i \le d$, the univariate $L^2(I)$–projector onto the polynomials of degree p in the variable $\hat{x}_i$; $\hat{\Pi}_p$ will denote the $L^2(\hat{\kappa})$–projector onto the tensor product polynomials of degree p in each variable. Then

$$\hat{\Pi}_p = \hat{\Pi}^{(1)}_p\,\hat{\Pi}^{(2)}_p \ldots \hat{\Pi}^{(d)}_p\,. \tag{3.18}$$

Error estimates for $\hat{u} - \hat{\Pi}_p\hat{u}$ can now be obtained from (3.17). Thus consider $d = 2$, for example; then $\hat{u} - \hat{\Pi}_p\hat{u} = \hat{u} - \hat{\Pi}^{(1)}_p\,\hat{\Pi}^{(2)}_p\hat{u} = \hat{u} - \hat{\Pi}^{(1)}_p\hat{u} + \hat{\Pi}^{(1)}_p(\hat{u} - \hat{\Pi}^{(2)}_p\hat{u})$. Hence, recalling that $\hat{\Pi}^{(i)}_p$, $i = 1,2$, are bounded linear operators in $L^2(\hat{\kappa})$ with norm 1,

$$\|\hat{u} - \hat{\Pi}_p\hat{u}\|_{L^2(\hat{\kappa})} \le \|\hat{u} - \hat{\Pi}^{(1)}_p\hat{u}\|_{L^2(\hat{\kappa})} + \|\hat{u} - \hat{\Pi}^{(2)}_p\hat{u}\|_{L^2(\hat{\kappa})}\,. \tag{3.19}$$

If $d > 2$, we iterate (3.19) and obtain

$$\|\hat{u} - \hat{\Pi}_p \hat{u}\|_{L^2(\hat{\kappa})} \leq \sum_{i=1}^{d} \|\hat{u} - \hat{\Pi}_p^{(i)} \hat{u}\|_{L^2(\hat{\kappa})} \ . \tag{3.20}$$

Employing the bound (3.17) in (3.20) we arrive at the following result.

Lemma 4 *Let $\hat{\kappa} = (-1,1)^d$, $d \geq 1$, and $\hat{u} \in H^k(\hat{\kappa})$ for some integer $k \geq 1$. Let further $\hat{\Pi}_p \hat{u}$ be the $L^2(\hat{\kappa})$ projection of $\hat{u}$ onto $\mathcal{Q}_p(\hat{\kappa})$ with $p \geq 0$; then, for any integer s, $0 \leq s \leq \min(p+1, k)$, and $W_i = W_i(\hat{x}_i) = (1 - \hat{x}_i^2)^{1/2}$, we have*

$$\begin{aligned} \|\hat{u} - \hat{\Pi}_p \hat{u}\|_{L^2(\hat{\kappa})} &\leq \left(\frac{\Gamma(p+2-s)}{\Gamma(p+2+s)} \right)^{\frac{1}{2}} \sum_{i=1}^{d} \|W_i^s \partial_i^s \hat{u}\|_{L^2(\hat{\kappa})} \\ &\leq C(d) \left(\frac{\Gamma(p+2-s)}{\Gamma(p+2+s)} \right)^{\frac{1}{2}} |\hat{u}|_{H^s(\hat{\kappa})} \ . \end{aligned} \tag{3.21}$$

Applying Lemma 4, we can now deduce a bound on the L^2–projection error η on each element $\kappa \in \mathcal{T}$. Recall that $\mathcal{T}$ is shape–regular and that $\kappa = F_\kappa(\hat{\kappa})$ with F_κ affine. Setting $\hat{u} = u \circ F_\kappa$ in (3.21) and noting that $(\hat{\Pi}_p \hat{u})(\hat{x}) = (\Pi_p u)(F_\kappa(\hat{x}))$ for all $\hat{x} \in \hat{\kappa}$, we find that for any $\kappa \in \mathcal{T}$ and $u \in H^s(\kappa)$,

$$\|u - \Pi_p u\|_{L^2(\kappa)} \leq C(d)\, h^s \left(\frac{\Gamma(p+2-s)}{\Gamma(p+2+s)} \right)^{\frac{1}{2}} |u|_{H^s(\kappa)} , \quad 0 \leq s \leq \min(p+1, k) \ , \tag{3.22}$$

where Π_p is defined by (3.7).

We see from (3.9) that in addition to bounding $\|u - \Pi_p u\|_{L^2(\kappa)}$ we also need to estimate $\|u - \Pi_p u\|_{L^2(\partial\kappa)}$. The usual approach to handling the latter term is to apply the multiplicative trace inequality

$$\|\eta\|_{L^2(\partial\kappa)}^2 \leq C(d) \left(\|\eta\|_{L^2(\kappa)} \|\nabla \eta\|_{L^2(\kappa)} + h_\kappa^{-1} \|\eta\|_{L^2(\kappa)}^2 \right) \tag{3.23}$$

with $\eta = u - \Pi_p u$ and employ estimates of η and $\nabla\eta$ in the $L^2(\kappa)$ norm. While (3.22) and Stirling's formula show that $\|\eta\|_{L^2(\kappa)}$ exhibits an hp–optimal rate of convergence, all available bounds in the literature on $\|\nabla\eta\|_{L^2(\kappa)}$ are h–optimal but p–suboptimal, resulting in a p–suboptimal bound on the term $\|\eta\|_{L^2(\partial\kappa)}$. To overcome this problem, instead, we directly estimate the $L^2(\kappa)$–projection error $u - \Pi_p u$ in the $L^2(\partial\kappa)$ norm. First, we map κ onto the canonical element $\hat{\kappa}$ and consider, without loss of generality, the trace of $\hat{u} - \hat{\Pi}_p \hat{u}$ on the face $x_1 = 1$. We shall, again, use a tensor product argument, the main ingredients of which will be (3.17) and the following result.

Lemma 5 *Let $I = (-1,1)$, $\hat{u} \in H^k(I)$ for some integer $k \geq 1$, and let $\hat{\Pi}_p\,\hat{u} \in \mathcal{P}_p(I)$ be its $L^2(I)$–projection with $p \geq 0$; then,*

$$\begin{aligned} |(\hat{u} - \hat{\Pi}_p\,\hat{u})(1)|^2 &\leq \frac{1}{2p+1}\;\frac{\Gamma(p+1-t)}{\Gamma(p+1+t)}\;\|W^t\hat{u}^{(t+1)}\|^2_{L^2(I)} \\ &\leq \frac{1}{2p+1}\;\frac{\Gamma(p+1-t)}{\Gamma(p+1+t)}\;|\hat{u}|^2_{H^{t+1}(I)}\;, \end{aligned}$$

for any integer t, $0 \leq t \leq \min(p, k-1)$, where $W(\hat{x}) = (1-\hat{x}^2)^{1/2}$.

Proof For $p = 0$ the proof is trivial. Let us suppose, therefore, that $p \geq 1$. We develop $\hat{u}' \in L^2(I)$ into a Legendre series as a function of $\hat{x} \in I = (-1,1)$:

$$\hat{u}' = \sum_{i=0}^{\infty} b_i L_i\,, \quad b_i = \frac{2i+1}{2}\;\int_{-1}^{1} \hat{u}'(\hat{x})\,L_i(\hat{x})\mathrm{d}\hat{x}\;,$$

where $L_i(\hat{x})$ is the Legendre polynomial of degree i on $(-1,1)$; then

$$\hat{u}(\hat{x}) = \hat{u}(-1) + \sum_{i=0}^{\infty} b_i\;\int_{-1}^{\hat{x}} L_i(\zeta)\,\mathrm{d}\zeta\;.$$

Since

$$\int_{-1}^{\hat{x}} L_i(\zeta)\mathrm{d}\zeta = \left(L_{i+1}(\hat{x}) - L_{i-1}(\hat{x})\right)/\left(2i+1\right), \qquad i \geq 1\,,$$

we find that

$$\hat{u} = \left(b_0 + \hat{u}(-1)\right)\,L_0 + b_0 L_1 + \sum_{i=2}^{\infty}\;\frac{b_{i-1}}{2i-1}\,L_i - \sum_{i=0}^{\infty}\;\frac{b_{i+1}}{2i+3}\,L_i\;.$$

Comparing coefficients with $\hat{u} = \sum_{i=0}^{\infty}\,\hat{u}_i L_i$ gives

$$\hat{u}_i = \frac{b_{i-1}}{2i-1} - \frac{b_{i+1}}{2i+3}\;,\quad i \geq 2\;.$$

Thus, for $r \geq 2$,

$$\begin{aligned} \sum_{i=r}^{\infty}\,\hat{u}_i &= \sum_{i=r}^{\infty}\left(\frac{b_{i-1}}{2i-1} - \frac{b_{i+1}}{2i+3}\right) = \sum_{i=r-1}^{\infty}\;\frac{b_i}{2i+1} - \sum_{i=r+1}^{\infty}\;\frac{b_i}{2i+1} \\ &= \frac{b_{r-1}}{2r-1} + \frac{b_r}{2r+1}\;, \end{aligned}$$

and

$$\begin{aligned} \left(\sum_{i=r}^{\infty}\,\hat{u}_i\right)^2 &\leq \frac{2(b_{r-1})^2}{(2r-1)^2} + \frac{2(b_r)^2}{(2r+1)^2} \leq \frac{1}{2r-1}\;\sum_{i=r-1}^{\infty}\;\frac{2}{2i+1}\,|b_i|^2 \\ &\leq \frac{1}{2r-1}\;\|\hat{u}'\|^2_{L^2(I)}\;. \end{aligned} \tag{3.24}$$

Since $L_i(1) = 1$ for all i, (3.24) yields

$$|(\hat{u} - \hat{\Pi}_p \hat{u})(1)|^2 = \left(\sum_{i=p+1}^{\infty} \hat{u}_i \right)^2 \leq \frac{1}{2p+1} \, \|\hat{u}'\|^2_{L^2(I)} \, , \qquad p \geq 1 \, .$$

Now let $\hat{v} := \hat{u} - \hat{P}_p \, \hat{u}$ where the projector $\hat{P}_p$ is defined by

$$(\hat{P}_p w)(\hat{x}) := w(-1) + \int_{-1}^{\hat{x}} (\hat{\Pi}_{p-1}(w'))(\zeta) \, \mathrm{d}\zeta \, , \qquad p \geq 1 \, , \quad w \in H^1(I) \, ;$$

then, from (3.17) applied with $s = t$ and $p \to p-1$, we have that

$$\begin{aligned} |(\hat{u} - \hat{\Pi}_p \hat{u})(1)|^2 &= |\hat{u}(1) - (\hat{P}_p \, \hat{u})(1) + (\hat{P}_p \hat{u})(1) - (\hat{\Pi}_p \hat{u})(1)|^2 \\ &= |(\hat{u} - \hat{P}_p \hat{u})(1) - \hat{\Pi}_p (\hat{u} - \hat{P}_p \hat{u})(1)|^2 = |(\hat{v} - \hat{\Pi}_p \hat{v})(1)|^2 \\ &\leq \frac{1}{2p+1} \, \|\hat{v}'\|^2_{L^2(I)} = \frac{1}{2p+1} \, \|\hat{u}' - \hat{\Pi}_{p-1} \, \hat{u}'\|^2_{L^2(I)} \\ &\leq \frac{1}{2p+1} \, \frac{\Gamma(p+1-t)}{\Gamma(p+1+t)} \, \|W^t \hat{u}^{(t+1)}\|^2_{L^2(I)} \, . \end{aligned}$$

for $0 \leq t \leq \min(p, k-1)$. ∎

We shall now use Lemma 5 to estimate $\|\hat{u} - \hat{\Pi}_p \hat{u}\|_{L^2(\partial\hat{\kappa})}$, $\hat{\kappa} = (-1,1)^d$.

Lemma 6 *Suppose that $\hat{u} \in H^k(\hat{\kappa})$ for some integer $k \geq 1$, and let s be an integer such that $1 \leq s \leq \min(p+1, k)$, with $p \geq 0$; then, we have that*

$$\|\hat{u} - \hat{\Pi}_p \hat{u}\|_{L^2(\partial\hat{\kappa})} \leq C(d) \Phi_1(s,p) |\hat{u}|_{H^s(\hat{\kappa})} \, , \tag{3.25}$$

where $\Phi_1(s,p)$ is defined by

$$\begin{aligned} \Phi_1(s,p) = \frac{1}{\sqrt{2p+1}} \Bigg[\left(\frac{\Gamma(p+2-s)}{\Gamma(p+s)} \right)^{\frac{1}{2}} + \left(\frac{\Gamma(p+3-s)}{\Gamma(p+1+s)} \right)^{\frac{1}{2}} \Bigg] \\ + \left(\frac{\Gamma(p+2-s)}{\Gamma(p+2+s)} \right)^{\frac{1}{4}} \left(\frac{\Gamma(p+3-s)}{\Gamma(p+1+s)} \right)^{\frac{1}{4}} + \left(\frac{\Gamma(p+2-s)}{\Gamma(p+2+s)} \right)^{\frac{1}{2}} \, . \end{aligned} \tag{3.26}$$

Proof We write

$$\hat{\kappa} = I^{(1)} \times I^{(2)} \times \ldots \times I^{(d)} \, , \qquad \hat{x} = (\hat{x}_1, \ldots, \hat{x}_d) \equiv (\hat{x}_1, \hat{x}') \, , \quad \hat{x}_i \in I^{(i)} \, ,$$

where $I^{(i)}$ is the interval $(-1,1)$ in the ith–coordinate direction. Further, we define $\hat{\kappa}' \subset \partial\hat{\kappa}$ via $\hat{\kappa} = I^{(1)} \times \hat{\kappa}'$, and we split $\hat{\Pi}_p$ in (3.18) as $\hat{\Pi}_p = \hat{\Pi}_p^{(1)} \, \hat{\Pi}'_p$. We then have

$$\begin{aligned} \|(\hat{u} - \hat{\Pi}_p \hat{u})(1,\cdot)\|_{L^2(\hat{\kappa}')} &\leq \|(\hat{u} - \hat{\Pi}_p^{(1)} \hat{u})(1,\cdot)\|_{L^2(\hat{\kappa}')} + \|\hat{\Pi}_p^{(1)} (\hat{u} - \hat{\Pi}'_p \hat{u})(1,\cdot)\|_{L^2(\hat{\kappa}')} \\ &\equiv \mathrm{T}_1 + \mathrm{T}_2 \, . \end{aligned} \tag{3.27}$$

The term T_1 in (3.27) can be estimated using Lemma 5:

$$\mathrm{T}_1 \equiv \|(\hat{u} - \hat{\Pi}_p^{(1)} \hat{u})(1,\cdot)\|_{L^2(\hat{\kappa}')} \leq \frac{1}{\sqrt{2p+1}} \left(\frac{\Gamma(p+2-s)}{\Gamma(p+s)} \right)^{\frac{1}{2}} |\hat{u}|_{H^s(\hat{\kappa})} \, , \tag{3.28}$$

for $1 \leq s \leq \min(p+1,k)$, $k \geq 1$. We define $w := \hat{u} - \hat{\Pi}'_p\,\hat{u}$, and note that

$$\begin{aligned} \mathrm{T}_2 &\equiv \|\hat{\Pi}^{(1)}_p w(1,\cdot)\|_{L^2(\hat{\kappa}')} \leq \|w(1,\cdot)\|_{L^2(\hat{\kappa}')} + \|(w - \hat{\Pi}^{(1)}_p w)(1,\cdot)\|_{L^2(\hat{\kappa}')} \\ &\equiv \mathrm{T}_{21} + \mathrm{T}_{22}\ . \end{aligned} \tag{3.29}$$

Letting $\hat{\partial}_i \equiv \partial_{\hat{x}_i}$, we use Lemma 5 on the second term in (3.29) to deduce that

$$\mathrm{T}_{22} \equiv \|(w - \hat{\Pi}^{(1)}_p w)(1,\cdot)\|_{L^2(\hat{\kappa}')} \leq \frac{1}{\sqrt{2p+1}}\ \|\hat{\partial}_1 w\|_{L^2(\hat{\kappa})}\ . \tag{3.30}$$

Since $\hat{\partial}_1 w = \hat{\partial}_1 \hat{u} - \hat{\Pi}'_p\,(\hat{\partial}_1 \hat{u})$, we get from (3.21), applied with respect to $\hat{x}'$, the bound

$$\|\hat{\partial}_1 w\|_{L^2(\hat{\kappa})} \leq C(d)\ \left(\frac{\Gamma(p+3-s)}{\Gamma(p+1+s)}\right)^{\frac{1}{2}}\ |\hat{u}|_{H^s(\hat{\kappa})}\ . \tag{3.31}$$

Inserting (3.31) into (3.30) yields

$$\mathrm{T}_{22} \leq C(d)\ \frac{1}{\sqrt{2p+1}} \left(\frac{\Gamma(p+3-s)}{\Gamma(p+1+s)}\right)^{\frac{1}{2}}\ |\hat{u}|_{H^s(\hat{\kappa})}\,, \tag{3.32}$$

for $1 \leq s \leq \min(p+1,k)$, $k \geq 1$. To bound the term T_{21}, we note that by a univariate multiplicative trace inequality in the $\hat{x}_1$–direction, integrated over $\hat{x}' \in \hat{\kappa}'$:

$$\mathrm{T}_{21} = \|w(1,\cdot)\|_{L^2(\hat{\kappa}')} \leq C\left(\|w\|^{\frac{1}{2}}_{L^2(\hat{\kappa})} \|\hat{\partial}_1 w\|^{\frac{1}{2}}_{L^2(\hat{\kappa})} + \|w\|_{L^2(\hat{\kappa})}\right)\ . \tag{3.33}$$

Further, applying (3.21) with respect to $\hat{x}'$, we get

$$\|w\|_{L^2(\hat{\kappa})} = \|\hat{u} - \hat{\Pi}'_p \hat{u}\|_{L^2(\hat{\kappa})} \leq C(d) \left(\frac{\Gamma(p+2-s)}{\Gamma(p+2+s)}\right)^{\frac{1}{2}} |\hat{u}|_{H^s(\hat{\kappa})}\ . \tag{3.34}$$

On inserting (3.34) and (3.31) into (3.33), we find that

$$\mathrm{T}_{21} \leq C(d) \left(\left(\frac{\Gamma(p+2-s)}{\Gamma(p+2+s)}\right)^{\frac{1}{4}} \left(\frac{\Gamma(p+3-s)}{\Gamma(p+1+s)}\right)^{\frac{1}{4}} + \left(\frac{\Gamma(p+2-s)}{\Gamma(p+2+s)}\right)^{\frac{1}{2}}\right) |\hat{u}|_{H^s(\hat{\kappa})}\ ,$$

for $1 \leq s \leq \min(p+1,k)$, $k \geq 1$. Finally, substituting this last bound and (3.32) into (3.29), and then inserting the resulting inequality and (3.28) into (3.27), we get

$$\|(\hat{u} - \hat{\Pi}_p \hat{u})(1,\cdot)\|_{L^2(\hat{\kappa}')} \ \leq\ C(d)\Phi_1(s,p)|\hat{u}|_{H^s(\hat{\kappa})}\ ,$$

with Φ_1 as in (3.26) and $1 \leq s \leq \min(p+1,k)$, $k \geq 1$. An identical argument for each of the other faces of $\hat{\kappa}$ and merging the resulting bounds completes the proof. ∎

The next result is the weighted-norm-analogue of Lemma 6; its proof is analogous, except now it exploits the weighted-norm-bounds from Lemmas 4 and 5.

Lemma 7 *Suppose that $k \geq 1$ is an integer such that $N_\ell(k,u)$, $\ell = 1,2,3,4$, defined below are finite, and assume that s is an integer such that $1 \leq s \leq \min(p+1,k)$ with $p \geq 0$; then, we have that*

$$\begin{aligned}\|\hat{u} - \hat{\Pi}_p \hat{u}\|_{L^2(\partial\hat{\kappa})} \leq C(d) \{ & A_1(s,p)N_1(s,u) + A_2(s,p)N_2(s,u) \\ & + [A_3(s,p)N_3(s,u)A_4(s,p)N_4(s,u)]^{\frac{1}{2}} + A_3(s,p)N_3(s,u)\} ,\end{aligned}$$

where

$$A_1(s,p) = \frac{1}{\sqrt{2p+1}} \left(\frac{\Gamma(p+2-s)}{\Gamma(p+s)}\right)^{\frac{1}{2}}, \quad N_1(s,u) = \left(\sum_{i=1}^{d} \|W_i^{s-1}\hat{\partial}_i^s \hat{u}\|^2_{L^2(\hat{\kappa})}\right)^{\frac{1}{2}},$$

$$A_2(s,p) = \frac{1}{\sqrt{2p+1}} \left(\frac{\Gamma(p+3-s)}{\Gamma(p+1+s)}\right)^{\frac{1}{2}}, \quad N_2(s,u) = \left(\sum_{j=1}^{d}\sum_{i\neq j} \|W_i^{s-1}\hat{\partial}_j\hat{\partial}_i^{s-1} \hat{u}\|^2_{L^2(\hat{\kappa})}\right)^{\frac{1}{2}},$$

$$A_3(s,p) = \left(\frac{\Gamma(p+2-s)}{\Gamma(p+2+s)}\right)^{\frac{1}{2}}, \qquad N_3(s,u) = \left(\sum_{j=1}^{d}\sum_{i\neq j} \|W_i^{s}\hat{\partial}_i^{s} \hat{u}\|^2_{L^2(\hat{\kappa})}\right)^{\frac{1}{2}},$$

$$A_4(s,p) = \left(\frac{\Gamma(p+3-s)}{\Gamma(p+1+s)}\right)^{\frac{1}{2}}, \qquad N_4(s,u) = N_2(s,u) ,$$

with $\hat{\kappa} = (-1,1)^d$ and $W_i \equiv W_i(\hat{x}_i) = (1-\hat{x}_i^2)^{\frac{1}{2}}$ for $\hat{x}_i \in (-1,1)$, $i = 1,\ldots,d$.

Remark 1 *The analytical results in this section were stated under the assumption that $\hat{u} = u \circ F_\kappa$ belongs to an integer-order (weighted) Sobolev space on $\hat{\kappa}$. For the purposes of the discussion in this remark we note, however, that using the K-method of function space interpolation the bounds in Lemmas 3 to 7 can be extended to fractional-order spaces.*

Consider $\hat{u}(\hat{x}) = (1+\hat{x}_1)^\alpha$ with $\alpha > 1/2$ for $\hat{x} = (\hat{x}_1,\ldots,\hat{x}_d)$ in $\hat{\kappa} = (-1,1)^d$. Clearly, $\hat{\partial}_1^s \hat{u} = C_{\alpha,s}(1+\hat{x}_1)^{\alpha-s}$ and $\hat{\partial}_i \hat{u} = 0$ for $i \geq 2$, so $u \in H^s(\hat{\kappa})$ if and only if $s < \alpha + 1/2$. Hence, from (3.21), (3.25) via Stirling's formula, we get

$$\left(\|\hat{u} - \hat{\Pi}_p\hat{u}\|^2_{L^2(\hat{\kappa})} + \|\hat{u} - \hat{\Pi}_p\hat{u}\|^2_{L^2(\partial\hat{\kappa})}\right)^{1/2} \leq C_{\alpha,\varepsilon,d}\, p^{-\alpha+\varepsilon} ,$$

for ε, $0 < \varepsilon \ll 1$. It can be shown, however, that the expression on the left-hand side of this inequality decays faster than this bound predicts. Indeed, using the nomenclature introduced in Lemma 7, $N_2(s,u) = N_4(s,u) = 0$ and

$$\begin{aligned} A_1(s,p)N_1(s,u) &\leq C_{s,d}\, p^{-s+\frac{1}{2}} \|W_1^{s-1}\partial_1^s \hat{u}\|_{L^2(\hat{\kappa})} , \\ A_3(s,p)N_3(s,u) &\leq C_{s,d}\, p^{-s} \|W_1^{s}\partial_1^s \hat{u}\|_{L^2(\hat{\kappa})} . \end{aligned}$$

Now the expressions on the right–hand sides of these inequalities are finite provided that $s < 2\alpha$. Thus, we deduce from Lemma 7 that, for $0 < \varepsilon \ll 1$,

$$\|\hat{u} - \hat{\Pi}_p\hat{u}\|_{L^2(\partial\hat{\kappa})} \leq C_{\alpha,\varepsilon,d}\, p^{-\left(2\alpha-\frac{1}{2}\right)+\varepsilon} .$$

Further, upon choosing $s < 2\alpha + 1$ *in the first inequality of (3.21) to ensure that* $(1 - \hat{x}_1^2)^{s/2}\,(1+\hat{x}_1)^{\alpha-s}$ *lies in* $L^2(-1,1)$*, we have from Lemma 4 that*

$$\|\hat{u} - \hat{\Pi}_p \hat{u}\|_{L^2(\hat{\kappa})} \leq C_{\alpha,\varepsilon,d}\, p^{-(2\alpha+1)+\varepsilon}\,.$$

We thus conclude that

$$\left(\|\hat{u} - \hat{\Pi}_p \hat{u}\|^2_{L^2(\hat{\kappa})} + \|\hat{u} - \hat{\Pi}_p \hat{u}\|^2_{L^2(\partial\hat{\kappa})}\right)^{1/2} \leq C_{\alpha,\varepsilon,d}\, p^{-\left(2\alpha-\frac{1}{2}\right)+\varepsilon}\,. \tag{3.35}$$

Consider, on the other hand, the function $\hat{u}$ *defined by* $\hat{u}(\hat{x}) = (\max(0,\hat{x}_1))^\alpha$ *with* $\alpha > 0$*. Then, by an analogous argument we deduce that, in contrast with (3.35),*

$$\left(\|\hat{u} - \hat{\Pi}_p \hat{u}\|^2_{L^2(\hat{\kappa})} + \|\hat{u} - \hat{\Pi}_p \hat{u}\|^2_{L^2(\partial\hat{\kappa})}\right)^{1/2} \leq C_{\alpha,\varepsilon,d}\, p^{-\alpha+\varepsilon}$$

only. These observations will be of relevance in Section 6 where we investigate the sharpness of our error analysis through numerical experiments on model problems.

Now consider any $\kappa \in \mathcal{T}$. Recalling that the subdivision $\mathcal{T}$ is shape–regular and that $\kappa = F_\kappa(\hat{\kappa})$ with F_κ affine, on setting $\hat{u} = u \circ F_\kappa$ in (3.25) and noting that $(\hat{\Pi}_p \hat{u})(\hat{x}) = (\Pi_p u)(F_\kappa(\hat{x}))$, $\hat{x} \in \hat{\bar{\kappa}}$, we deduce from Lemma 6 the following result.

Lemma 8 *Let* $\kappa \in \mathcal{T}$ *and suppose that* $u \in H^k(\kappa)$ *for some integer* $k \geq 1$*. Then, for any integer* s*,* $1 \leq s \leq \min(p+1, k)$*, and* $p \geq 0$*, we have that*

$$\|u - \Pi_p u\|_{L^2(\partial\kappa)} \leq C(d)\Phi_1(s,p) h_\kappa^{s-\frac{1}{2}} |u|_{H^s(\kappa)}\,, \tag{3.36}$$

where $\Phi_1(s,p)$ *is defined by (3.26).*

Remark 2 *For fixed* $s \geq 1$*, by applying Stirling's formula we deduce that*

$$\Phi_1(p,s) \leq C(s)\,(p+1)^{-\left(s-\frac{1}{2}\right)}\,.$$

Consequently, (3.36) is of optimal order in both $p \geq 0$ *and* h*.*

3.4 *hp*–Convergence of the DGFEM

By (3.8), the triangle inequality and (3.9) we have that

$$\begin{aligned} |||u - u_{\mathrm{DG}}|||_{\mathrm{DG}} &\leq |||\xi|||_{\mathrm{DG}} + |||\eta|||_{\mathrm{DG}} \\ &\leq |||\eta|||_{\mathrm{DG}} + \left[\sum_{\kappa\in\mathcal{T}} \left(\gamma_1^2 \|c_0\eta\|^2_{L^2(\kappa)} + 2\|\eta^+\|^2_{\partial_+\kappa\cap\Gamma} + 2\|\eta^-\|^2_{\partial_-\kappa\setminus\Gamma}\right)\right]^{\frac{1}{2}}. \end{aligned} \tag{3.37}$$

To complete the error analysis, we substitute the estimates (3.22) and (3.36) into the right–hand side of (3.37). In addition to $\Phi_1(p,s)$ we also define

$$\Phi_2(p,s) = \left(\frac{\Gamma(p+2-s)}{\Gamma(p+2+s)}\right)^{\frac{1}{2}}\,.$$

The resulting error bound is formulated in the next theorem.

Theorem 9 *Let $\Omega \subset \mathbb{R}^d$ be a bounded polyhedral domain, $\mathcal{T} = \{\kappa\}$ a shape–regular subdivision into d-parallelepipeds κ with diameter h_κ. Let $u_{\rm DG} \in S^{\mathbf{p}}(\Omega, \mathcal{T}, \mathbf{F})$ be the discontinuous Galerkin approximation to u defined by (3.4) and suppose that $u|_\kappa \in H^{k_\kappa}(\kappa)$ for each $\kappa \in \mathcal{T}$, for integers $k_\kappa \geq 1$. Then, assuming that (3.2) and (3.6) are valid, the following error bound holds:*

$$|||u - u_{\rm DG}|||_{\rm DG}^2 \leq C \sum_{\kappa \in \mathcal{T}} h_\kappa^{2s_\kappa - 1} \left(\beta_\kappa \Phi_1^2(p_\kappa, s_\kappa) + \gamma_\kappa h_\kappa \Phi_2^2(p_\kappa, s_\kappa) \right) |u|_{H^{s_\kappa}(\kappa)}^2 ,$$

for any integers s_κ, $1 \leq s_\kappa \leq \min(p_\kappa + 1, k_\kappa)$, and $p_\kappa \geq 0$. Here,

$$\beta_\kappa = \|b\|_{L^\infty(\kappa)} , \qquad \gamma_\kappa = (1 + \gamma_1^2) \|c_0\|_{L^\infty(\kappa)}^2 ,$$

where c_0 and γ_1 are defined in (3.3) and Lemma 2, respectively, and C is a positive constant that depends only on the dimension d and the shape–regularity of $\mathcal{T}$.

Remark 3 *In particular, for uniform orders $p_\kappa = p \geq 0$, $s_\kappa = s$, $1 \leq s \leq \min(p+1, k)$, $k \geq 1$, and $h = \max_{\kappa \in \mathcal{T}} h_\kappa$, we get the bound*

$$|||u - u_{\rm DG}|||_{\rm DG} \leq C \left(\frac{h}{p+1} \right)^{s - \frac{1}{2}} |u|_{s,\mathcal{T}} . \tag{3.38}$$

The right–hand side in (3.38) is identical to the "optimal bound" $C(h/(p+1))^{s-\frac{1}{2}}|u|_{s,\mathcal{T}}$ which was obtained in [15] for a stabilized version of the hp–DGFEM for (3.1), with the streamline–diffusion stabilization parameter δ_κ in element κ chosen as $\delta_\kappa = h_\kappa/p_\kappa$. The present discussion corresponds to the case when $\delta_\kappa = 0$.

Remark 4 *The use of the L^2–projector Π_p in the definitions of ξ and η in (3.8) and the validity of the assumption (3.6) are essential ingredients of our analysis which relies on the fact that (3.14) holds. If (3.14) is violated, the present analysis yields an error bound in the $|||\cdot|||_{\rm DG}$ norm that is still optimal with respect to h but is p–suboptimal, by $p^{3/2}$. A possible remedy is to supplement the definition of the scheme with a streamline diffusion stabilization term as in [15], for example; this restores the hp–optimality of the error bound without hypothesis (3.6). Our numerical experiments in Section 6 indicate, however, that the method (3.4) is hp–optimal in the absence of hypothesis (3.6) even without streamline–diffusion stabilization. In fact, the numerical experiments in Section 6 show that the rate of p–convergence in the DG–norm may, in certain cases, even exceed the optimal rate of h–convergence. The source of this phenomenon has already been hinted at in Remark 1.*

4 Diffusion

4.1 DGFEM formulation

Now, let us consider the model problem (2.1) in the absence of the advection and reaction terms; i.e., we study the diffusion equation

$$\mathcal{L}_a u \equiv - \sum_{i,j=1}^{d} \partial_j (a_{ij}(x)\, \partial_i u) = f(x) , \quad x \in \Omega . \tag{4.1}$$

In the present section we shall assume that (4.1) is elliptic at each point $x \in \bar{\Omega}$; i.e., we strengthen (2.2) to

$$\zeta^{\mathrm{T}}\, a(x)\, \zeta > 0 \quad \forall \zeta \in \mathbb{R}^d \setminus \{0\}, \quad x \in \bar{\Omega}\ . \tag{4.2}$$

Then, it follows from (2.3) that $\Gamma \setminus \Gamma_0 = \emptyset$ in (2.4) and we complete (4.1) by the boundary conditions in (2.5), as in Section 2.1 but now with $\Gamma_- = \emptyset$ (still assuming that Γ_{D} is nonempty and relatively open in Γ). For simplicity of presentation, we suppose that the entries of the matrix a are constant on each element κ in $\mathcal{T}$; i.e.,

$$a \in \left[S^{\mathbf{0}}(\Omega, \mathcal{T}, \mathbf{F})\right]^{d\times d}_{\mathrm{sym}}\ . \tag{4.3}$$

Assuming that (4.3) holds, the matrix function a admits a unique square root $\sqrt{a} \in [S^{\mathbf{0}}(\Omega, \mathcal{T}, \mathbf{F})]^{d\times d}_{\mathrm{sym}}$ which again satisfies (4.2). We note that, with minor changes only, our results can be easily extended to the case of $\sqrt{a} \in [S^{\mathbf{q}}(\Omega, \mathcal{T}, \mathbf{F})]^{d\times d}_{\mathrm{sym}}$ where the composite polynomial degree vector $\mathbf{q}$ has nonnegative entries. In the following, we write $\bar{a} = |\sqrt{a}\,|^2_2$ where $|\cdot|_2$ denotes the matrix norm subordinate to the l^2 vector norm on $\mathbb{R}^d$ and $\bar{a}_\kappa = \bar{a}|_\kappa$; by $\bar{a}_{\tilde{\kappa}}$ we denote the arithmetic mean of the values $\bar{a}_{\kappa'}$ over those elements κ' (including κ itself) that share a $(d-1)$–dimensional face with κ.

We recall from Section 2.2 that $\mathcal{E}$ denotes the set of all open $(d-1)$–dimensional element faces associated with $\mathcal{T}$, and $\mathcal{E}_{\mathrm{int}}$ (resp. $\mathcal{E}_{\mathrm{D}}$) is the set all those open faces in $\mathcal{E}$ that lie inside Ω (resp. on Γ_{D}). Given that $e \in \mathcal{E}_{\mathrm{int}}$, there exist indices i and j such that $i > j$ and κ_i and κ_j share the interface e; we define the (element–numbering–dependent) jump of $v \in H^1(\Omega, \mathcal{T})$ 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. We note that, in general, $[v]$ is distinct from the jump $\lfloor v \rfloor$ defined in Section 3.1 in that the latter is independent of the element numbering. With each face $e \in \mathcal{E}_{\mathrm{int}}$ we associate the unit normal vector ν which points from κ_i to κ_j; on boundary faces, we put $\nu = \mu$. With this notation, we introduce the bilinear form

$$D(w, v) \;=\; B_a(w, v) + B_s(w, v)\ ,$$

where (cf. [1, 3, 24, 27, 33])

$$\begin{aligned} B_a(w, v) \;=\; & \sum_{\kappa\in\mathcal{T}} \int_\kappa a\nabla w \cdot \nabla v \mathrm{d}x + \int_{\Gamma_{\mathrm{D}}} \{w((a\nabla v)\cdot\mu) - ((a\nabla w)\cdot\mu)v\}\mathrm{d}s \\ & + \int_{\Gamma_{\mathrm{int}}} \{[w]\langle(a\nabla v)\cdot\nu\rangle - \langle(a\nabla w)\cdot\nu\rangle[v]\}\mathrm{d}s\ , \end{aligned} \tag{4.4}$$

$$B_s(w, v) \;=\; \int_{\Gamma_{\mathrm{D}}} \sigma w v \mathrm{d}s + \int_{\Gamma_{\mathrm{int}}} \sigma[w][v]\mathrm{d}s\ , \tag{4.5}$$

and the linear functional

$$\ell_{\mathrm{D}}(v) = \ell_a(v) + \ell_s(v)\ ,$$

where

$$\ell_a(v) = \sum_{\kappa\in\mathcal{T}} \int_\kappa f v \mathrm{d}x + \int_{\Gamma_\mathrm{D}} g_\mathrm{D}((a\nabla v)\cdot\mu)\mathrm{d}s + \int_{\Gamma_\mathrm{N}} g_\mathrm{N} v \mathrm{d}s \ ,$$
$$\ell_s(v) = \int_{\Gamma_\mathrm{D}} \sigma g_\mathrm{D} v \mathrm{d}s \ .$$

Here σ is called the *discontinuity–penalization* parameter, and is defined by

$$\sigma|_e = \sigma_e \quad \text{for} \ \ e \in \mathcal{E}_\mathrm{int} \cup \mathcal{E}_\mathrm{D} \ ,$$

where σ_e is a nonnegative constant on edge e. The precise choice of σ_e depends on a and the discretization parameters, and will be discussed in detail in the next section.

The *hp*–DGFEM for (4.1), (2.5) is: find $u_\mathrm{DG} \in S^\mathbf{p}(\Omega,\mathcal{T},\mathbf{F})$ such that

$$D(u_\mathrm{DG}, v) = \ell_\mathrm{D}(v) \quad \forall v \in S^\mathbf{p}(\Omega,\mathcal{T},\mathbf{F}) \ . \tag{4.6}$$

In order for (4.6) to be meaningful, it is necessary to assume that $p_\kappa \geq 1$, $\kappa \in \mathcal{T}$. Also, to ensure that the Galerkin orthogonality property $D(u - u_\mathrm{DG}, v) = 0$ holds for all $v \in S^\mathbf{p}(\Omega,\mathcal{T},\mathbf{F})$, we shall suppose throughout that the solution u to the elliptic boundary value problem under consideration is sufficiently smooth: namely, $u \in H^2(\Omega,\mathcal{T})$ and the functions u and $(a\nabla u)\cdot\nu$ are continuous across each face e in $\mathcal{E}_\mathrm{int}$. If this smoothness requirement is violated (as, for example, in an elliptic transmission problem), the discretization method (4.6) has to be modified accordingly.

Remark 5 *We note that when the discontinuity–penalization parameter σ is set to zero, the hp–DGFEM (4.6) is identical to the method introduced in [3, 24]. Other variants of the DGFEM for second–order uniformly elliptic problems have also been considered in the literature; see [24] for a comprehensive review.*

4.2 Analytical results

Our first result concerns the positivity of the bilinear form $D(\cdot,\cdot)$ and the existence and uniqueness of a solution to (4.6).

Theorem 10 *Let (4.2) and (4.3) hold; then, for every $w \in H^2(\Omega,\mathcal{T})$ we have*

$$|||w|||_\mathrm{DG}^2 \equiv D(w,w) = \sum_{\kappa\in\mathcal{T}} \|\sqrt{a}\,\nabla w\|_{L^2(\kappa)}^2 + \int_{\Gamma_\mathrm{D}} \sigma w^2 \,\mathrm{d}s + \int_{\Gamma_\mathrm{int}} \sigma [w]^2 \,\mathrm{d}s \ , \tag{4.7}$$

with $\sqrt{a}$ denoting the (positive definite) square–root of the symmetric matrix a, and σ is the (nonnegative) discontinuity–penalization parameter. Furthermore, if σ is positive on $\Gamma_\mathrm{int} \cup \Gamma_\mathrm{D}$ then the hp–DGFEM (4.6) has a unique solution u_DG in $S^\mathbf{p}(\Omega,\mathcal{T},\mathbf{F})$.

Proof Identity (4.7) follows trivially from (4.4) and (4.5). If now, in addition, σ is positive on $\Gamma_{\rm int} \cup \Gamma_{\rm D}$ then, since $a(x)$ is positive definite at each $x \in \bar{\Omega}$, it follows from (4.7) that $D(w,w) > 0$ for all w in $S^{\mathbf{p}}(\Omega, \mathcal{T}, \mathbf{F}) \setminus \{0\}$, and hence we deduce the uniqueness of the solution $u_{\rm DG}$. As the linear space $S^{\mathbf{p}}(\Omega, \mathcal{T}, \mathbf{F})$ is finite–dimensional and (4.6) is a linear problem, the existence of the solution to (4.6) follows from the fact that its homogeneous counterpart has the unique solution $u_{\rm DG} \equiv 0$. ■

As in Section 3.2, we decompose the global error as $u - u_{\rm DG} = \eta + \xi$ where $\eta = u - \Pi u$, $\xi = \Pi u - u_{\rm DG}$ and Π is a certain projector (whose specific choice is of no relevance at this point) onto the finite element space $S^{\mathbf{p}}(\Omega, \mathcal{T}, \mathbf{F})$.

Lemma 11 *Let $\mathcal{T}$ be a shape–regular subdivision of Ω and assume that the parameter σ is positive on $\Gamma_{\rm int} \cup \Gamma_{\rm D}$. Then, the following inequality holds, with C a positive constant that depends only on the dimension d and the shape regularity of $\mathcal{T}$:*

$$
\begin{aligned}
|||\xi|||^2_{\rm DG} \leq C \Bigg(& \int_{\Gamma_{\rm D}} \sigma |\eta|^2 \, {\rm d}s + \int_{\Gamma_{\rm int}} \sigma [\eta]^2 \, {\rm d}s + \sum_{\kappa \in \mathcal{T}} \| \sqrt{a} \, \nabla \eta \|^2_{L^2(\kappa)} \\
& + \sum_{\kappa \in \mathcal{T}} \left(\| \sqrt{\tau} \, \eta \|^2_{L^2(\partial\kappa \cap \Gamma_{\rm D})} + \bar{a}^2_\kappa \left\| \frac{1}{\sqrt{\sigma}} \nabla \eta \right\|^2_{L^2(\partial\kappa \cap \Gamma_{\rm D})} \right) \\
& + \sum_{\kappa \in \mathcal{T}} \left(\| \sqrt{\tau} \, [\eta] \|^2_{L^2(\partial\kappa \cap \Gamma_{\rm int})} + \bar{a}^2_\kappa \left\| \frac{1}{\sqrt{\sigma}} \nabla \eta \right\|^2_{L^2(\partial\kappa \cap \Gamma_{\rm int})} \right) \Bigg) \ , \qquad (4.8)
\end{aligned}
$$

where $\tau_e = \langle \bar{a} p^2 \rangle_e / h_e$ and h_e is the diameter of face $e \in \mathcal{E}_{\rm int} \cup \mathcal{E}_{\rm D}$, with the convention that for $e \in \mathcal{E}_{\rm D}$ contributions from outside Ω in the definition of τ_e are set to 0.

Proof By virtue of Theorem 10, we have

$$
|||\xi|||^2_{\rm DG} = D(\xi, \xi) = D((u - u_{\rm DG}) - \eta, \xi) = -D(\eta, \xi) \ ,
$$

where we have used the Galerkin orthogonality property $D(u - u_{\rm DG}, \xi) = 0$ which follows from (4.6) with $v = \xi$ and the definition of the boundary value problem (4.1), (2.5), given the assumed smoothness of u. Thus, we deduce that

$$
|||\xi|||^2_{\rm DG} \leq |B_a(\eta, \xi)| + |B_s(\eta, \xi)| \ .
$$

Now, from (4.5) we have that

$$
|B_s(\eta, \xi)| \leq |||\xi|||_{\rm DG} \left(\int_{\Gamma_{\rm D}} \sigma |\eta|^2 {\rm d}s + \int_{\Gamma_{\rm int}} \sigma [\eta]^2 {\rm d}s \right)^{\frac{1}{2}} \ . \qquad (4.9)
$$

Next,

$$
|B_a(\eta, \xi)| \leq I + II + III \ ,
$$

where

$$
I \equiv \left| \sum_{\kappa \in \mathcal{T}} \int_\kappa a \nabla \eta \cdot \nabla \xi {\rm d}x \right| \ , \quad II \equiv \left| \int_{\Gamma_{\rm D}} \{ \eta ((a \nabla \xi) \cdot \mu) - ((a \nabla \eta) \cdot \mu) \xi \} {\rm d}s \right| \ ,
$$

$$
III \equiv \left| \int_{\Gamma_{\rm int}} \{ [\eta] \langle (a \nabla \xi) \cdot \nu \rangle - \langle (a \nabla \eta) \cdot \nu \rangle [\xi] \} \, {\rm d}s \right| \ .
$$

For term I we have

$$I^2 \;\leq\; |||\xi|||^2_{\mathrm{DG}} \sum_{\kappa\in\mathcal{T}} \|\sqrt{a}\,\nabla\eta\|^2_{L^2(\kappa)}\,. \tag{4.10}$$

To deal with term II, we first note that

$$\begin{aligned} II \;\leq\; & \left(\sum_{\kappa\in\mathcal{T}} \frac{\bar{a}_\kappa}{\gamma_\kappa}\|\eta\|^2_{L^2(\partial\kappa\cap\Gamma_{\mathrm{D}})}\right)^{\frac{1}{2}} \left(\sum_{\kappa\in\mathcal{T}} \gamma_\kappa \|\sqrt{a}\,\nabla\xi\|^2_{L^2(\partial\kappa\cap\Gamma_{\mathrm{D}})}\right)^{\frac{1}{2}} \\ & + \left(\sum_{\kappa\in\mathcal{T}} \bar{a}^2_\kappa \left\|\frac{1}{\sqrt{\sigma}}\nabla\eta\right\|^2_{L^2(\partial\kappa\cap\Gamma_{\mathrm{D}})}\right)^{\frac{1}{2}} \left(\sum_{\kappa\in\mathcal{T}} \|\sqrt{\sigma}\,\xi\|^2_{L^2(\partial\kappa\cap\Gamma_{\mathrm{D}})}\right)^{\frac{1}{2}} \end{aligned} \tag{4.11}$$

for any set of positive real numbers γ_κ. As, by hypothesis, a is a constant matrix on each element $\kappa\in\mathcal{T}$, we can apply the inverse inequality

$$\|\sqrt{a}\,\nabla\xi\|^2_{L^2(\partial\kappa\cap\Gamma_{\mathrm{D}})} \leq C\,\frac{p^2_\kappa}{h_\kappa}\,\|\sqrt{a}\,\nabla\xi\|^2_{L^2(\kappa)}\,, \tag{4.12}$$

where C depends only on the shape–regularity of $\mathcal{T}$ (see [29], (4.6.4) of Theorem 4.76). On substituting (4.12) into (4.11), letting $\gamma_\kappa = h_\kappa/p^2_\kappa$ and defining $\tau_e = \bar{a}_\kappa p^2_\kappa/2h_e$ for a $(d-1)$-dimensional face $e\subset\partial\kappa\cap\Gamma_{\mathrm{D}}$, we arrive at the desired bound on II:

$$II^2 \leq C\,|||\xi|||^2_{\mathrm{DG}} \sum_{\kappa\in\mathcal{T}} \left(\|\sqrt{\tau}\,\eta\|^2_{L^2(\partial\kappa\cap\Gamma_{\mathrm{D}})} + \bar{a}^2_\kappa \left\|\frac{1}{\sqrt{\sigma}}\nabla\eta\right\|^2_{L^2(\partial\kappa\cap\Gamma_{\mathrm{D}})}\right)\,. \tag{4.13}$$

Similarly, we have

$$III^2 \;\leq\; C\,|||\xi|||^2_{\mathrm{DG}} \sum_{\kappa\in\mathcal{T}} \left(\|\sqrt{\tau}\,[\eta]\|^2_{L^2(\partial\kappa\cap\Gamma_{\mathrm{int}})} + \bar{a}^2_\kappa \left\|\frac{1}{\sqrt{\sigma}}\nabla\eta\right\|^2_{L^2(\partial\kappa\cap\Gamma_{\mathrm{int}})}\right). \tag{4.14}$$

Collecting the bounds (4.9), (4.10), (4.13) and (4.14) gives the desired result. ∎

Before embarking on the *a priori* error analysis of the *hp*–DGFEM (4.6), we state an approximation result for the finite element space $S^{\mathbf{p}}(\Omega,\mathcal{T},\mathbf{F})$.

Lemma 12 *Suppose that $\kappa\in\mathcal{T}$ is a d–simplex or a d–parallelepiped of diameter h_κ. Suppose further that $u|_\kappa \in H^{k_\kappa}(\kappa)$, $k_\kappa\geq 0$, for $\kappa\in\mathcal{T}$. Then, there exists a sequence $z^{h_\kappa}_{p_\kappa}(u)$ in $\mathcal{R}_{p_\kappa}(\kappa)$, $p_\kappa = 1,2,...$, such that for $0\leq q\leq k_\kappa$,*

$$\|u - z^{h_\kappa}_{p_\kappa}(u)\|_{H^q(\kappa)} \leq C\frac{h^{s_\kappa-q}_\kappa}{p^{k_\kappa-q}_\kappa}\|u\|_{H^{k_\kappa}(\kappa)}\,, \tag{4.15}$$

where $s_\kappa = \min(p_\kappa+1, k_\kappa)$ and C is a constant independent of u, h_κ and p_κ, but dependent of $k = \max_{\kappa\in\mathcal{T}} k_\kappa$.

Proof See Lemma 4.5 in [2] for $d = 2$; when $d > 2$ the proof is analogous. ∎

For $u \in H^2(\Omega, \mathcal{T})$, we now define $\Pi_p^h u \in S^{\mathbf{p}}(\Omega, F)$ by

$$(\Pi_p^h u)|_\kappa = z_{p_\kappa}^{h_\kappa}(u|_\kappa)\,, \qquad \kappa \in \mathcal{T}\ . \tag{4.16}$$

We shall assume in what follows that the polynomial degree vector $\mathbf{p}$, with $p_\kappa \geq 1$ for each $\kappa \in \mathcal{T}$, has *bounded local variation*, i.e., there exists a constant $\rho \geq 1$ such that, for any pair of elements κ and κ' which share a $(d-1)$–dimensional face,

$$\rho^{-1} \leq p_\kappa / p_{\kappa'} \leq \rho\ . \tag{4.17}$$

Our next result concerns the accuracy of the *hp*–version of the DGFEM (4.6).

Theorem 13 *Let $\Omega \subset \mathbb{R}^d$ be a bounded polyhedral domain, $\mathcal{T} = \{\kappa\}$ a shape–regular subdivision of Ω into d-parallelepipeds and $\mathbf{p}$ a polynomial degree vector of bounded local variation. Assign to each face $e \in \mathcal{E}_{\rm int} \cup \mathcal{E}_{\rm D}$ a positive real number*

$$\sigma_e = \frac{\langle \bar{a} p \rangle_e}{h_e}\ , \tag{4.18}$$

where h_e is the diameter of e, with the convention that for $e \in \mathcal{E}_{\rm D}$ contributions from outside Ω in the definition of σ_e are set to 0. Then, assuming that $u|_\kappa \in H^{k_\kappa}(\kappa)$, $k_\kappa \geq 2$, for $\kappa \in \mathcal{T}$, the solution $u_{\rm DG} \in S^{\mathbf{p}}(\Omega, \mathcal{T}, \mathbf{F})$ of (4.6) obeys the error bound

$$|||u - u_{\rm DG}|||_{\rm DG}^2 \leq C \sum_{\kappa \in \mathcal{T}} \alpha_\kappa \frac{h_\kappa^{2s_\kappa - 2}}{p_\kappa^{2k_\kappa - 3}} \|u\|^2_{H^{k_\kappa}(\kappa)}\ , \tag{4.19}$$

with $1 \leq s_\kappa \leq \min(p_\kappa + 1, k_\kappa)$, $p_\kappa \geq 1$ for $\kappa \in \mathcal{T}$, where $\alpha_\kappa = \bar{a}_{\tilde{\kappa}}$ and C is a positive constant depending only on d, the parameter ρ from (4.17), $k = \max_{\kappa \in \mathcal{T}} k_\kappa$, and the shape–regularity of $\mathcal{T}$.

Proof Consider the decomposition of the global error

$$u - u_{\rm DG} = (u - \Pi_p^h u) + (\Pi_p^h u - u) \equiv \eta + \xi\ ,$$

where Π_p^h is the projector defined by (4.16). Using the triangle inequality we get

$$|||u - u_{\rm DG}|||_{\rm DG} \leq |||\eta|||_{\rm DG} + |||\xi|||_{\rm DG} \equiv I + II\ .$$

Recalling the definition of the DG–norm (4.7), we have that

$$|||\eta|||_{\rm DG}^2 = \sum_{\kappa \in \mathcal{T}} \|\sqrt{a}\,\nabla \eta\|^2_{L^2(\kappa)} + \int_{\Gamma_{\rm D}} \sigma |\eta|^2 {\rm d}s + \int_{\Gamma_{\rm int}} \sigma [\eta]^2 {\rm d}s \equiv I_1 + I_2 + I_3.$$

We see from (4.8), with $\Pi = \Pi_p^h$, that the terms I_1, I_2 and I_3 are included in the right–hand side of (4.8); thus to obtain bounds on $|||\xi|||_{\rm DG}$ and $|||\eta|||_{\rm DG}$ it suffices to estimate each of the terms on the right–hand side of (4.8).

Now, the terms on the right–hand side of (4.8) fall into two categories; they either involve the L^2 norm over κ or the L^2 norm over (part of the) boundary of κ. For terms from the first category we find, using Lemma 12 with $q = 1$, that

$$\|\sqrt{a}\,\nabla\eta\|^2_{L^2(\kappa)} \leq C\bar{a}_\kappa \frac{h_\kappa^{2s_\kappa-2}}{p_\kappa^{2k_\kappa-2}} \|u\|^2_{H^{k_\kappa}(\kappa)} \,. \tag{4.20}$$

In order to deal with terms from the second category, we use (4.15) with $q = 0, 1$ and the multiplicative trace inequality (3.23) to deduce that

$$\|\eta\|^2_{L^2(\partial\kappa)} \leq C \frac{h_\kappa^{2s_\kappa-1}}{p_\kappa^{2k_\kappa-1}} \|u\|^2_{H^{k_\kappa}(\kappa)} \,. \tag{4.21}$$

Analogously, we have

$$\|\nabla\eta\|^2_{L^2(\partial\kappa)} \leq C \frac{h_\kappa^{2s_\kappa-3}}{p_\kappa^{2k_\kappa-3}} \|u\|^2_{H^{k_\kappa}(\kappa)} \,. \tag{4.22}$$

Applying these inequalities in the right–hand side of (4.8), choosing σ_e as in (4.18) and noting (4.17) and the shape regularity of $\mathcal{T}$ to relate h_e to h_κ, we deduce that

$$|||\xi|||^2_{\mathrm{DG}} \leq C \sum_{\kappa\in\mathcal{T}} \alpha_\kappa \left(\frac{h_\kappa^{2s_\kappa-2}}{p_\kappa^{2k_\kappa-2}} + \frac{p_\kappa^2}{h_\kappa} \frac{h_\kappa^{2s_\kappa-1}}{p_\kappa^{2k_\kappa-1}} \right) \|u\|^2_{H^{k_\kappa}(\kappa)} \,,$$

and hence (4.19). ∎

Remark 6 *If $k_\kappa = k \geq 2$ and $p_\kappa = p \geq k-1$ for each $\kappa \in \mathcal{T}$, then the estimate (4.19) becomes*

$$|||u - u_{\mathrm{DG}}|||_{\mathrm{DG}} \leq C \frac{h^{k-1}}{p^{k-3/2}} \|u\|_{k,\mathcal{T}} \,,$$

where $\|\cdot\|_{k,\mathcal{T}}$ is the broken H^k norm defined in (2.6). This bound is optimal in h and suboptimal in p by $p^{\frac{1}{2}}$, and coincides with that of Rivière, Wheeler and Girault [27].

Our error analysis, both in the case of pure advection considered in Section 3 and the case of pure diffusion here, is based on decomposing the global error as $u - u_{\mathrm{DG}} = (u - \Pi u) + (\Pi u - u_{\mathrm{DG}}) \equiv \xi + \eta$, where Π is a certain projector onto the finite element space $S^{\mathbf{p}}(\Omega, \mathcal{T}, \mathbf{F})$. However, our proofs in the two cases required different choices of Π so as to maximize the asymptotic convergence rates of the error bounds: in Section 3, Π was chosen as Π_p, the orthogonal projector in $L^2(\Omega)$ onto $S^{\mathbf{p}}(\Omega, \mathcal{T}, \mathbf{F})$, whereas in Theorem 13 we selected as Π the projector Π_p^h defined by (4.16). In the next section we shall consider advection–diffusion equations. The analysis there relies on combining the error bounds derived in the hyperbolic and elliptic cases. Thus, we also formulate a variant of Theorem 13 where, instead of Π_p^h, the proof makes use of the orthogonal projector in $L^2(\Omega)$ onto $S^{\mathbf{p}}(\Omega, \mathcal{T}, \mathbf{F})$ as Π. The resulting bound is still optimal in h, but is now p–suboptimal by a full power of p.

Theorem 14 *Let $\Omega \subset \mathbb{R}^d$ be a bounded polyhedral domain, $\mathcal{T} = \{\kappa\}$ a shape–regular subdivision of Ω into d-parallelepipeds and $\mathbf{p}$ a polynomial degree vector of bounded local variation. Let each face $e \in \mathcal{E}_{\rm int} \cup \mathcal{E}_{\rm D}$ be assigned a positive real number*

$$\sigma_e = \frac{\langle \bar{a} p^2 \rangle_e}{h_e} \ , \tag{4.23}$$

where h_e is the diameter of e. Then, assuming that $u|_\kappa \in H^{k_\kappa}(\kappa)$, $k_\kappa \geq 2$, for $\kappa \in \mathcal{T}$, the solution $u_{\rm DG} \in S^{\mathbf{p}}(\Omega, \mathcal{T}, \mathbf{F})$ of (4.6) obeys the error bound

$$|||u - u_{\rm DG}|||^2_{\rm DG} \leq C \sum_{\kappa \in \mathcal{T}} \alpha_\kappa \, \frac{h_\kappa^{2s_\kappa - 2}}{p_\kappa^{2k_\kappa - 4}} \|u\|^2_{H^{k_\kappa}(\kappa)} \ , \tag{4.24}$$

with $1 \leq s_\kappa \leq \min(p_\kappa + 1, k_\kappa)$, $p_\kappa \geq 1$, $\alpha_\kappa = \bar{a}_{\tilde{\kappa}}$ for $\kappa \in \mathcal{T}$; C is a constant depending on d, the parameter ρ from (4.17), $k = \max_\kappa k_\kappa$, and the shape–regularity of $\mathcal{T}$.

Proof The structure of the proof is the same as for Theorem 13, except now we decompose the global error as

$$u - u_{\rm DG} = (u - \Pi_p u) + (\Pi_p u - u_{\rm DG}) \equiv \eta + \xi \ , \tag{4.25}$$

where Π_p denotes the orthogonal projector in $L^2(\Omega)$ onto $S^{\mathbf{p}}(\Omega, \mathcal{T}, \mathbf{F})$. A new ingredient of the present proof is that, in contrast with Π_p^h, Π_p is not known to satisfy an hp–optimal bound of the type (4.15), except for $q = 0$. Thus, instead of bounding $\|\sqrt{a}\,\nabla\eta\|_{L^2(\kappa)}$, $\|\eta\|_{L^2(\partial\kappa)}$ and $\|\nabla\eta\|_{L^2(\partial\kappa)}$ directly, we now further decompose η as

$$\eta = u - \Pi_p u = (u - \Pi_p^h u) + \Pi_p^h (u - \Pi_p u) \equiv \eta_1 + \eta_2 \ . \tag{4.26}$$

The term η_1 is bounded directly using (4.15) and the multiplicative trace inequality (3.23), as in (4.20), (4.21) and (4.22). Norms of η_2, on the other hand, are dealt with by switching them to $\|\eta_2\|_{L^2(\kappa)}$ by means of the inverse inequalities

$$\|\sqrt{a}\,\nabla\eta_2\|^2_{L^2(\kappa)} \leq C\bar{a}_\kappa \frac{p_\kappa^4}{h_\kappa^2} \|\eta_2\|^2_{L^2(\kappa)} \ , \tag{4.27}$$

$$\|\eta_2\|^2_{L^2(\partial\kappa)} \leq C\frac{p_\kappa^2}{h_\kappa} \|\eta_2\|^2_{L^2(\kappa)} \, , \quad \|\nabla\eta_2\|^2_{L^2(\partial\kappa)} \leq C\frac{p_\kappa^6}{h_\kappa^3} \|\eta_2\|^2_{L^2(\kappa)} \ . \tag{4.28}$$

Now $\|\eta_2\|_{L^2(\kappa)}$ is further bounded above by

$$\|\eta_2\|_{L^2(\kappa)} = \|\Pi_p^h (u - \Pi_p u)\|_{L^2(\kappa)} \leq C\|u - \Pi_p u\|_{L^2(\kappa)} \leq C\frac{h_\kappa^{s_\kappa}}{p_\kappa^{s_\kappa}} |u|_{H^{s_\kappa}(\kappa)} \ , \tag{4.29}$$

for $1 \leq s_\kappa \leq \min(p_\kappa + 1, k_\kappa)$, where the first inequality follows from (4.15) with $q = k_\kappa = s_\kappa = 0$ and the second inequality is a consequence of (3.22). Hence,

$$\|\eta_2\|_{L^2(\kappa)} \leq C\frac{h_\kappa^{s_\kappa}}{p_\kappa^{k_\kappa}} \|u\|_{H^{k_\kappa}(\kappa)} \ , \tag{4.30}$$

for $1 \leq s_\kappa \leq \min(p_\kappa + 1, k_\kappa)$; we note that in the transition from (4.29) to (4.30), the generic constant C is increased by the factor $(k_\kappa - 1)^{k_\kappa - 1}$. Substituting (4.30) into (4.27) and (4.28), collecting the resulting bounds on the various norms of η_2 and the corresponding bounds on η_1, we deduce from (4.26) that

$$\|\sqrt{a}\,\nabla\eta\|^2_{L^2(\kappa)} \leq C\bar{a}_\kappa \frac{h_\kappa^{2s_\kappa-2}}{p_\kappa^{2k_\kappa-4}} \|u\|^2_{H^{k_\kappa}(\kappa)}\,,$$

$$\|\eta\|^2_{L^2(\partial\kappa)} \leq C\frac{h_\kappa^{2s_\kappa-1}}{p_\kappa^{2k_\kappa-2}} \|u\|^2_{H^{k_\kappa}(\kappa)}\,, \quad \|\nabla\eta\|^2_{L^2(\partial\kappa)} \leq C\frac{h_\kappa^{2s_\kappa-3}}{p_\kappa^{2k_\kappa-6}} \|u\|^2_{H^{k_\kappa}(\kappa)}\,.$$

Inserting these bounds on η into (4.8) and noting the definition (4.23) of σ_e, (4.17) and the shape regularity of $\mathcal{T}$ to relate h_e to h_κ, we get

$$|||\xi|||^2_{\mathrm{DG}} \leq C \sum_{\kappa\in\mathcal{T}} \alpha_\kappa \frac{h_\kappa^{2s_\kappa-2}}{p_\kappa^{2k_\kappa-4}} \|u\|^2_{H^{k_\kappa}(\kappa)}\,,$$

for $1 \leq s_\kappa \leq \min(p_\kappa + 1, k_\kappa)$, $\kappa \in \mathcal{T}$. An identical bound holds for $|||\eta|||^2_{\mathrm{DG}}$. The estimate (4.24) then follows from (4.25) via the triangle inequality. ∎

5 Partial Differential Equations with nonnegative characteristic form

Let us return to the general problem (2.1), (2.5). The associated *hp*–DGFEM is now defined as follows: find $u_{\mathrm{DG}} \in S^{\mathbf{p}}(\Omega, \mathcal{T}, \mathbf{F})$ such that

$$B_{\mathrm{DG}}(u_{\mathrm{DG}}, v) = \ell_{\mathrm{DG}}(v) \quad \forall v \in S^{\mathbf{p}}(\Omega, \mathcal{T}, \mathbf{F})\,, \tag{5.1}$$

where

$$\begin{aligned} B_{\mathrm{DG}}(w,v) = \sum_{\kappa\in\mathcal{T}} \Bigg(& \int_\kappa a\nabla w\cdot\nabla v\mathrm{d}x + \int_\kappa (b\cdot\nabla w + cw)v\mathrm{d}x \\ & - \int_{\partial_-\kappa\cap(\Gamma_{\mathrm{D}}\cup\Gamma_-)} (b\cdot\mu)\, w^+ v^+\mathrm{d}s - \int_{\partial_-\kappa\setminus\Gamma} (b\cdot\mu)\lfloor w\rfloor\, v^+\mathrm{d}s\Bigg) \\ & + \int_{\Gamma_{\mathrm{D}}} \{w((a\nabla v)\cdot\mu) - ((a\nabla w)\cdot\mu)v\}\mathrm{d}s + \int_{\Gamma_{\mathrm{D}}} \sigma wv\mathrm{d}s \\ & + \int_{\Gamma_{\mathrm{int}}} \{[w]\langle(a\nabla v)\cdot\nu\rangle - \langle(a\nabla w)\cdot\nu\rangle[v]\}\mathrm{d}s + \int_{\Gamma_{\mathrm{int}}} \sigma[w][v]\mathrm{d}s\,, \end{aligned}$$

and

$$\begin{aligned} \ell_{\mathrm{DG}}(v) = \sum_{\kappa\in\mathcal{T}} & \left(\int_\kappa fv\mathrm{d}x - \int_{\partial_-\kappa\cap(\Gamma_{\mathrm{D}}\cup\Gamma_-)} (b\cdot\mu)\, g_{\mathrm{D}}\, v^+\mathrm{d}s\right) \\ & + \int_{\Gamma_{\mathrm{D}}} g_{\mathrm{D}}((a\nabla v)\cdot\mu)\mathrm{d}s + \int_{\Gamma_{\mathrm{N}}} g_{\mathrm{N}}v\mathrm{d}s + \int_{\Gamma_{\mathrm{D}}} \sigma g_{\mathrm{D}}v\mathrm{d}s\,, \end{aligned}$$

with σ a positive parameter whose precise choice will be given in the next theorem. Still assuming (3.2) and with c_0 defined by (3.3), we introduce the DG–norm

$$|||w|||_{\mathrm{DG}}^2 = \sum_{\kappa\in\mathcal{T}} \left(\|\sqrt{a}\,\nabla w\|_{L^2(\kappa)}^2 + \|c_0 w\|_{L^2(\kappa)}^2 + \frac{1}{2}\|w^+\|_{\partial_-\kappa\cap(\Gamma_{\mathrm{D}}\cup\Gamma_-)}^2 \right.$$
$$\left. + \frac{1}{2}\|w^+ - w^-\|_{\partial_-\kappa\setminus\Gamma}^2 + \frac{1}{2}\|w^+\|_{\partial_+\kappa\cap\Gamma}^2 \right) + \int_{\Gamma_{\mathrm{D}}} \sigma w^2 \mathrm{d}s + \int_{\Gamma_{\mathrm{int}}} \sigma [w]^2 \mathrm{d}s \ .$$

The discontinuous Galerkin finite element methods considered in Sections 3 and 4 for the purely hyperbolic and the elliptic, purely diffusive, problems, respectively, are special cases of (5.1); the same is true of the norms $|||\cdot|||_{\mathrm{DG}}$ associated with the bilinear forms of those methods. Now consider $B_{\mathrm{DG}}(w,w)$. On writing $(b\cdot\nabla w)w = \frac{1}{2} b\cdot\nabla w^2$ for $w\in H^2(\Omega,\mathcal{T})$, after integration by parts and recalling that by hypothesis $b\cdot\mu \geq 0$ on Γ_{N} and therefore $|\partial_-\kappa\cap(\Gamma_{\mathrm{N}}\cup\Gamma_+)| = 0$ for each $\kappa\in\mathcal{T}$, we have

$$\sum_{\kappa\in\mathcal{T}} \left(\int_\kappa (b\cdot\nabla w + cw) w \mathrm{d}x - \int_{\partial_-\kappa\cap(\Gamma_{\mathrm{D}}\cup\Gamma_-)} (b\cdot\mu)\,|w^+|^2 \mathrm{d}s - \int_{\partial_-\kappa\setminus\Gamma} (b\cdot\mu) \lfloor w \rfloor\, w^+ \mathrm{d}s \right)$$
$$= \sum_{\kappa\in\mathcal{T}} \left(\|c_0 w\|_{L^2(\kappa)}^2 + \frac{1}{2}\|w^+\|_{\partial_-\kappa\cap(\Gamma_{\mathrm{D}}\cup\Gamma_-)}^2 + \frac{1}{2}\|w^+ - w^-\|_{\partial_-\kappa\setminus\Gamma}^2 + \frac{1}{2}\|w^+\|_{\partial_+\kappa\cap\Gamma}^2 \right).$$

Hence, trivially,

$$|||w|||_{\mathrm{DG}}^2 = B_{\mathrm{DG}}(w,w) \qquad \forall w\in H^2(\Omega,\mathcal{T}) \ .$$

In order to ensure that the Galerkin orthogonality property $B_{\mathrm{DG}}(u-u_{\mathrm{DG}},v)=0$ holds for all $v\in S^{\mathbf{p}}(\Omega,\mathcal{T},\mathbf{F})$, we suppose that the solution u to the boundary value problem under consideration is sufficiently smooth: namely, $u\in H^2(\Omega,\mathcal{T})$ and the functions u and $(a\nabla u)\cdot\nu$ are continuous across each face e in $\mathcal{E}_{\mathrm{int}}$ that intersects the subdomain of ellipticity, $\{x\in\bar{\Omega} \,:\, \zeta^{\mathrm{T}} a(x)\zeta > 0 \ \ \forall \zeta\in\mathbb{R}^d\}$. If this smoothness requirement is violated, the discretization method has to be modified accordingly (cf. Sec. 6.4 for an example). Thereby, combining the error bounds from Theorems 9 and 14, we arrive at the following *a priori* error estimate for the *hp*–DGFEM (5.1).

Theorem 15 *Let $\Omega\subset\mathbb{R}^d$ be a bounded polyhedral domain, $\mathcal{T}=\{\kappa\}$ a shape–regular subdivision of Ω into d-parallelepipeds, and $\mathbf{p}$ a polynomial degree vector of bounded local variation. Suppose that on face $e\in\mathcal{E}_{\mathrm{int}}\cup\mathcal{E}_{\mathrm{D}}$ the parameter σ_e is defined as in (4.23). Then, assuming that the conditions (3.2), (3.6) and (4.3) on the data hold, and $u|_\kappa\in H^{k_\kappa}(\kappa)$, $k_\kappa\geq 2$, for $\kappa\in\mathcal{T}$, the solution $u_{\mathrm{DG}}\in S^{\mathbf{p}}(\Omega,\mathcal{T},\mathbf{F})$ of (5.1) obeys the error bound*

$$|||u-u_{\mathrm{DG}}|||_{\mathrm{DG}}^2 \leq C \sum_{\kappa\in\mathcal{T}} \left(\alpha_\kappa \frac{h_\kappa^{2s_\kappa-2}}{p_\kappa^{2k_\kappa-4}} + \beta_\kappa \frac{h_\kappa^{2s_\kappa-1}}{p_\kappa^{2k_\kappa-1}} + \gamma_\kappa \frac{h_\kappa^{2s_\kappa}}{p_\kappa^{2k_\kappa}} \right) \|u\|_{H^{k_\kappa}(\kappa)}^2 \ ,$$

for $1\leq s_\kappa\leq \min(p_\kappa+1,k_\kappa)$, $p_\kappa\geq 1$, $\kappa\in\mathcal{T}$, where $\alpha_\kappa=\bar{a}_{\tilde{\kappa}}$; β_κ and γ_κ are as in Theorem 9 and C is a constant depending on the dimension d, the parameter ρ from (4.17), $k=\max_\kappa k_\kappa$, and the shape–regularity of $\mathcal{T}$.

We highlight the fact that since the discontinuity–penalization σ involves the norm of the matrix $\sqrt{a}$, in the hyperbolic limit of $a \equiv 0$ the terms that contain σ in $B_{\rm DG}(\cdot,\cdot)$ and $\ell_{\rm DG}$ all vanish. This is a desirable property, since linear hyperbolic equations may possess solutions that are discontinuous across characteristic hypersurfaces, and penalizing discontinuities across faces which belong to these seems unnatural.

A further bound on $u - u_{\rm DG}$ can be obtained from the decomposition

$$\begin{aligned} u - u_{\rm DG} &= (u - \Pi_p u) + (\Pi_p u - u_{\rm DG}) \equiv \eta + \xi \\ &= (u - \tilde{\Pi}_p^h u) + \tilde{\Pi}_p^h (u - \Pi_p u) + (\Pi_p u - u_{\rm DG}) \equiv \eta_1 + \eta_2 + \xi \ , \end{aligned}$$

where $\tilde{\Pi}_p^h$ is the projector from (4.5.20) in [29]. Suppose further that u is *element-wise analytic* on $\mathcal{T}$ in the sense that, for each $\kappa \in \mathcal{T}$, $u|_\kappa$ has analytic extension to an open set, independent of h_κ, containing $\bar{\kappa}$. Then,

$$\exists d_\kappa > 0 \quad \exists C = C(u) > 0 \quad \forall s \geq 0 \qquad |u|_{H^s(\kappa)} \leq C(u) (d_\kappa)^s \, s! \, |{\rm meas}(\kappa)|^{\frac{1}{2}} \ . \tag{5.2}$$

Since $\tilde{\Pi}_p^h$ obeys bounds similar to (3.22) and (3.25) with constants whose dependence on the Sobolev index is given explicitly in terms of the Gamma–function, by means of Stirling's formula, after a rather straightforward but lengthy calculation, as in [15], we deduce the following result.

Theorem 16 *Let $\Omega \subset \mathbb{R}^d$ be a bounded polyhedral domain, $\mathcal{T} = \{\kappa\}$ a shape–regular subdivision of Ω into d-parallelepipeds, and $\mathbf{p}$ a polynomial degree vector of bounded local variation. Suppose that on face $e \in \mathcal{E}_{\rm int} \cup \mathcal{E}_{\rm D}$ the parameter σ_e is defined as in (4.23), u is element-wise analytic on $\mathcal{T}$, and (3.2), (3.6) and (4.3) hold. Then, the solution $u_{\rm DG} \in S^{\mathbf{p}}(\Omega, \mathcal{T}, \mathbf{F})$ of (5.1), with $p_\kappa \geq 1$ for $\kappa \in \mathcal{T}$, obeys the error bound*

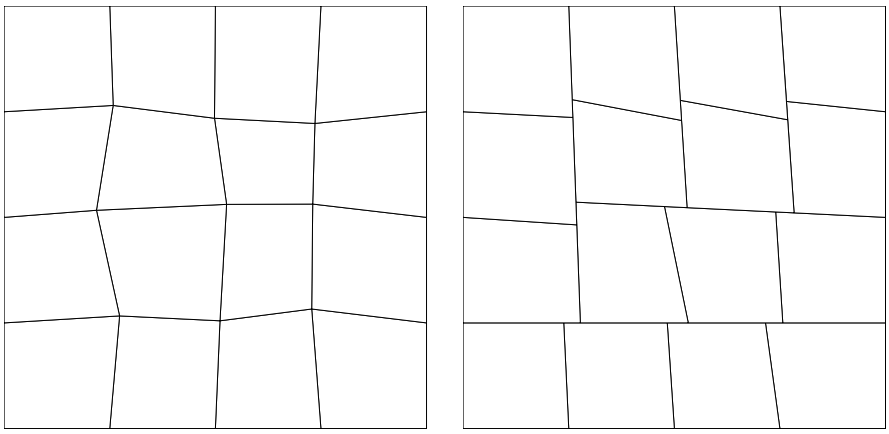
$$|||u - u_{\rm DG}|||_{\rm DG}^2 \leq C(u) \sum_{\kappa \in \mathcal{T}} \left(\alpha_\kappa h_\kappa^{-2} + \beta_\kappa h_\kappa^{-1} + \gamma_\kappa \right) {\rm e}^{-2 p_\kappa (\chi_\kappa + \varepsilon_\kappa |\ln h_\kappa|)} |{\rm meas}(\kappa)| \ ,$$

where $C(u)$ is a constant depending on u, the dimension d, the parameter ρ from (4.17), and the shape–regularity of $\mathcal{T}$; $\alpha_\kappa = \bar{a}_{\tilde{\kappa}}$, β_κ and γ_κ are as in Theorem 9, $\varepsilon_\kappa = (1 + d_\kappa^2)^{-1/2}$ with d_κ as in (5.2), and $\chi_\kappa = -\frac{1}{4} \ln F(\varepsilon_\kappa, d_\kappa) > 0$, where $F(\varepsilon, d) = (\varepsilon d)^{2\varepsilon} (1 - \varepsilon)^{1-\varepsilon} (1 + \varepsilon)^{-1-\varepsilon}$.

Consequently, if the solution u is element-wise analytic on $\mathcal{T}$ then the DGFEM exhibits an exponential rate of convergence as $p_\kappa \to \infty$, and if p_κ is sufficiently large an increase in the exponential rate of convergence occurs as h_κ is reduced.

6 Numerical experiments

In this section we present a number of numerical experiments to illustrate the *a priori* error estimates derived for the *hp*–DGFEM. We begin, in Section 6.1, by considering a strictly hyperbolic problem ($\{a_{ij}\}_{i,j=1}^d = 0$); in Section 6.2 we study a self–adjoint elliptic problem ($b = 0$); in Section 6.3 we look at a singularly–perturbed isotropic advection–diffusion problem ($a = \varepsilon I$, $0 < \varepsilon \ll 1$); and finally in Section 6.4 we consider an advection–diffusion problem with degenerate, anisotropic diffusion matrix a. In all cases we investigate the performance of the DGFEM in two space–dimensions ($d = 2$) on quadrilateral meshes.



(a) (b)

Figure 2: Example 1. (a) Quadrilateral mesh (i); (b) Quadrilateral mesh (ii).

6.1 Example 1

In this example we let $\Omega = (-1,1)^2$, and select

$$\{a_{ij}\}_{i,j=1}^{2} = 0\,, \qquad b = (2-y^2, 2-x)\,, \qquad c = 1 + (1+x)(1+y)^2\,;$$

the forcing function f is chosen so that the analytical solution to (2.1) is given by

$$u(x,y) = 1 + \sin(\pi(1+x)(1+y)^2/8)\,. \tag{6.1}$$

This is a variant of the hyperbolic test problem considered in [4, 15].

We investigate the asymptotic behavior of the *hp*–DGFEM on a sequence of successively finer square and quadrilateral meshes for different values of the polynomial degree p. In each case we consider two types of quadrilateral mesh which are constructed from a uniform $N \times N$ square mesh by (i) randomly perturbing each of the interior nodes by up to 10% of the local mesh size, cf. Figure 2(a); (ii) randomly *splitting* each of the interior nodes by a displacement of up to 10% of the local mesh size, cf. Figure 2(b). We note that the latter meshes are constructed so that all the nodes in the interior of Ω are irregular (i.e., hanging).

In Figure 3 we plot the DG–norm of the error against the mesh function h for p between 1 and 5. For consistency, $|||u - u_{\mathrm{DG}}|||_{\mathrm{DG}}$ is plotted against h_u for each mesh type, where h_u denotes the mesh-size of the uniform $N \times N$ square mesh; this ensures that a fair comparison between the error per degree of freedom for each mesh type can be made. We see that $|||u - u_{\mathrm{DG}}|||_{\mathrm{DG}}$ converges at the rate $\mathcal{O}(h^{p+\frac{1}{2}})$ as h tends to zero for each (fixed) p, thereby confirming Theorem 15 (see also Theorem 9) in the case of a variable velocity vector b that *does not* satisfy the condition (3.6).

In particular, we observe that while the error on the square mesh is smaller than on the randomly generated quadrilateral mesh (i), as we would expect; the error is consistently

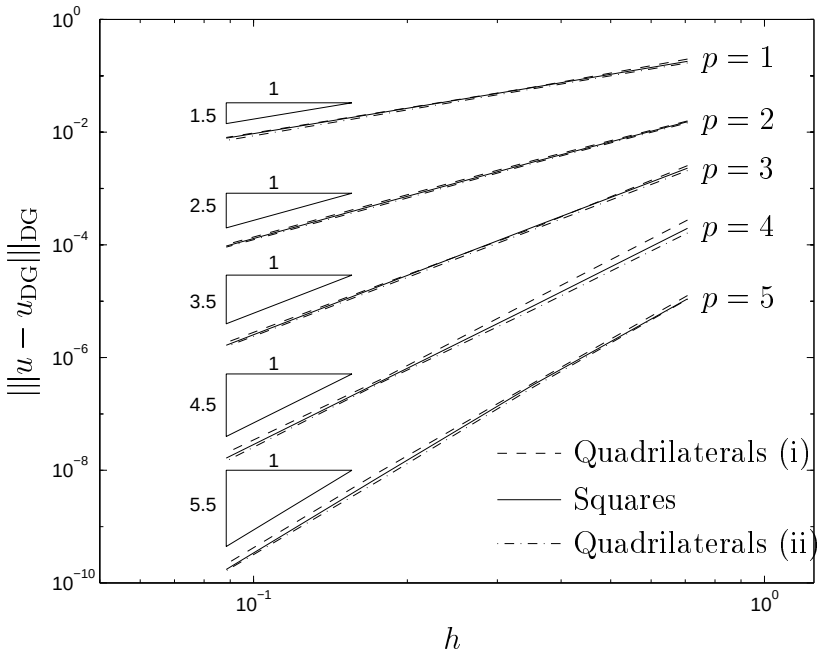


Figure 3: Example 1. Convergence of the DGFEM with h–refinement.

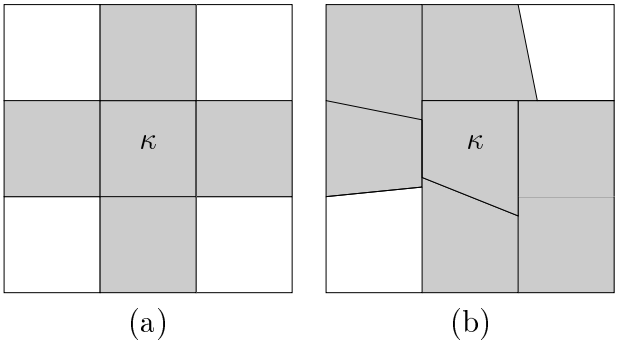


Figure 4: Inter-element communication: (a) Regular mesh; (b) Irregular mesh with hanging nodes.

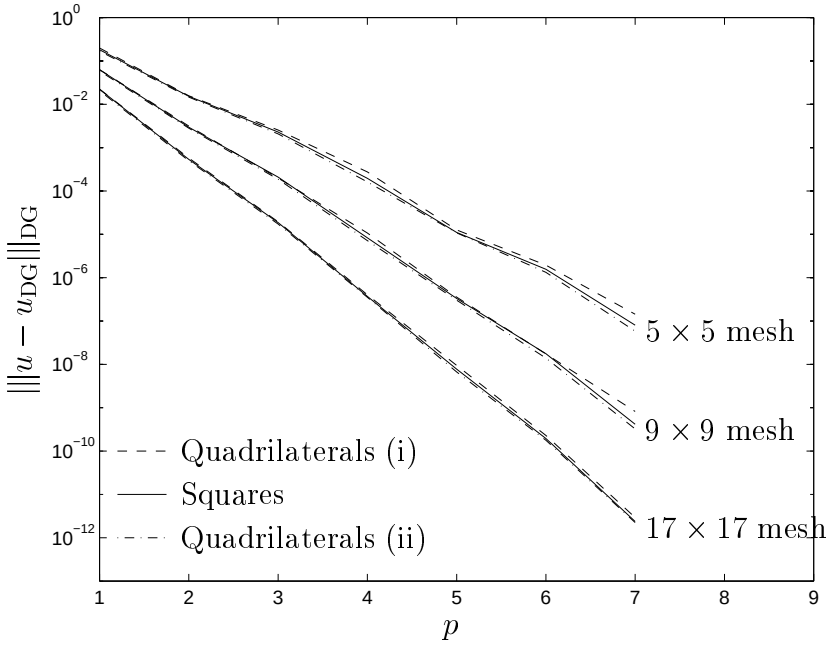


Figure 5: Example 1. Convergence of the DGFEM under p–refinement.

smaller when the irregular quadrilateral mesh is employed. We attribute this improvement in $|||u - u_{DG}|||_{DG}$ to the increase in inter-element communication on the meshes (ii); when no hanging nodes are present in the mesh, elements may only communicate with their four immediate neighbors, cf. Figure 4(a). On the other hand, on irregular meshes elements may now communicate with all of their neighbors which share a common node, cf. Figure 4(b). However, we note that the improvement in the error when the mesh consists of irregular hanging nodes is relatively small; moreover, the increased inter-element communication leads to an increase in the number of nonzero entries arising in the matrix system, which obviously increases the storage requirement for the method.

Next, we investigate the convergence of the DGFEM under p–refinement for fixed h. In Figure 5 we first plot the DG–norm of the error against p on three different square and quadrilateral meshes (meshes (i) and (ii)). In each case, we observe that on the linear–log scale, the convergence plots become straight lines as the degree of the approximating polynomial is increased, thereby indicating exponential convergence in p, cf. Theorem 16. Furthermore, we observe that the p–convergence of the DGFEM is robust with respect to mesh distortion.

Finally, to ensure that the DGFEM converges at (least at) the optimal algebraic rate predicted by Remark 2 as the polynomial degree p is increased, even when b *does not* satisfy condition (3.6), we now consider a slightly different test problem for which the precise regularity of the analytical solution u is known. To this end, we keep the functions

Table 1: Example 1. Convergence of the DGFEM under p–refinement on a 6×6 uniform square mesh (singularity lies in the interior of a strip of elements).

p	$\alpha = 3/2$		$\alpha = 5/2$		$\alpha = 7/2$	
	$\|\|\|u - u_{\mathrm{DG}}\|\|\|_{\mathrm{DG}}$	k	$\|\|\|u - u_{\mathrm{DG}}\|\|\|_{\mathrm{DG}}$	k	$\|\|\|u - u_{\mathrm{DG}}\|\|\|_{\mathrm{DG}}$	k
1	0.1707	-	0.2038	-	0.2718	-
2	0.2034E-01	3.07	0.1408E-01	3.86	0.2246E-01	3.60
3	0.8171E-02	2.25	0.1598E-02	5.37	0.1060E-02	7.53
4	0.4807E-02	1.84	0.6122E-03	3.34	0.1435E-03	6.95
5	0.3186E-02	1.84	0.3117E-03	3.02	0.4983E-04	4.74
6	0.2267E-02	1.87	0.1820E-03	2.95	0.2267E-04	4.32
7	0.1700E-02	1.87	0.1158E-03	2.93	0.1196E-04	4.15
8	0.1320E-02	1.89	0.7847E-04	2.91	0.6945E-05	4.07
9	0.1057E-02	1.89	0.5565E-04	2.92	0.4327E-05	4.02
10	0.8640E-03	1.91	0.4096E-04	2.91	0.2841E-05	3.99
11	0.7209E-03	1.90	0.3102E-04	2.92	0.1946E-05	3.97
12	0.6095E-03	1.93	0.2408E-04	2.91	0.1379E-05	3.96
13	0.5231E-03	1.91	0.1906E-04	2.92	0.1005E-05	3.95
14	0.4530E-03	1.94	0.1535E-04	2.92	0.7499E-06	3.95

Table 2: Example 1. Convergence of the DGFEM under p–refinement on a 5×5 uniform square mesh (singularity lies on inter-element boundaries).

p	$\alpha = 3/2$		$\alpha = 5/2$		$\alpha = 7/2$	
	$\|\|\|u - u_{\mathrm{DG}}\|\|\|_{\mathrm{DG}}$	k	$\|\|\|u - u_{\mathrm{DG}}\|\|\|_{\mathrm{DG}}$	k	$\|\|\|u - u_{\mathrm{DG}}\|\|\|_{\mathrm{DG}}$	k
1	0.2313	-	0.2876	-	0.3762	-
2	0.2295E-01	3.33	0.2251E-01	3.68	0.3971E-01	3.24
3	0.4423E-02	4.06	0.1363E-02	6.92	0.1779E-02	7.66
4	0.1963E-02	2.82	0.2326E-03	6.15	0.1011E-03	9.97
5	0.1053E-02	2.79	0.7593E-04	5.02	0.1573E-04	8.34
6	0.6286E-03	2.83	0.3125E-04	4.87	0.4118E-05	7.35
7	0.4036E-03	2.88	0.1480E-04	4.85	0.1372E-05	7.13
8	0.2731E-03	2.92	0.7743E-05	4.85	0.5370E-06	7.03
9	0.1922E-03	2.98	0.4367E-05	4.86	0.2362E-06	6.97
10	0.1392E-03	3.06	0.2611E-05	4.88	0.1137E-06	6.94
11	0.1031E-03	3.16	0.1634E-05	4.92	0.5873E-07	6.93
12	0.7746E-04	3.28	0.1059E-05	4.98	0.3213E-07	6.93
13	0.5875E-04	3.45	0.7058E-06	5.07	0.1841E-07	6.96
14	0.4470E-04	3.69	0.4789E-06	5.23	0.1093E-07	7.03

a, b and c as above, and choose the forcing function f so that

$$u(x,y) = \begin{cases} \cos(\pi y/2) & \text{in } (-1,0)\times(-1,1)\,, \\ \cos(\pi y/2) + x^{\alpha} & \text{in } (0,1)\times(-1,1)\,, \end{cases}$$

where α is a nonnegative constant. The solution u belongs to $H^{\alpha+\frac{1}{2}-\varepsilon}(\Omega)$, for any $\varepsilon > 0$, but does not belong to $H^{\alpha+\frac{1}{2}}(\Omega)$; cf. Castillo *et al.* [6].

In Tables 1 & 2 we show the DG–norm of the error and the convergence rate k as p is increased for $\alpha = 3/2$, $5/2$ and $7/2$, on uniform $N \times N$ square meshes, with $N = 6$ and $N = 5$, respectively. For N even (cf. Table 1 for $N = 6$), the singularity lies in the interior of the strip of elements in the mesh which intersect the line $x = 0$. In this case, $|||u - u_{\rm DG}|||_{\rm DG}$ converges at the rate (approximately) $\mathcal{O}(p^{-(\alpha+\frac{1}{2})})$ as p tends to infinity for fixed h.

In contrast, for N odd (cf. Table 2 for $N = 5$), the singularity lies on inter-element boundaries; here, the DG–norm of the error behaves (on average) like $\mathcal{O}(p^{-2\alpha})$ as p tends to infinity for fixed h. Thus, in each case we observe an improvement of approximately $p^{-1/2}$ over the theoretical predictions in Remark 1. More importantly, the results show that the optimal rate of p-convergence predicted by our theory when b satisfies the condition (3.6) is retained by the method even if (3.6) is violated.

6.2 Example 2

In this second example, we let $\Omega = (0,1)^2$,

$$a = \begin{pmatrix} \varepsilon + x & xy \\ xy & \varepsilon + y \end{pmatrix},$$

where $\varepsilon = 1/10$, $b = (0,0)$, $c = \mathrm{e}^{-(x^2+y^2)}$ and f is chosen so that

$$u(x,y) = \frac{(1+x)^2}{4}\sin(2\pi xy)\,.$$

Here, we study the convergence rate of the hp–DGFEM in the presence of the discontinuity–penalization term $\sigma \equiv \sigma(m)$; we recall that

$$\sigma|_e \equiv \sigma_e(m) = \frac{\langle \bar{a} p^m \rangle_e}{h_e} \qquad \text{for } \; e \in \mathcal{E}_{\rm int} \cup \mathcal{E}_{\rm D}\,,$$

where $m = 1$ or $m = 2$, cf. (4.18) and (4.23), respectively. Figure 6 presents a comparison of the error in the DGFEM with the mesh function h for $1 \le p \le 5$ on uniform square meshes. For consistency, here the error is measured in the (broken) energy norm, $\|\cdot\|_E$, rather than the DG–norm, since the definition of the latter norm is dependent on the presence and the size of the discontinuity–penalization in the scheme. We observe that the error behaves like $\mathcal{O}(h^p)$ which is the h–optimal rate of convergence, since for this uniformly elliptic problem the energy norm is equivalent to the broken H^1-norm. In particular, we

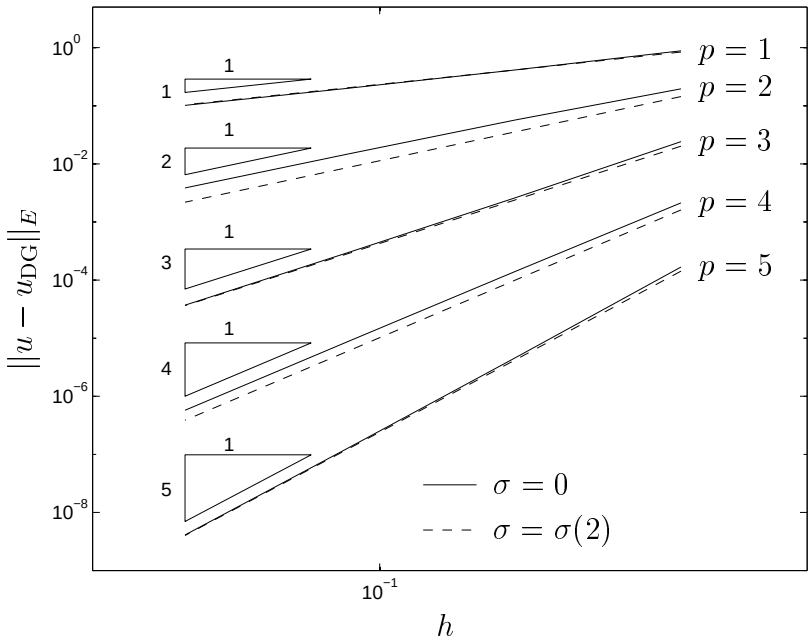


Figure 6: Example 2. Convergence of the DGFEM in the energy norm under h–refinement.

note that, for each p, the presence of $\sigma(m)$, with $m = 2$, reduces the error in the DGFEM in the broken energy norm; this is much more evident for even p than for odd p. For the purposes of clarity, the numerical results for $m = 1$ have been omitted; in this case the presence of the discontinuity–penalization term $\sigma(1)$ still improves the error in the DGFEM compared with $\sigma = 0$, though by a smaller factor than when $m = 2$.

$$\sigma|_e \equiv \sigma_e(m) = \frac{\langle \bar{a} p^m \rangle_e}{h_e} \quad \text{for} \quad e \in \mathcal{E}_{\mathrm{int}} \cup \mathcal{E}_{\mathrm{D}} \ ,$$

where $m = 1$ or $m = 2$, cf. (4.18) and (4.23), respectively. Figure 6 presents a comparison of the error in the DGFEM with the mesh function h for $1 \leq p \leq 5$ on uniform square meshes. For consistency, here the error is measured in the (broken) energy norm, $\|\cdot\|_E$, rather than the DG–norm, since the definition of the latter norm is dependent on the presence and the size of the discontinuity–penalization in the scheme. We observe that the error behaves like $\mathcal{O}(h^p)$ which is the h–optimal rate of convergence, since for this uniformly elliptic problem the energy norm is equivalent to the broken H^1-norm. In particular, we note that, for each p, the presence of $\sigma(m)$, with $m = 2$, reduces the error in the DGFEM in the broken energy norm; this is much more evident for even p than for odd p. For the purposes of clarity, the numerical results for $m = 1$ have been omitted; in this case the presence of the discontinuity–penalization term $\sigma(1)$ still improves the error

In Figure 7 we now only plot the (c_0-weighted) L^2-norm of the error against h; we observe that the error in the L^2-norm behaves like $\mathcal{O}(h^{p+1})$ for odd p and like $\mathcal{O}(h^p)$ for

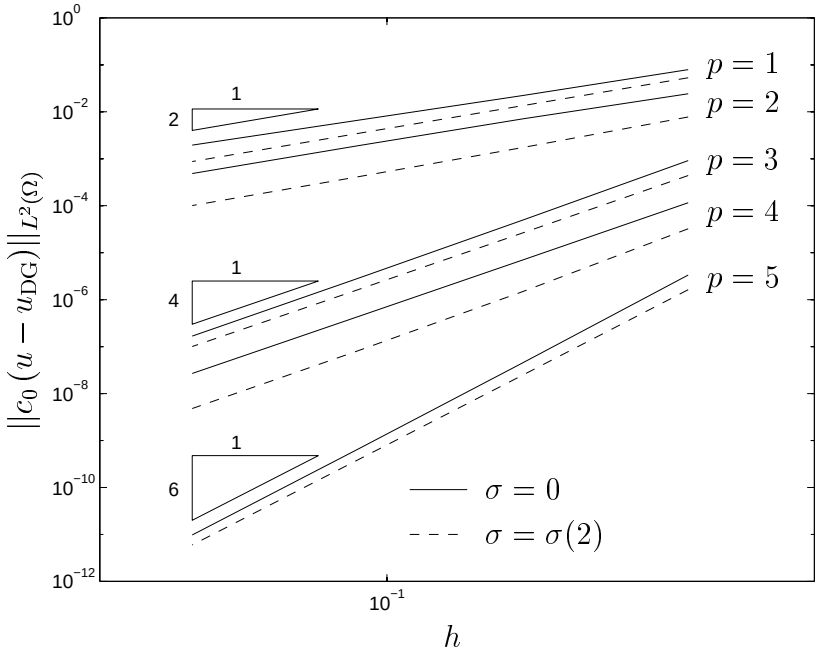


Figure 7: Example 2. Convergence of the DGFEM in the (c_0-weighted) $L^2(\Omega)$ norm under h–refinement.

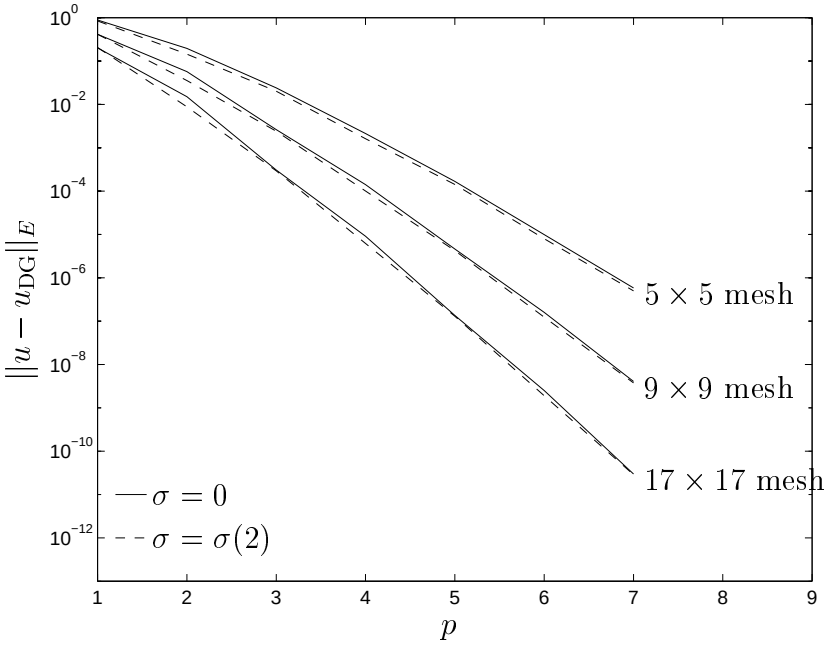


Figure 8: Example 2. Convergence of the DGFEM under p–refinement.

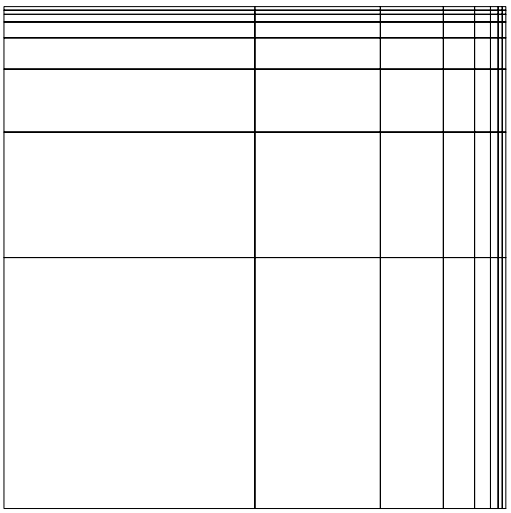


Figure 9: Example 3. Geometrically refined mesh for $n_\varepsilon = 9$.

even p, cf. [24]. Thus, for second–order elliptic problems the DGFEM is numerically observed to be h–optimal in the energy norm for all p; but h–optimality in the L^2-norm is only seen for p odd. In each case the presence of the discontinuity–penalization with $m = 2$ reduces the error of the DGFEM in the L^2-norm, though the improvement is smaller when p is odd.

Finally, in Figure 8 we investigate the convergence of the DGFEM under p–refinement for fixed h. On a linear–log scale, the figure shows the robust exponential convergence of the method on uniform square meshes for both $\sigma = \sigma(2)$ and $\sigma = 0$. We note that identical behavior is observed on quasi-uniform quadrilateral meshes with both h–refinement and p–refinement, and with $\sigma = \sigma(1)$.

6.3 Example 3

We consider the following singularly perturbed advection–diffusion problem:

$$-\varepsilon\Delta u + u_x + u_y \;=\; f\,, \quad (x,y) \in (0,1)^2\,,$$

where, $0 < \varepsilon \ll 1$ and f is chosen so that

$$u(x,y) = x + y(1-x) + \frac{\mathrm{e}^{-1/\varepsilon} - \mathrm{e}^{-(1-x)(1-y)/\varepsilon}}{1 - \mathrm{e}^{-1/\varepsilon}}\,. \tag{6.2}$$

This is a multi-dimensional variant of the 1D advection–diffusion problem considered by Melenk & Schwab [21]. For $0 < \varepsilon \ll 1$ the solution (6.2) has boundary layers along $x = 1$ and $y = 1$.

In this numerical experiment we test the robustness of the hp–DGFEM on highly stretched anisotropic quadrilateral meshes as the physical diffusion ε decreases; in each

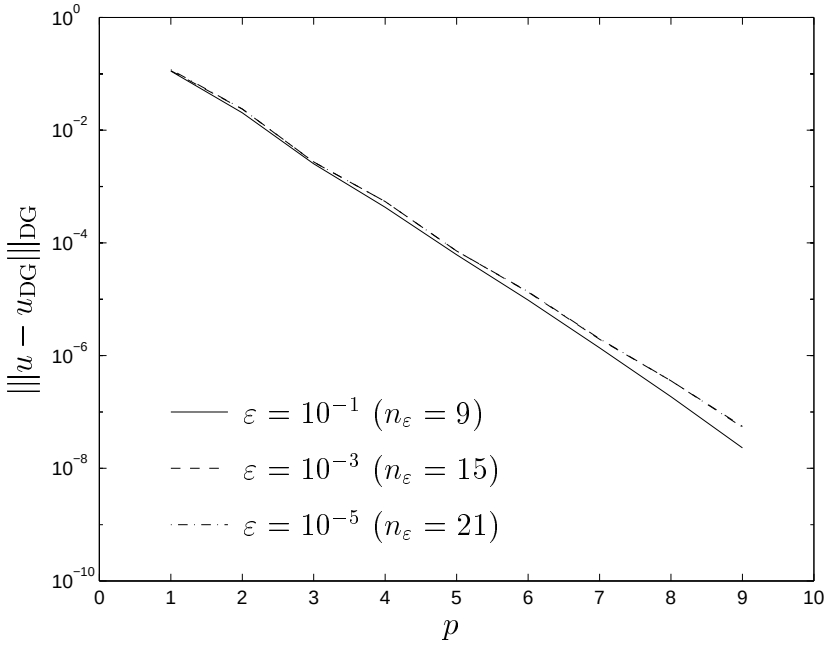


Figure 10: Example 3. Convergence of the DGFEM with $\sigma = \sigma(2)$ under p–refinement for $\varepsilon = 10^{-1}, 10^{-3}, 10^{-5}$.

case the discontinuity–penalization $\sigma = \sigma(2)$. The meshes are constructed by geometrical refinement into the boundary layers along $x = 1$ and $y = 1$, and are parameterized by n_ε which denotes the number of points in the x and y directions. In Figure 9 we show a typical mesh for $n_\varepsilon = 9$. Figure 10 shows a plot of the DG–norm of the error against the polynomial degree p on a linear–log scale for $\varepsilon = 10^{-1}, 10^{-3}, 10^{-5}$, on geometrically refined quadrilateral meshes with $n_\varepsilon = 9, 15, 21$, respectively. In each case we observe robust exponential convergence as the polynomial degree is increased; we note that the largest cell–aspect ratio in each of the three meshes used is 64 for $n_\varepsilon = 9$, 4096 for $n_\varepsilon = 15$ and over a quarter of a million for $n_\varepsilon = 21$.

6.4 Example 4

In this final example we consider a partial differential equation with nonnegative characteristic form which has mixed type on the domain Ω. Let $\Omega = (-1, 1)^2$, suppose that ε is a positive constant and c_1 and c_2 are nonnegative constants. We consider the following model problem with type-change across the horizontal line $y = 0$:

$$\begin{cases} -\varepsilon u_{yy} + u_x + c_1 u &= 0 \quad -1 \le x \le 1\,, \quad y > 0\,, \\ u_x + c_2 u &= 0 \quad -1 \le x \le 1\,, \quad y \le 0\,. \end{cases} \tag{6.3}$$

The problem is parabolic for $y > 0$ and hyperbolic for $y \le 0$.

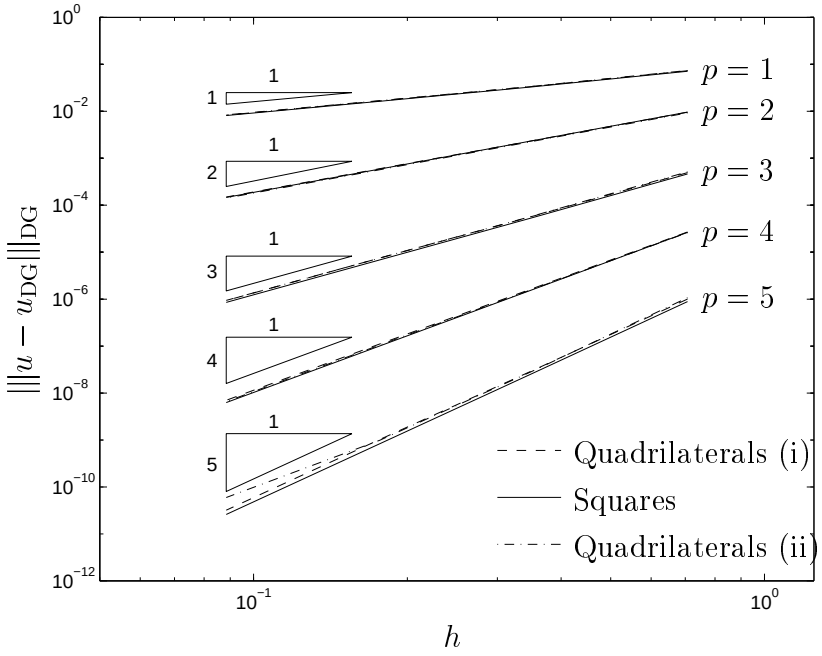


Figure 11: Example 4. Convergence of the DGFEM with $\sigma = \sigma(2)$ under h–refinement.

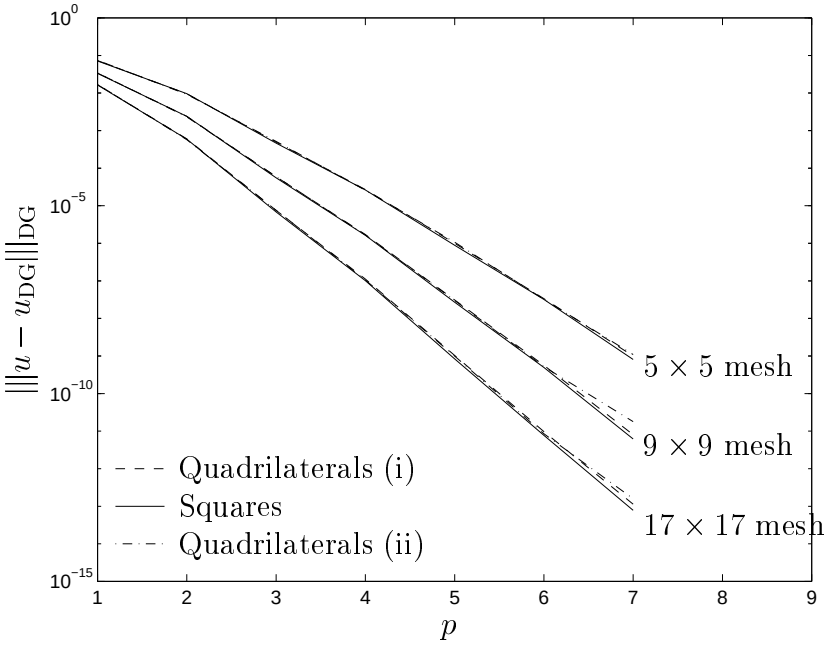


Figure 12: Example 4. Convergence of the DGFEM with $\sigma = \sigma(2)$ under p–refinement.

In order to ensure continuity of the normal flux across the interface where the partial differential equation changes type from parabolic to hyperbolic, the analytical solution is (in general) forced to be discontinuous across $y = 0$. Thereby, for appropriate boundary data, the solution to (6.3) is given by

$$u(x,y) = \begin{cases} \sin(\pi(1+y)/2)\,\mathrm{e}^{-(c_1+\varepsilon\pi^2/4)(1+x)} & -1 \le x \le 1\,, \quad y > 0\,, \\ \sin(\pi(1+y)/2)\,\mathrm{e}^{-c_2(1+x)} & -1 \le x \le 1\,, \quad y \le 0\,. \end{cases} \tag{6.4}$$

In the special case when $c_1 + \varepsilon\pi^2/4 = c_2$, the solution (6.4) is in fact continuous. The DG–norm of the error then, again, converges optimally as h tends to zero for a fixed polynomial degree p, and exponentially as the polynomial degree is increased for fixed h, as we would expect. For brevity, these results are omitted.

A more interesting situation occurs when the solution u is discontinuous; thereby the smoothness assumptions required in Section 5 for the proof of Theorem 15 are violated. In this case, the numerical scheme (5.1) must be modified to ensure that the discontinuity–penalization term σ is *inactive* on edges along $y = 0$ across which the underlying partial differential equation changes type. To demonstrate the performance of the *hp*–DGFEM for (6.3), we set $\varepsilon = 5 \times 10^{-2}$, $c_1 = c_2 = 1/10$ (i.e., $c_1 + \varepsilon\pi^2/4 \neq c_2$) and $\sigma = \sigma(2)$.

To highlight one of the advantages of using discontinuous elements, we consider uniform $N \times N$ (N odd) square meshes and quasi-uniform quadrilateral meshes for which the discontinuity in the analytical solution lies on element interfaces only; here the quadrilateral meshes are constructed as in Examples 1 and 2, except that the internal nodes in the mesh on the line $y = 0$ are kept fixed. In this case the *hp*–DGFEM behaves as if the analytical solution were smooth; i.e., optimal algebraic rates of convergence are observed under *h*–refinement, cf. Figure 11, and exponential rates of convergence are observed under *p*–refinement, cf. Figure 12. In fact, since the analytic extensions of the two pieces of the solution (above and below the line $y = 0$) are entire analytic functions, the method exhibits super-exponential convergence under *p*–refinement, with the asymptotic convergence curves in Figure 12 dipping downwards from the preasymptotic linear slope with increasing p. Furthermore, we note that in contrast to Example 1, the DG–norm of the error measured on the quadrilateral meshes (ii) consisting of irregular hanging nodes is now larger than when either the uniform square meshes or the conforming quadrilateral meshes (i) are used. As the flow is now aligned with the grid lines of the uniform square meshes, the increased communication on the meshes (ii) is no longer very important; instead, the error in the DG–norm is marginally increased due to the mesh distortion introduced by randomly splitting the interior nodes (excluding those on the line $y = 0$).

7 Appendix

Under the smoothness hypotheses on the data and with the notational conventions of Section 2.1, we consider the question of well-posedness of the following homogeneous Dirichlet–Neumann boundary value problem for a partial differential equation with nonnegative

characteristic form: find u such that

$$\begin{aligned} \mathcal{L}u \equiv -\nabla\cdot(a\nabla u) + b\cdot\nabla u + cu &= f \quad \text{in } \Omega\ , \\ u &= 0 \quad \text{on } \Gamma_{\rm D}\cup\Gamma_-\ , \\ (a\nabla u)\cdot\mu &= 0 \quad \text{on } \Gamma_{\rm N}\ . \end{aligned} \tag{7.1}$$

It is helpful to note the following simple result [16].

Lemma 17 *Suppose that M is a symmetric nonnegative definite $d\times d$ matrix. If $\zeta\in\mathbb{R}^d$ satisfies $\zeta^T M\zeta = 0$, then $M\zeta = 0$.*

It follows from the definition of Γ_0 that

$$\sum_{i,j=1}^{d} a_{ij}(x)\mu_i\mu_j = 0 \qquad \text{for } x\in\Gamma\setminus\Gamma_0\ . \tag{7.2}$$

Since $(a_{ij}(x))$ is a symmetric nonnegative definite $d\times d$ matrix for x in $\Gamma\setminus\Gamma_0$, we deduce from (7.2) and Lemma 17 with $M=a$ and $\zeta=\mu$ that

$$\sum_{j=1}^{d} a_{ij}(x)\mu_j = 0 \qquad \text{for } x\in\Gamma\setminus\Gamma_0\ , \qquad i=1,\ldots,d\ . \tag{7.3}$$

Now, suppose for a moment that (7.1) has a solution u and $u\in H^2(\Omega)$. By (7.3),

$$\sum_{i,j=1}^{d}\int_\Gamma a_{ij}(x)\frac{\partial u}{\partial x_i}\mu_j v\,{\rm d}s = 0 \qquad \text{for all } v\in\mathcal{V}\ , \tag{7.4}$$

where

$$\mathcal{V} = \left\{ v\in H^1(\Omega) : v(x) = 0 \ \text{ for } \ x\in\Gamma_{\rm D}\right\}\ .$$

This observation will be of key importance. On multiplying the partial differential equation in (7.1) by $v\in\mathcal{V}$ and integrating by parts, we find that

$$(a\nabla u,\nabla v) - (u,\nabla\cdot(bv)) + (cu,v) + \langle u,v\rangle_{\Gamma_+\cup\Gamma_{\rm N}} = (f,v) \qquad \text{for all } v\in\mathcal{V}\ , \tag{7.5}$$

where $(\cdot,\cdot)$ denotes the L^2 inner–product over Ω, $\langle\cdot,\cdot\rangle_{\Gamma_+\cup\Gamma_{\rm N}} = \langle\cdot,\cdot\rangle_{\Gamma_+} + \langle\cdot,\cdot\rangle_{\Gamma_{\rm N}}$

$$\langle\cdot,\cdot\rangle_{\Gamma_-\cup\Gamma_+\cup\Gamma_{\rm N}} = \langle\cdot,\cdot\rangle_{\Gamma_-} + \langle\cdot,\cdot\rangle_{\Gamma_+} + \langle\cdot,\cdot\rangle_{\Gamma_{\rm N}}\ ,$$

with

$$\langle w,v\rangle_{\Gamma_\pm} = \int_{\Gamma_\pm} |b\cdot\mu| wv\,{\rm d}s \qquad \text{and} \qquad \langle w,v\rangle_{\Gamma_{\rm N}} = \int_{\Gamma_{\rm N}} |b\cdot\mu| wv\,{\rm d}s\ .$$

We note that in the transition to (7.5) the boundary integral term on Γ which arises in the course of partial integration from the $-\nabla \cdot (a\nabla u)$ term vanishes by virtue of (7.4), while the boundary integral term on $\Gamma \setminus (\Gamma_+ \cup \Gamma_{\rm N}) = \Gamma_{\rm D} \cup \Gamma_-$ resulting from the $b \cdot \nabla u$ term on partial integration disappears since $u = 0$ on this set in (7.1). The form of (7.5) serves as motivation for the statement of the weak formulation of (7.1) which is presented below. We consider the inner product $(\cdot, \cdot)_{\mathcal{H}}$ defined by

$$(w, v)_{\mathcal{H}} = (a\nabla w, \nabla v) + (w, v) + \langle w, v \rangle_{\Gamma_- \cup \Gamma_+ \cup \Gamma_{\rm N}}$$

and denote by $\mathcal{H}$ the closure of $\mathcal{V}$ in $L^2(\Omega)$ with respect to the norm $\|\cdot\|_{\mathcal{H}}$ defined by $\|w\|_{\mathcal{H}} = (w, w)_{\mathcal{H}}^{1/2}$. Clearly, $\mathcal{H}$ is a Hilbert space. For $w \in \mathcal{H}$ and $v \in \mathcal{V}$, we now consider the bilinear form $B(\cdot, \cdot) : \mathcal{H} \times \mathcal{V} \to \mathbb{R}$ defined by

$$B(w, v) = (a\nabla w, \nabla v) - (w, \nabla \cdot (bv)) + (cw, v) + \langle w, v \rangle_{\Gamma_+ \cup \Gamma_{\rm N}}$$

and for $v \in \mathcal{V}$ we introduce the linear functional $\ell : \mathcal{V} \to \mathbb{R}$ by

$$\ell(v) = (f, v) \ .$$

We shall say that $u \in \mathcal{H}$ is a weak solution to the boundary value problem (7.1) if

$$B(u, v) = \ell(v) \quad \forall v \in \mathcal{V} \ . \tag{7.6}$$

In order to ensure the existence of a unique $u \in \mathcal{H}$ satisfying (7.6), in addition to assuming that $b \cdot \mu$ is nonnegative on $\Gamma_{\rm N}$, we also suppose that (3.2) holds.

The next theorem paraphrases Theorem 1.4.1 on p.29 in Oleinik & Radkevič [23], in the more general setup of mixed Dirichlet–Neumann boundary conditions and under less restrictive hypotheses on the coefficients a_{ij} than in [23]. The loosening of the regularity requirements on the coefficients here is related to the fact that the principal part of the operator in [23] is written in the non-divergence form

$$\sum_{i,j=1}^{d} a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} \ ,$$

so the associated Fichera function involves partial derivatives of the a_{ij}, and the counterpart of (3.2) contains second partial derivatives of the a_{ij} thus requiring that $a_{ij} \in W^{2,\infty}(\Omega)$, which is more demanding than the piecewise continuity of the a_{ij} assumed here (in fact, when $\Gamma_{\rm N} = \emptyset$ even $a_{ij} \in L^\infty(\Omega)$ will suffice).

Theorem 18 *Assuming that (3.2) is valid, for each $f \in L^2(\Omega)$ there exists $u \in \mathcal{H}$ such that (7.6) holds. Moreover, there exists a Hilbert subspace $\mathcal{H}'$ of $\mathcal{H}$ such that $u \in \mathcal{H}'$, and u is the unique element in $\mathcal{H}'$ such that (7.6) holds.*

Proof The proof is based on the Riesz representation theorem. For $v \in \mathcal{V}$ fixed, by the Cauchy–Schwarz inequality,

$$B(w,v) \le K_1 \|w\|_{\mathcal{H}} \|v\|_{H^1(\Omega)} \qquad \forall w \in \mathcal{H} \,,$$

where we have used the trace theorem for $H^1(\Omega)$. Thus $B(\cdot, v)$ is a bounded linear functional on $\mathcal{H}$. By the Riesz representation theorem, there exists a unique element of $\mathcal{H}$, denoted by $T(v)$ such that

$$B(w,v) = (w, T(v))_{\mathcal{H}} \qquad \forall w \in \mathcal{H} \,.$$

Since B is bilinear, it follows that $T : v \to T(v)$ is a linear operator from $\mathcal{V}$ into $\mathcal{H}$. Next we show that T is injective. Note that

$$B(v,v) = (a\nabla v, \nabla v) - (v, \nabla \cdot (bv)) + (cv, v) + \langle v, v\rangle_{\Gamma_+ \cup \Gamma_{\rm N}} \qquad \forall v \in \mathcal{V} \,.$$

On integrating by parts in the second term on the right–hand side and applying (3.2),

$$B(v,v) \ge (a\nabla v, \nabla v) + \gamma_0 \|v\|^2 + \frac{1}{2}\langle v, v\rangle_{\Gamma_- \cup \Gamma_+ \cup \Gamma_{\rm N}} \ge K_0 \|v\|^2_{\mathcal{H}} \qquad \forall v \in \mathcal{V} \,,$$

where $K_0 = \min(\gamma_0, \frac{1}{2})$. Hence

$$(v, T(v))_{\mathcal{H}} \ge K_0 \|v\|^2_{\mathcal{H}} \qquad \forall v \in \mathcal{V} \,; \tag{7.7}$$

so $T : v \to T(v)$ is an injection from $\mathcal{V}$ onto the range $\mathcal{R}(T)$ of T contained in $\mathcal{H}$. Let $\mathcal{H}'$ denote the closure of $\mathcal{R}(T)$ in $\mathcal{H}$ with respect to the norm $\|\cdot\|_{\mathcal{H}}$. Clearly, by (7.7)

$$|\ell(v)| \le \|f\|_{L^2(\Omega)} \|v\|_{\mathcal{H}} \le K_0^{-1} \|f\|_{L^2(\Omega)} \|T(v)\|_{\mathcal{H}} \qquad \forall v \in \mathcal{V} \,. \tag{7.8}$$

Given $w \in \mathcal{H}'$, consider a sequence (w_n) in $\mathcal{R}(T)$ such that $w_n \to w$ in $\mathcal{H}$. Define

$$g(w) = \lim_{n \to \infty} \ell(T^{-1} w_n) \,. \tag{7.9}$$

As $\|T^{-1} w\|_{\mathcal{H}} \le (1/K_0)\|w\|_{\mathcal{H}}$ for all $w \in \mathcal{R}(T)$, it is easily seen using (7.8) that the definition of g is correct in the sense that the limit exists and is independent of the choice of the sequence (w_n). By (7.8) and (7.9), on noting that $\lim_{n\to\infty} \|w_n\|_{\mathcal{H}} = \|w\|_{\mathcal{H}}$, we deduce that

$$|g(w)| \le K_0^{-1} \|f\|_{L^2(\Omega)} \|w\|_{\mathcal{H}} \,.$$

Thus, g is a bounded linear functional on $\mathcal{H}'$. Since $\mathcal{H}'$ is closed (by definition) in the norm of $\mathcal{H}$, it is a Hilbert subspace of $\mathcal{H}$. By the Riesz representation theorem, there exists a unique $u \in \mathcal{H}'$ such that $g(w) = (u, w)_{\mathcal{H}}$ for all w in $\mathcal{H}'$; in particular,

$$g(Tv) = (u, Tv)_{\mathcal{H}} \qquad \forall v \in \mathcal{V} \,. \tag{7.10}$$

However, as the (constant) sequence (w_n) with $w_n = Tv$, $n = 1, 2, \ldots$, converges to $w = Tv$ and since the value of $g(w)$ is independent of the choice of the sequence (w_n), it follows from (7.9) that

$$g(Tv) = g(w) = \lim_{n \to \infty} \ell(T^{-1} w_n) = \ell(v) \qquad \forall v \in \mathcal{V} \,. \tag{7.11}$$

From (7.10) and (7.11) we have that $\ell(v) = (u, Tv)_{\mathcal{H}}$ for all v in $\mathcal{V}$. Thus we have shown the existence of a unique $u \in \mathcal{H}'(\subset \mathcal{H})$ such that

$$B(u,v) \equiv (u,Tv)_{\mathcal{H}} = \ell(v) \qquad \forall v \in \mathcal{V} \ .$$

■

Theorem 18 is a generalization of the well-posedness result in our paper [16] to the case of mixed Dirichlet–Neumann boundary conditions.

Acknowledgements

We wish to express are sincere gratitude to Andrea Toselli (Courant Institute, New York University), and Max Jensen and Emmanuil Georgoulis (University of Oxford) for helpful comments on an earlier version of this paper.

References

[1] D.N. Arnold, *An interior penalty finite element method with discontinuous elements.* SIAM J. Numer. Anal., 19:742–760, 1982.

[2] I. Babuška and M. Suri, *The hp–version of the finite element method with quasiuniform meshes.* M^2AN Mathematical Modelling and Numerical Analysis, 21:199–238, 1987.

[3] C. Baumann, *An hp–adaptive discontinuous Galerkin FEM for computational fluid dynamics.* Doctoral Dissertation, TICAM, UT Austin, Texas, 1997.

[4] K.S. Bey and J.T. Oden, *hp–Version discontinuous Galerkin methods for hyperbolic conservation laws.* Comput. Methods Appl. Mech. Engrg., 133:259–286, 1996.

[5] D. Braess, *Finite Elements. Theory, Fast Solvers, and Applications in Solid Mechanics.* Cambridge University Press, 1997.

[6] P. Castillo, B. Cockburn, I. Perugia and D. Schötzau, *An a priori error analysis of the local discontinuous Galerkin method for elliptic problems.* (Submitted to SIAM J. Numer. Anal.)

[7] B. Cockburn, G. Kanschat, I. Perugia and D. Schötzau, *Superconvergence of the local discontinuous Galerkin method for elliptic problems on Cartesian Grids.* (Submitted to SIAM J. Numer. Anal.)

[8] B. Cockburn, S. Hou and C.-W. Shu, *TVB Runge–Kutta local projection discontinuous Galerkin finite elements for hyperbolic conservation laws.* Math. Comp., 54:545–581, 1990.

[9] B. Cockburn and C.-W. Shu, *TVB Runge–Kutta local projection discontinuous Galerkin finite element method for scalar conservation laws II: General framework,* Math. Comp., 52:411–435, 1989.

[10] B. Cockburn and C.-W. Shu, *The Runge–Kutta local projection P^1–discontinuous Galerkin method for scalar conservation laws,* RAIRO Modé. Math. Anal. Numér., 25:337–361, 1991.

[11] B. Cockburn and C.-W. Shu, *The local discontinuous Galerkin method for time–dependent reaction–diffusion systems.* SIAM J. Numer. Anal., 35:2440–2463, 1998.

[12] B. Cockburn and C.-W. Shu, *The Runge–Kutta discontinuous Galerkin method for conservation laws: Multidimensional systems,* J. Comput. Phys. 141:199–244, 1998.

[13] R.S. Falk and G.R. Richter, *Local error estimates for a finite element method for hyperbolic and convection–diffusion equations.* SIAM J. Numer. Anal., 29:730–754, 1992.

[14] J.E. Flaherty, R.M. Loy, M.S. Shephard and J.D. Teresco, *Software for parallel adaptive solution of conservation laws by discontinuous Galerkin methods,* In: B. Cockburn, G.E. Karniadakis and C.-W. Shu (Eds.), Discontinuous Galerkin Methods: Theory, Computation and Applications, Lecture Notes in Computational Science and Engineering, Vol. 11, Springer–Verlag, Berlin, 2000, pp.113–123.

[15] P. Houston, Ch. Schwab and E. Süli, *Stabilized hp–finite element methods for first–order hyperbolic problems.* SIAM J. Numer. Anal., 37(5):1618–1643, 2000.

[16] P. Houston and E. Süli, *Stabilized hp-finite element approximation of partial differential equations with non-negative characteristic form.* (Submitted to Computing).

[17] C. Johnson, U. Nävert and J. Pitkäranta, *Finite element methods for linear hyperbolic problems.* Comp. Meth. Appl. Mech. Engrg., 45:285–312, 1984.

[18] C. Johnson and J. Pitkäranta, *An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation.* Math. Comp. 46:1–26, 1986.

[19] G.E. Karniadakis and S. Sherwin, *Spectral/hp Finite Element Methods in CFD.* Oxford University Press, 1999.

[20] P. Lesaint and P.A. Raviart, *On a finite element method for solving the neutron transport equation.* In: Mathematical aspects of Finite Elements in Partial Differential equations, C.A. deBoor (Ed.), Academic Press, pp. 89–123, 1974.

[21] J.M. Melenk and Ch. Schwab, *An hp finite element method for convection–diffusion problems.* Seminar for Applied Mathematics, ETH Zürich Technical Report No 97-05.

[22] J. Nitsche, *Über ein Variationsprinzip zur Lösung von Dirichlet Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind.* Abh. Math. Sem. Univ. Hamburg, 36:9–15, 1971.

[23] O.A. Oleinik and E.V. Radkevič. *Second Order Equations with Nonnegative Characteristic Form.* American Mathematical Society, 1973.

[24] J.T. Oden, I. Babuška and C. Baumann, *A discontinuous hp–FEM for diffusion problems.* J. Comput. Phys., 146:491–519, 1998.

[25] G. Richter, *An optimal–order error estimate for the discontinuous Galerkin method.* Math. Comp. 50:75–88, 1988.

[26] G. Richter, *The discontinuous Galerkin method with diffusion.* Math. Comp. 58:631–643, 1992.

[27] B. Rivière, M.F. Wheeler and V. Girault, Improved energy estimates for interior penalty, constrained and discontinuous Galerkin methods for elliptic problems. Part I. Computational Geosciences 3 (1999)Vol. 3-4: 337–360, 1999.

[28] W.H. Reed and T.R. Hill, *Triangular mesh methods for the neutron transport equation.* Technical Report LA-UR-73-479, Los Alamos Scientific Laboratory, 1973.

[29] Ch. Schwab, *p– and hp–Finite Element Methods. Theory and Applications to Solid and Fluid Mechanics.* Oxford University Press, 1998.

[30] Ch. Schwab, *hp–FEM for fluid flow simulation.* In: Lecture Notes on high order methods in CFD, T. Barth and M. Deconinck (Eds.), Lecture Notes in Computational Science and Engineering, Vol. 9, Springer–Verlag, Berlin, 1999.

[31] E. Süli, P. Houston and Ch. Schwab, *hp–Finite element methods for hyperbolic problems.* In: J R Whiteman, editor, The Mathematics of Finite Elements and Applications. MAFELAP X. Elsevier, 2000.

[32] E. Süli, Ch. Schwab and P. Houston, *hp–DGFEM for partial differential equations with nonnegative characteristic form.* In: B. Cockburn, G.E. Karniadakis and C.-W. Shu (Eds.), Discontinuous Galerkin Methods: Theory, Computation and Applications, Lecture Notes in Computational Science and Engineering, Vol. 11, Springer–Verlag, Berlin, 2000, pp.221–230.

[33] M.F. Wheeler, *An elliptic collocation finite element method with interior penalties.* SIAM J. Numer. Anal., 15:152–161, 1978.