A Péclet-robust discontinuous Galerkin method for nonlinear diffusion with advection

Lourenço Beirão da Veiga, Daniele Antonio Di Pietro, Kirubell Biniam Haile

A Péclet-robust discontinuous Galerkin method for nonlinear diffusion with advection

Lourenço Beirão da Veiga[1,2], Daniele A. Di Pietro[3], and Kirubell B. Haile[1]

[1]Dipartimento di Matematica e Applicazioni, Università di Milano Bicocca,
lourenco.beirao@unimib.it, k.haile@campus.unimib.it
[2]IMATI-PV, CNR, Pavia, Italy
[3]IMAG, Univ. Montpellier, CNRS, Montpellier, France, daniele.di-pietro@umontpellier.fr

February 14, 2024

Abstract

We analyze a Discontinuous Galerkin method for a problem with linear advection-reaction and p-type diffusion, with Sobolev indices $p \in (1, \infty)$. The discretization of the diffusion term is based on the full gradient including jump liftings and interior-penalty stabilization while, for the advective contribution, we consider a strengthened version of the classical upwind scheme. The developed error estimates track the dependence of the local contributions to the error on local Péclet numbers. A set of numerical tests supports the theoretical derivations.

Key words. Discontinuous Galerkin methods, diffusion-advection-reaction problems, p-Laplacian, Péclet-robust error estimates

MSC2010. 65N30, 65N08, 35K55

1 Introduction

Discontinuous Galerkin (DG) methods were introduced in the 70s [4, 27] and have gained significant popularity starting from the late 90s [1, 2, 6, 7, 12–15, 20, 21]. They are nowadays widely regarded as the reference methods for advection-dominated problems. When a polynomial degree $k \geq 1$ is used, classical error estimates for linear diffusion-advection(-reaction) problems show that the error contribution stemming from diffusive terms is $O(h^k)$ (with h denoting the meshsize), while the one stemming from advective terms is $O(h^{k+\frac{1}{2}})$; see, e.g., [3] and also [22] and [21, Section 4.6] for an analysis covering the locally degenerate case. Pre-asymptotic convergence rates between k and $k + \frac{1}{2}$ can be observed, in practice, when sufficiently coarse meshes are considered. Standard estimates do not usually allow, however, a quantitative assessment of this phenomenon. Error estimates are, on the other hand, completely missing for problems with non-linear diffusion terms.

The goal of this work is to fill the above gaps by deriving Péclet-dependent error estimates for a problem with linear advection-reaction and p-type diffusion, for Sobolev indices $p \in (1, \infty)$. The discretization of the diffusion term is, similarly to [12, 16], based on the full gradient including jump liftings and interior-penalty stabilization. For the advective contribution, on the other hand, we consider a strengthened version of the classical upwind scheme obtained interpreting the latter as a penalty contribution in the spirit of [11]. The peculiarity of our error estimates is that they track the dependence of the local contributions to the error on local Péclet numbers. To improve the estimates of certain terms, we provide a new extension to the nonconforming case of the techniques of [25], based in turn on the

results of [23] (see also [5]). This requires a certain number of subtleties, both in the adaptation of the argument and in the definition of the face Péclet numbers (which need to account for both the physical and numerical diffusion). To the best of our knowledge, our Péclet-dependent error estimates are the first of this kind for a nonlinear problem, and enable a quantitative assessment of pre-asymptotic convergence rates. In the linear case, corresponding to $p = 2$, local Péclet numbers can be computed based on the sole knowledge of the problem data and the mesh, making it possible to identify a priori advection- and diffusion-dominated elements/faces. Incidentally, new error estimates for the DG discretization of the p-Laplace problem are also recovered as a special case (the previous works [12, 16] only considered convergence by compactness). The theoretical results are supported by extensive numerical validation.

The present contribution furthermore sets the stage for future publications developing pressure robust and advection-robust finite elements for time-dependent Navier–Stokes type equations (e.g. [9, 24]) modeling incompressible fluid flows with non-Newtonian rheology.

The rest of this work is organized as follows. In Section 2 we describe the continuous problem. After presenting some definitions and preliminary results in Section 3, the numerical scheme is introduced in Section 4 along with the main theoretical results. The proofs of the latter are given in Section 5. Finally, numerical tests are collected in Section 6.

2 The continuous problem

Let $\Omega \subset \mathbb{R}^d$, $d \geq 1$, denote a bounded, connected polyhedral domain. We develop a Péclet-robust discontinuous Galerkin (DG) method for the following problem: Find $u : \Omega \to \mathbb{R}$ such that

$$\begin{aligned} -\nabla \cdot [\sigma(\nabla u) - \beta u] + \mu u &= f \qquad \text{in } \Omega, \\ u &= 0 \qquad \text{on } \partial\Omega. \end{aligned} \tag{1}$$

Here above, we assume that the velocity field satisfies $\beta \in W^{1,\infty}(\Omega)^d$ and, for the sake of simplicity, that $\nabla \cdot \beta = 0$ almost everywhere in Ω. Furthermore, we assume $\mu(x) \geq \underline{\mu} > 0$ for almost every $x \in \Omega$ with $\underline{\mu} \in \mathbb{R}$. The extension to non-incompressible velocity fields is standard, and essentially requires to assume a positive lower bound on the quantity $\mu + \frac{1}{2}(\nabla \cdot \beta)$ instead of μ. The function σ represents the diffusive flux function, which we describe below.

For given real number $p \geq 1$ and integer $n \geq 1$, we consider the power flux function

$$\sigma_n : \mathbb{R}^n \ni x \mapsto |x|^{p-2}x \in \mathbb{R}^n$$

with $|\cdot|$ denoting the Euclidian norm. In what follows, for the sake of brevity, we omit the subscript when $n = d$, i.e., we set $\sigma := \sigma_d$. The following derivations can be extended to more general flux functions satisfying appropriate p-monotonicity and p-continuity properties characterizing Leray–Lions-type operators and their generalizations; see, e.g., [23, 26].

In what follows, to alleviate the notation, we will use the symbol $c(\delta)$ for a generic constant, possibly different at each occurrence, which depends on the parameter δ but is independent of the meshsize (see below), the problem data and solution.

Lemma 1 (Modified monotonicity of the power flux function)**.** *Let $p \in (1, \infty)$ and an integer $n \geq 1$ be given. For all $x, y, z \in \mathbb{R}^n$ and any real number $\delta > 0$, it holds*

$$(\sigma_n(x) - \sigma_n(z)) \cdot (x - y) \leq \delta(\sigma_n(y) - \sigma_n(x)) \cdot (y - x) + c(\delta)\big(|x| + |z|\big)^{p-2}|x - z|^2, \tag{2}$$

with $c(\delta)$ positive constant depending only on p and δ.

Proof. Throughout this proof, $a \lesssim b$ means $a \leq b$ with hidden constant only depending on p, while $a \simeq b$ stands for "$a \lesssim b$ and $b \lesssim a$". Let $\varphi : \mathbb{R}^+ \ni t \mapsto \frac{1}{p}t^p \in \mathbb{R}^+$ and, for any $a \geq 0$ and any $t \geq 0$, let

$\varphi_a(t) \coloneqq \int_0^t \varphi'(a+s)\frac{s}{a+s}\,\mathrm{d}s$. By [23, Eq. (6.28)], it holds, for all $s, t \in \mathbb{R}^+$ and all $\delta > 0$,

$$\varphi_a'(s)t + \varphi_a'(t)s \le \delta\varphi_a(s) + c(\delta)\varphi_a(t). \tag{3}$$

Moreover, by [23, Lemma 3], we have, for all $x, y \in \mathbb{R}^n$,

$$(\sigma_n(y) - \sigma_n(x)) \cdot (y - x) \simeq \varphi_{|x|}(|x - y|) \simeq (|x| + |y|)^{p-2}|x - y|^2 \tag{4}$$

and, as observed in [25, Lemma 2.3], for all $x, z \in \mathbb{R}^n$,

$$|\sigma_n(x) - \sigma_n(z)| \lesssim \varphi'_{|x|}(|x - z|). \tag{5}$$

Let now $x, y, z \in \mathbb{R}^n$. Applying (3) with $a = |x|$, $s = |x-y|$, and $t = |x-z|$, and noticing that $\varphi_a'(s)t \ge 0$, we get

$$\varphi'_{|x|}(|x - z|)|x - y| \le \delta\varphi_{|x|}(|x - y|) + c(\delta)\varphi_{|x|}(|x - z|). \tag{6}$$

The conclusion follows first observing that $(\sigma_n(x) - \sigma_n(z)) \cdot (x - y) \le |\sigma_n(x) - \sigma_n(z)|\,|x - y|$ and using (5) to estimate the left-hand side of (2) with the left-hand side of (6), then applying, respectively, the left-most equivalence in (4) to estimate $\varphi_{|x|}(|x - y|)$ and the right-most equivalence in (4) to estimate $\varphi_{|x|}(|x - z|)$ in the right-hand side of (6). □

3 The discrete setting and preliminary results

We denote by $\mathcal{T}_h$ a mesh of Ω belonging to an admissible sequence $\{\mathcal{T}_h\}_h$ in the sense of [21, Section 1.4]. For any $T \in \mathcal{T}_h$, we denote by ω_T the union of the mesh elements sharing at least one face with T, which are collected in the set $\mathcal{T}_T$. We moreover denote by $\mathcal{F}_h$ the set of faces, partitioned into boundary faces collected in $\mathcal{F}_h^{\mathrm{b}}$ and interfaces collected in $\mathcal{F}_h^{\mathrm{i}}$. Given a face $F \in \mathcal{F}_h$, we denote by $\mathcal{T}_F$ the set of mesh elements sharing F and by ω_F their union. Furthermore, for any $T \in \mathcal{T}_h$ we denote by $\mathcal{F}_T$ the set of its faces. For any mesh element or face $Y \in \mathcal{T}_h \cup \mathcal{F}_h$, we denote by h_Y its diameter and set $h \coloneqq \max_{T\in\mathcal{T}_h} h_T$. Since $\mathcal{T}_h$ belongs to an admissible mesh sequence, the maximum number of faces of a mesh element is bounded uniformly in h.

To avoid naming generic constants, from this point on we will use the notation $a \lesssim b$ to express the inequality $a \le Cb$ with C independent of the meshsize, of the problem data and solution, but possibly depending on other quantities including the domain, the ambient dimension d, the mesh regularity parameter, and the Sobolev index p. We will write $a \simeq b$ in lieu of "$a \lesssim b$ and $b \lesssim a$".

Remark 2 (Polytopal meshes). We underline that the present results apply not only to standard type of grids, but also to general polytopal meshes. For a few (among many) examples of other polytopal schemes in a similar context, see for example [8, 10, 17, 19].

3.1 Local and broken spaces

Given a polynomial degree $k \ge 0$ and a mesh element $T \in \mathcal{T}_h$, we denote by $\mathcal{P}^k(T)$ the space spanned by the restriction to T of d-variate polynomial functions. At the global level, we define the broken polynomial space

$$\mathcal{P}^k(\mathcal{T}_h) \coloneqq \left\{ v_h \in L^1(\Omega) \mid v_T \coloneqq (v_h)_{|T} \in \mathcal{P}^k(T) \text{ for all } T \in \mathcal{T}_h \right\}.$$

The L^2-orthogonal projector on $\mathcal{P}^k(\mathcal{T}_h)$ is denoted by π_h^k and is obtained patching together the L^2-orthogonal projectors π_T^k on $\mathcal{P}^k(T)$, $T \in \mathcal{T}_h$. The same notations are used for vector versions of these projectors mapping on $\mathcal{P}^k(\mathcal{T}_h)^d$ or $\mathcal{P}^k(T)^d$ and acting component-wise. Letting $Y \in \mathcal{T}_h \cup \mathcal{F}_h \cup \{\Omega\}$, we denote by $W^{q,p}(Y)$ the usual Sobolev space on Y and we set

$$W^{q,p}(\mathcal{T}_h) \coloneqq \left\{ v \in L^p(\Omega) \mid v_{|T} \in W^{q,p}(T) \text{ for all } T \in \mathcal{T}_h \right\}.$$

We will also need broken spaces defined on local patches $\mathcal{T}_Y$, $Y \in \mathcal{T}_h \cup \mathcal{F}_h$, defined in a similar way. Finally, for the Hilbertian case $p = 2$, we will also use the habitual abridged notations $H^q := W^{q,2}$ and $L^2 := H^0$.

For future use, for any $p \in (1, +\infty)$ we define the conjugate index p' such that

$$\frac{1}{p} + \frac{1}{p'} = 1 \iff p' = \frac{p}{p-1}. \tag{7}$$

The above definition can be generalized to $p \in \{1, \infty\}$ setting $\frac{1}{\infty} := 0$ and $\frac{1}{0} := \infty$.

3.2 Trace operators and integration by parts formula

For each interface $F \in \mathcal{F}_h^{\mathrm{i}}$, we fix once and for all an orientation for the unit normal vector n_F. Denoting by T_1 and T_2 the elements sharing F ordered so that n_F points out of T_1, we define the jump and average operators such that, for any $\varphi \in W^{1,1}(\mathcal{T}_h)$,

$$[\varphi]_F := \varphi_{|T_1} - \varphi_{|T_2}, \qquad \{\varphi\}_F := \frac{1}{2}\left(\varphi_{|T_1} + \varphi_{|T_2}\right).$$

When applied to vector-valued functions, these operators act component-wise. The above operators are extended to boundary faces $F \in \mathcal{F}_h^{\mathrm{b}}$ setting

$$[\varphi]_F = \{\varphi\}_F := \varphi.$$

We recall the following integration by parts formula: For all $\tau : \Omega \to \mathbb{R}^d$ and $v : \Omega \to \mathbb{R}$ smooth enough,

$$\int_\Omega \tau \cdot \nabla_h v = -\int_\Omega (\nabla_h \cdot \tau)\, v + \sum_{F \in \mathcal{F}_h^{\mathrm{i}}} \int_F \{\tau\}_F \cdot n_F\, [v]_F + \sum_{F \in \mathcal{F}_h^{\mathrm{i}}} \int_F [\tau]_F \cdot n_F\, \{v\}_F + \int_{\partial\Omega} (\tau \cdot n)\, v, \tag{8}$$

where n denotes the unit normal vector field on $\partial\Omega$ pointing out of Ω.

3.3 Jump liftings and discrete gradient

The jumps of smooth enough functions can be lifted to polynomial functions defined over Ω. Specifically, given an integer $k \geq 0$, for each $F \in \mathcal{F}_h$ we define the local trace lifting $r_F^k : L^1(F) \to \mathcal{P}^k(\mathcal{T}_h)^d$ such that, for all $\psi \in L^1(F)$,

$$\int_\Omega r_F^k \psi \cdot \tau_h = \int_F \psi\, \{\tau_h\}_F \cdot n_F \qquad \forall \tau_h \in \mathcal{P}^k(\mathcal{T}_h)^d \tag{9}$$

and we let $R_h^k : W^{1,1}(\mathcal{T}_h) \to \mathcal{P}^k(\mathcal{T}_h)^d$ be the global face jumps lifting such that, for any $\varphi \in W^{1,1}(\mathcal{T}_h)$,

$$R_h^k \varphi := \sum_{F \in \mathcal{F}_h} r_F^k([\varphi]_F). \tag{10}$$

Finally, we define the discrete gradient $G_h^k : W^{1,1}(\mathcal{T}_h) \to L^1(\Omega)^d$ setting

$$G_h^k \varphi := \nabla_h \varphi - R_h^k \varphi, \tag{11}$$

where ∇_h denotes the broken gradient on $\mathcal{T}_h$.

For any $p \in [1, \infty)$, we define the following broken norm: For all $\varphi \in W^{1,p}(\mathcal{T}_h)$,

$$\|\varphi\|_{1,p,h} := \left(\|\nabla_h \varphi\|_{L^p(\Omega)^d}^p + |\varphi|_{1,p,h}^p\right)^{\frac{1}{p}} \text{ with } |\varphi|_{1,p,h} := \left(\sum_{F \in \mathcal{F}_h} h_F^{1-p} \|[\varphi]_F\|_{L^p(F)}^p\right)^{\frac{1}{p}}, \tag{12}$$

which extends as follows to the case $p = \infty$:

$$\|\varphi\|_{1,\infty,h} \coloneqq \|\nabla_h\varphi\|_{L^\infty(\Omega)^d} + |\varphi|_{1,\infty,h} \text{ with } |\varphi|_{1,\infty,h} \coloneqq \max_{F\in\mathcal{F}_h} h_F^{-1}\|[\varphi]_F\|_{L^\infty(F)}. \tag{13}$$

In local estimates, we will also need the following local versions of the norms (12) and (13): For all $T \in \mathcal{T}_h$ and all $\varphi \in W^{1,p}(\mathcal{T}_T)$,

$$\|\varphi\|_{1,p,T} \coloneqq \left(\|\nabla_h\varphi\|_{L^p(T)^d}^p + |\varphi|_{1,p,T}^p\right)^{\frac{1}{p}} \text{ with } |\varphi|_{1,p,T} \coloneqq \left(\sum_{F\in\mathcal{F}_T} h_F^{1-p}\|[\varphi]_F\|_{L^p(F)}^p\right)^{\frac{1}{p}} \tag{14}$$

and

$$\|\varphi\|_{1,\infty,T} \coloneqq \|\nabla_h\varphi\|_{L^\infty(T)^d} + |\varphi|_{1,\infty,T} \text{ with } |\varphi|_{1,\infty,T} \coloneqq \max_{F\in\mathcal{F}_T} h_F^{-1}\|[\varphi]_F\|_{L^\infty(F)}.$$

It is easy to check that, for all $\varphi \in W^{1,p}(\mathcal{T}_h)$, $p \in [1,\infty]$,

$$\|\varphi\|_{1,p,h}^p \simeq \sum_{T\in\mathcal{T}_h} \|\varphi\|_{1,p,T}^p \quad \text{and} \quad |\varphi|_{1,p,h}^p \simeq \sum_{T\in\mathcal{T}_h} |\varphi|_{1,p,T}^p.$$

Lemma 3 (Properties of the jump lifting)**.** *It holds, for any integer $k \geq 0$ and any $p \in [1,\infty]$:*

1. Boundedness. *For all $\varphi \in W^{1,p}(\mathcal{T}_h)$, it holds*

$$\|r_F^k([\varphi]_F)\|_{L^p(\Omega)^d} \lesssim h_F^{\frac{1-p}{p}}\|[\varphi]_F\|_{L^p(F)} \qquad \forall F \in \mathcal{F}_h \tag{15}$$

with the convention that $h_F^{\frac{1-\infty}{\infty}} \coloneqq h_F^{-1}$ and

$$\|R_h^k\varphi\|_{L^p(T)^d} \lesssim |\varphi|_{1,p,T} \qquad \forall T \in \mathcal{T}_h. \tag{16}$$

2. Approximation. *For any $w \in W_0^{1,p}(\Omega)$ (with $W_0^{1,p}(\Omega)$ denoting the closure of $C_c^\infty(\Omega)$ in $W^{1,p}(\Omega)$) such that $w \in W^{r+1,p}(\mathcal{T}_h)$ for some $r \in \{0,\dots,k\}$,*

$$\|R_h^k\pi_h^k w\|_{L^p(T)^d} \lesssim h_T^r|w|_{W^{r+1,p}(\mathcal{T}_T)} \qquad \forall T \in \mathcal{T}_h. \tag{17}$$

Proof. *Proof of* (15)–(16). It holds, for all $F \in \mathcal{F}_h$,

$$\begin{aligned}\|r_F^k([\varphi]_F)\|_{L^p(\Omega)^d} &= \sup_{\tau\in L^{p'}(\Omega)^d\setminus\{0\}} \frac{\int_\Omega r_F^k([\varphi]_F)\cdot\tau}{\|\tau\|_{L^{p'}(\Omega)^d}} \\ &= \sup_{\tau\in L^{p'}(\Omega)^d\setminus\{0\}} \frac{\int_\Omega r_F^k([\varphi]_F)\cdot\pi_h^k\tau}{\|\tau\|_{L^{p'}(\Omega)^d}} \overset{(9)}{=} \sup_{\tau\in L^{p'}(\Omega)^d\setminus\{0\}} \frac{\int_F [\varphi]_F\{\pi_h^k\tau\}_F\cdot n_F}{\|\tau\|_{L^{p'}(\Omega)^d}},\end{aligned}$$

where the introduction of the L^2-orthogonal projector π_h^k in the second equality is made possible by its definition. We next write, setting $h^{-\frac{1}{p'}} \coloneqq 1$ if $p' = \infty$,

$$\left|\int_F [\varphi]_F\{\pi_h^k\tau\}_F\cdot n_F\right| \lesssim h_F^{-\frac{1}{p'}}\|[\varphi]_F\|_{L^p(F)}\|\pi_h^k\tau\|_{L^{p'}(\mathcal{T}_F)^d} \lesssim h_F^{-\frac{1}{p'}}\|[\varphi]_F\|_{L^p(F)}\|\tau\|_{L^{p'}(\mathcal{T}_F)^d},$$

where we have used a Hölder inequality with exponents (p,p',∞) along with the fact that $\|n_F\|_{L^\infty(F)^d} \leq 1$ followed by the discrete trace inequality [18, Lemma 1.32] in the first passage, while the second passage

is a consequence of the $L^{p'}$-boundedness of the L^2-orthogonal projector (cf. [18, Lemma 1.44]). Additionally noticing that, by (7), $-\frac{1}{p'} = \frac{1-p}{p}$ and that $\|\tau\|_{L^{p'}(\mathcal{T}_F)^d} \le \|\tau\|_{L^{p'}(\Omega)^d}$, yields (15).

Let now $T \in \mathcal{T}_h$. In order to estimate $\|R_h^k\varphi\|_{L^p(T)^d}$, we first recall that, for any $F \in \mathcal{F}_h$, the support of $r_F^k([\varphi]_F)$ is ω_F, then use a triangle inequality together with (15), and finally use $\mathrm{card}(\mathcal{F}_T) \lesssim 1$:

$$\|R_h^k\varphi\|_{L^p(T)^d} = \left\| \sum_{F\in\mathcal{F}_T} r_F^k([\varphi]_F) \right\|_{L^p(T)^d} \lesssim \sum_{F\in\mathcal{F}_T} h_F^{\frac{1-p}{p}} \|[\varphi]_F\|_{L^p(F)} \lesssim |\varphi|_{1,p,T},$$

which is the bound (16).

Proof of (17). If $p \in [1,\infty)$, using the result proved in the previous point, we can write, for all $T \in \mathcal{T}_h$,

$$\begin{aligned}\|R_h^k\pi_h^k w\|_{L^p(T)^d}^p &\lesssim \sum_{F\in\mathcal{F}_T} h_F^{1-p} \|[\pi_h^k w]_F\|_{L^p(F)}^p \\ &= \sum_{F\in\mathcal{F}_T} h_F^{1-p} \|[\pi_h^k w - w]_F\|_{L^p(F)}^p \lesssim h_T^{pr} \sum_{F\in\mathcal{F}_T} |w|_{W^{r+1,p}(\mathcal{T}_F)}^p,\end{aligned}$$

where, to insert w in the second passage, we have used the fact that its jumps vanish across interfaces and its trace on $\partial\Omega$ is zero, while the conclusion follows from scaled trace inequalities and approximation properties of the L^2-orthogonal projector, additionally recalling that $h_T \lesssim h_{T'}$ for all $T' \in \mathcal{T}_T$ by mesh regularity. Using $\mathrm{card}(\mathcal{F}_T) \lesssim 1$, (17) follows. If $p = \infty$, we have

$$\|R_h^k\pi_h^k w\|_{L^\infty(T)^d} \lesssim \sum_{F\in\mathcal{F}_T} h_F^{-1} \|[\pi_h^k w]_F\|_{L^\infty(F)} = \sum_{F\in\mathcal{F}_T} h_F^{-1} \|[\pi_h^k w - w]_F\|_{L^\infty(F)} \lesssim h_T^r |w|_{W^{r+1,\infty}(\mathcal{T}_T)},$$

which concludes the proof. □

Lemma 4 (Approximation properties of the discrete gradient)**.** *For all integer $k \ge 0$, all $p \in [1,\infty]$, and all $w \in W_0^{1,p}(\Omega) \cap W^{r+1,p}(\mathcal{T}_h)$ with $r \in \{0,\dots,k\}$, it holds*

$$\|G_h^k\pi_h^k w - \nabla w\|_{L^p(T)^d} \lesssim h_T^r |w|_{W^{r+1,p}(\mathcal{T}_T)} \qquad \forall T \in \mathcal{T}_h. \tag{18}$$

Proof. Using (11) and a triangle inequality, we obtain

$$\|G_h^k\pi_h^k w - \nabla w\|_{L^p(T)^d} \le \|\nabla\pi_T^k w - \nabla w\|_{L^p(T)^d} + \|R_h^k\pi_h^k w\|_{L^p(T)^d}.$$

The conclusion follows using the approximation properties of the L^2-orthogonal projector for the first term and (17) for the second. □

Remark 5 (Local boundedness of $G_h^k \circ \pi_h^k$). For any $q \in [1,\infty]$, combining a triangle inequality with (18) written for $r = 0$, it is readily inferred that, for all $\varphi \in W^{1,q}(\mathcal{T}_h)$,

$$\|G_h^k\pi_h^k\varphi\|_{L^q(T)^d} \lesssim |\varphi|_{W^{1,q}(\mathcal{T}_T)} \qquad \forall T \in \mathcal{T}_h. \tag{19}$$

4 Discrete problem and main results

4.1 Discrete problem

From this point on, we let a Sobolev exponent $p \in (1,\infty)$ and polynomial degree $k \ge 1$ be fixed. The diffusion term is discretized, similarly to what is proposed in [12], by the function $a_h : \mathcal{P}^k(\mathcal{T}_h) \times \mathcal{P}^k(\mathcal{T}_h) \to \mathbb{R}$ such that, for all $(w_h, v_h) \in \mathcal{P}^k(\mathcal{T}_h) \times \mathcal{P}^k(\mathcal{T}_h)$,

$$a_h(w_h, v_h) \coloneqq \int_\Omega \sigma(G_h^k w_h) \cdot G_h^k v_h + s_h(w_h, v_h), \tag{20}$$

where

$$s_h(w_h, v_h) \coloneqq \sum_{F\in\mathcal{F}_h} h_F^{1-p} \int_F \sigma_1([w_h]_F)[v_h]_F = \sum_{F\in\mathcal{F}_h} h_F^{1-p} \int_F |[w_h]_F|^{p-2}[w_h]_F[v_h]_F.$$

The discretization of the advection-reaction terms hinges on the bilinear form $b_h : \mathcal{P}^k(\mathcal{T}_h) \times \mathcal{P}^k(\mathcal{T}_h) \to \mathbb{R}$ such that, for all $(w_h, v_h) \in \mathcal{P}^k(\mathcal{T}_h) \times \mathcal{P}^k(\mathcal{T}_h)$,

$$\begin{aligned} b_h(w_h, v_h) = &- \int_\Omega w_h(\beta\cdot\nabla_h v_h) + \int_\Omega \mu w_h v_h + \sum_{F\in\mathcal{F}_h} \int_F (\beta\cdot n_F)\{w_h\}_F[v_h]_F \\ &+ \frac{1}{2}\sum_{F\in\mathcal{F}_h} \hat\beta_F \int_F [w_h]_F[v_h]_F, \end{aligned} \tag{21}$$

where, for all $F \in \mathcal{F}_h$, we have introduced the face reference velocity

$$\hat\beta_F \coloneqq \|\beta\cdot n_F\|_{L^\infty(F)}.$$

Notice that the stabilization term is not the classical upwind, but rather a stronger version based on the reinterpretation as jump penalty provided in [11].

Remark 6 (Generalizations). The bilinear form b_h includes suitable terms that will be used to control the diffusive and advection terms on advection-dominated faces. The above formulation (and, in many cases, also the theoretical results that follow) could be easily extended to other choices, such as including cross-wind or making $\hat\beta_F$ dependent on some computable estimate of the local Péclet number. In particular, one could switch to standard upwind stabilization on boundary faces to correctly treat boundary conditions in the vanishing diffusion case.

The discrete problem reads: Find $u_h \in \mathcal{P}^k(\mathcal{T}_h)$ such that

$$a_h(u_h, v_h) + b_h(u_h, v_h) = \int_\Omega f v_h \qquad \forall v_h \in \mathcal{P}^k(\mathcal{T}_h). \tag{22}$$

4.2 Main results

In this section we collect the main results of the analysis of problem (22). The error estimate accounts for the different regimes in each mesh element/face, as identified by local Péclet numbers (for a similar local approach in a different context, see, for instance, [17]).

4.2.1 Dimensionless numbers and reference quantities

In order to state these convergence results, we need to define here key reference quantities and dimensionless numbers. For any function $w \in W^{1,p}(\Omega)$ and any mesh element $T \in \mathcal{T}_h$, we define the element Péclet number as follows. If $\hat\beta_T \coloneqq \|\beta\|_{L^\infty(T)^d}$ vanishes, we set $\mathrm{Pe}_T(w) = 0$; otherwise,

$$\mathrm{Pe}_T(w) \coloneqq \frac{\hat\beta_T h_T}{\hat K_T(w)} \quad \text{with} \quad \hat K_T(w) \coloneqq \||\nabla w|^{p-2}\|_{L^\infty(\mathcal{T}_T)}, \tag{23}$$

with the convention that $\hat K_T(w) = +\infty$ (and thus $\mathrm{Pe}_T(w) = 0$) if the restriction of $|\nabla w|^{p-2}$ is not in $L^\infty(\omega_T)$. Furthermore, we define the reference time:

$$\hat\tau_T \coloneqq \frac{1}{\max(\|\mu\|_{L^\infty(T)}, |\beta|_{W^{1,\infty}(T)^d})}. \tag{24}$$

Similarly, for any $F \in \mathcal{F}_h$, we define the face Péclet number as follows. If $\hat{\beta}_F = 0$, we set $\mathrm{Pe}_F(w) = 0$; otherwise

$$\mathrm{Pe}_F(w) \coloneqq \frac{\hat{\beta}_F h_F}{\hat{K}_F(w)} \quad \text{with} \quad \hat{K}_F(w) \coloneqq \max\left(\||\nabla w|^{p-2}\|_{L^\infty(F)}, h_F^{2-p}\||[\pi_h^k w]_F|^{p-2}\|_{L^\infty(F)}\right), \tag{25}$$

where again $\hat{K}_F(w) = +\infty$ (and thus $\mathrm{Pe}_F(w) = 0$) whenever the involved functions are not in $L^\infty(F)$.

Notice that the face Péclet number accounts for the fact that the stabilization term introduces additional numerical diffusion. In practical situations, this numerical diffusion can be expected to be small compared to the physical one.

We partition the sets of mesh elements and faces based on the values of the local Péclet numbers. Specifically, given a smooth enough function $w : \Omega \to \mathbb{R}$, we set

$$\mathcal{T}_h^{\mathrm{a}}(w) \coloneqq \{T \in \mathcal{T}_h \mid \mathrm{Pe}_T(w) > 1\}, \qquad \mathcal{T}_h^{\mathrm{d}}(w) \coloneqq \mathcal{T}_h \setminus \mathcal{T}_h^{\mathrm{a}}(w),$$
$$\mathcal{F}_h^{\mathrm{a}}(w) \coloneqq \{F \in \mathcal{F}_h \mid \mathrm{Pe}_F(w) > 1\}, \qquad \mathcal{F}_h^{\mathrm{d}}(w) \coloneqq \mathcal{F}_h \setminus \mathcal{F}_h^{\mathrm{a}}(w).$$

4.2.2 Norms

The relevant norm for the analysis of the diffusion terms is $\|\cdot\|_{1,p,h}$ (cf. (12)) as well its restriction to an element $T \in \mathcal{T}_h$ (cf. (14)). The norm for the advective and reactive terms is, on the other hand, given by

$$\|v_h\|_{\beta,\mu,h} \coloneqq \left(\frac{1}{2}\sum_{F\in\mathcal{F}_h} \hat{\beta}_F\|[v_h]_F\|_{L^2(F)}^2 + \|\mu^{\frac{1}{2}} v_h\|_{L^2(\Omega)}^2\right)^{\frac{1}{2}} \qquad \forall v_h \in \mathcal{P}^k(\mathcal{T}_h). \tag{26}$$

This choice of advection-reaction norm is justified as follows. By standard arguments (which essentially amount to applying the integration by parts formula (8) with $(\tau, v) = (\beta w_h, v_h)$ to the first term in the right-hand side of (21), using the continuity of $\beta \cdot n_F$ across interfaces, and recalling that $\nabla \cdot \beta = 0$), it is easy to check that

$$b_h(v_h, v_h) = \|v_h\|_{\beta,\mu,h}^2 \qquad \forall v_h \in \mathcal{P}^k(\mathcal{T}_h), \tag{27}$$

showing that b_h is coercive with respect to the norm defined by (26) with coercivity constant equal to 1.

4.2.3 Error estimate

The following theorem contains an estimate of the error between the solution of the discrete problem (22) and the projection of the continuous solution that tracks the dependence of the convergence rate on the local regime. We remark that the regularity conditions required below for u are implied, for instance, by the simpler but less sharp requirement $w \in W^{1,p}(\Omega) \cap W^{r+1,\overline{p}}(\mathcal{T}_h)$ with $\overline{p} = \max\{2, 2p-2, p'\}$.

Theorem 7 (Convergence)**.** *Denote, respectively, by $u \in W^{1,p}(\Omega)$ and by $u_h \in \mathcal{P}^k(\mathcal{T}_h)$ the solutions of the weak formulation of problem* (1) *and of the discrete problem* (22)*. Additionally assume that, for some $r \in \{0, \dots, k\}$,*

- $u_{|\omega_T} \in W^{r+1,p}(\mathcal{T}_T)$ *for all* $T \in \mathcal{T}_h^{\mathrm{d}}(u)$*;*
- $u_{|T} \in H^{r+1}(T)$ *for all* $T \in \mathcal{T}_h$*;*
- $u_{|\omega_F} \in W^{r+1,p}(\mathcal{T}_F) \cap W^{r+1,p'}(\mathcal{T}_F)$ *and* $\sigma(\nabla u)_{|\omega_F} \in W^{r,p'}(\mathcal{T}_F)^d$ *for all* $F \in \mathcal{F}_h^{\mathrm{d}}(u)$*;*
- $u_{|\omega_F} \in H^{r+1}(\mathcal{T}_F)$ *and* $\sigma(\nabla u)_{|\omega_F} \in H^{r+\frac{1}{2}}(\mathcal{T}_F)^d$ *for all* $F \in \mathcal{F}_h^{\mathrm{a}}(u)$.

Then, letting

$$q \coloneqq \begin{cases} 2 & \text{if } p < 2, \\ p & \text{if } p \geq 2, \end{cases} \tag{28}$$

it holds

$$\begin{aligned} &\|u_h - \pi_h^k u\|_{1,p,h}^q + \|u_h - \pi_h^k u\|_{\beta,\mu,h}^2 \\ &\quad \lesssim \sum_{T \in \mathcal{T}_h} \hat{\tau}_T^{-2} \underline{\mu}_T^{-1} h_T^{2(r+1)} |u|_{H^{r+1}(T)}^2 \\ &\qquad + \sum_{T \in \mathcal{T}_h^{\mathrm{a}}(u)} \hat{\beta}_T h_T^{2r+1} |u|_{H^{r+1}(T)}^2 + \sum_{T \in \mathcal{T}_h^{\mathrm{d}}(u)} \begin{cases} h_T^{rp} |u|_{W^{r+1,p}(\mathcal{T}_T)}^p & \text{if } p < 2 \\ h_T^{2r} |u|_{W^{r+1,p}(\mathcal{T}_T)}^2 & \text{if } p \geq 2 \end{cases} \\ &\qquad + \sum_{F \in \mathcal{F}_h^{\mathrm{a}}(u)} h_F^{2r+1} \left(\hat{K}_F(u)^{-1} |\sigma(\nabla u)|_{H^{r+\frac{1}{2}}(\mathcal{T}_F)^d}^2 + \hat{\beta}_F |u|_{H^{r+1}(\mathcal{T}_F)}^2 \right) + \sum_{F \in \mathcal{F}_h^{\mathrm{d}}(u)} h_F^{rp} |u|_{W^{r+1,p}(\mathcal{T}_F)}^p \\ &\qquad + \left[\sum_{F \in \mathcal{F}_h^{\mathrm{d}}(u)} h_F^{rp'} \left(|\sigma(\nabla u)|_{W^{r,p'}(\mathcal{T}_F)^d}^{p'} + \hat{K}_F(u)^{p'} |u|_{W^{r+1,p'}(\mathcal{T}_F)}^{p'} \right) \right]^{\frac{q'}{p'}}. \end{aligned} \tag{29}$$

Proof. See Section 5.3. □

The above convergence result is fully local, being able to deliver sharp estimates also in situations where diffusion or advection dominate in different areas of the domain. This feature is particularly important in the present nonlinear situation, where the distinction among the two cases depends on the solution itself and not only on some data given a priori. Notice that, for the sake of conciseness, here we do not consider the trivial case of dominating reaction.

For the more interesting case $p < 2$, the above estimates are "optimal" in the sense that, for regular solutions, the bound yields the same asymptotic order of convergence as for conforming Finite Element (FE) schemes, i.e., $O(h^{\frac{rp}{2}})$ [5, 25]. Furthermore, in the pre-asymptotic regime, our estimate underlines a better error reduction rate in the areas of the domain where advection dominates (behaving as $h^{r+\frac{1}{2}}$). In this respect, note that the negative power of $\hat{K}_F$ appearing in the bound above is balanced by the associated σ term, see Remark 10. The case $p = 2$ corresponds to a linear diffusion-advection-reaction problem, for which classical estimates are recovered (see, e.g., [3, 22] and also [21, Section 4.6]). In the case $p > 2$, the same observations apply, except for the fact that the asymptotic convergence rate now compares unfavorably to the conforming FE case, due to the presence of an $O(h^{rp'})$ term in the right hand side (to be compared with $O(h^{2r})$). This aspect could be possibly improved by introducing a stronger jump term s_h (which, on the other, hand would lead to a weaker pre-asymptotic reduction rate in advection dominated regimes) or by introducing some suitable tweaks in the analysis, see Remark 12.

5 Theoretical analysis

5.1 Properties of the diffusion function

Lemma 8 (Stability of a_h)**.** *For any $w_h, v_h \in \mathcal{P}^k(\mathcal{T}_h)$, recalling the definition* (28) *of q and assuming that $\|w_h\|_{1,p,h} + \|v_h\|_{1,p,h} \lesssim 1$ if $p < 2$, there exists C_a independent of h (but possibly depending on Ω, p, and the mesh regularity parameter) such that*

$$C_a \|w_h - v_h\|_{1,p,h}^q \lesssim a_h(w_h, w_h - v_h) - a_h(v_h, w_h - v_h). \tag{30}$$

Proof. The proof is a straightforward adaptation of the monotonicity properties of σ and the arguments of [18, Point (ii) of Theorem 6.19]. □

We start by estimating the error stemming from the diffusion term.

Lemma 9 (Estimate of the discrete diffusion error)**.** *Let $w \in W^{1,p}(\Omega)$ be such that $\sigma(\nabla w) \in W^{1,p'}(\mathcal{T}_h)^d$ and $\nabla\cdot\sigma(\nabla w) \in L^{p'}(\Omega)$. Let's define the diffusion error linear form $\mathcal{E}_{a,h}^k : \mathcal{P}^k(\mathcal{T}_h) \to \mathbb{R}$ such that, for all $v_h \in \mathcal{P}^k(\mathcal{T}_h)$,*

$$\mathcal{E}_{a,h}^k(w; v_h) \coloneqq -\int_\Omega \nabla\cdot\sigma(\nabla w)\, v_h - a_h(\pi_h^k w, v_h). \tag{31}$$

Additionally assume that, for some $r \in \{0,\dots,k\}$,

- *$w_{|\omega_T} \in W^{r+1,p}(\mathcal{T}_T)$ for all $T \in \mathcal{T}_h^{\mathrm{d}}(w)$;*
- *$w_{|T} \in H^{r+1}(T)$ for all $T \in \mathcal{T}_h^{\mathrm{a}}(w)$;*
- *$w_{|\omega_F} \in W^{r+1,p}(\mathcal{T}_F)$ and $\sigma(\nabla w)_{|\omega_F} \in W^{r,p'}(\mathcal{T}_F)^d$ for all $F \in \mathcal{F}_h^{\mathrm{d}}(w)$;*
- *$w_{|\omega_F} \in H^{r+1}(\mathcal{T}_F)$ and $\sigma(\nabla w)_{|\omega_F} \in H^{r+\frac12}(\mathcal{T}_F)^d$ for all $F \in \mathcal{F}_h^{\mathrm{a}}(w)$.*

Then, recalling (28)*, it holds, for any $w_h \in \mathcal{P}^k(\mathcal{T}_h)$ and any real number $\delta > 0$,*

$$\begin{aligned}
&\mathcal{E}_{a,h}^k(w; w_h - \pi_h^k w)\\
&\quad\le \delta\left(a_h(w_h, w_h-\pi_h^k w) - a_h(\pi_h^k w, w_h - \pi_h^k w) + |w_h-\pi_h^k w|_{1,p,h}^q + \|w_h-\pi_h^k w\|_{\beta,\mu,h}^2\right)\\
&\qquad + c(\delta)\left(\sum_{T\in\mathcal{T}_h^{\mathrm{a}}(w)} \hat\beta_T h_T^{2r+1}|w|^2_{H^{r+1}(T)} + \sum_{T\in\mathcal{T}_h^{\mathrm{d}}(w)} \begin{cases} h_T^{rp}|w|^p_{W^{r+1,p}(\mathcal{T}_T)} & \text{if } p<2\\ h_T^{2r}|w|^2_{W^{r+1,p}(\mathcal{T}_T)} & \text{if } p\ge 2\end{cases}\right)\\
&\qquad + c(\delta)\sum_{F\in\mathcal{F}_h^{\mathrm{a}}(w)} \hat K_F^{-1}(w) h_F^{2r+1}|\sigma(\nabla w)|^2_{H^{r+\frac12}(\mathcal{T}_F)^d} + c(\delta)\left(\sum_{F\in\mathcal{F}_h^{\mathrm{d}}(w)} h_F^{rp'}|\sigma(\nabla w)|^{p'}_{W^{r,p'}(\mathcal{T}_F)}\right)^{\frac{q'}{p'}}\\
&\qquad + c(\delta)\left(\sum_{F\in\mathcal{F}_h^{\mathrm{a}}(w)} \hat\beta_F h_F^{2r+1}|w|^2_{H^{r+1}(\mathcal{T}_F)} + \sum_{F\in\mathcal{F}_h^{\mathrm{d}}(w)} h_F^{rp}|w|^p_{W^{r+1,p}(\mathcal{T}_F)}\right),
\end{aligned} \tag{32}$$

with $c(\delta)$ independent of the particular mesh in $\{\mathcal{T}_h\}_h$ and the function w.

Proof. Let, for the sake of brevity,

$$v_h \coloneqq w_h - \pi_h^k w.$$

Using the integration by parts formula (8) for the first term in the right-hand side of (31) along with the fact that $[\sigma(\nabla w)]_F\cdot n_F$ vanishes for all $F\in\mathcal{F}_h^{\mathrm{i}}$ (which expresses the continuity of normal fluxes), expanding a_h according to its definition (20), adding $0 = \int_\Omega \pi_h^k\sigma(\nabla w)\cdot R_h^k v_h - \sum_{F\in\mathcal{F}_h}\int_F \{\pi_h^k\sigma(\nabla w)\}_F\cdot n_F\,[v_h]_F$ (cf. (10) and (9)), and adding and subtracting $\int_\Omega \sigma(\nabla w)\cdot R_h^k v_h$, we arrive at the following decomposition of the error:

$$\begin{aligned}
&\mathcal{E}_{a,h}^k(w; v_h)\\
&\quad= \underbrace{\int_\Omega \left[\sigma(G_h^k\pi_h^k w) - \sigma(\nabla w)\right]\cdot G_h^k(\pi_h^k w - w_h)}_{\mathfrak{T}_1} \underbrace{- \sum_{F\in\mathcal{F}_h}\int_F \{\sigma(\nabla w) - \pi_h^k\sigma(\nabla w)\}_F\cdot n_F\,[v_h]_F}_{\mathfrak{T}_2}\\
&\qquad + \cancel{\int_\Omega \left[\sigma(\nabla w) - \pi_h^k\sigma(\nabla w)\right]\cdot R_h^k v_h} + \underbrace{s_h(\pi_h^k w, \pi_h^k w - w_h)}_{\mathfrak{T}_3},
\end{aligned} \tag{33}$$

where the cancellation follows from the definition of π_h^k after recalling that $R_h^k v_h \in \mathcal{P}^k(\mathcal{T}_h)^d$. We next proceed to estimate the other terms in the right-hand side.

Estimate of $\mathfrak{T}_1$. For the first term, we start by writing $\mathfrak{T}_1 = \sum_{T\in\mathcal{T}_h} \mathfrak{T}_1(T)$ and consider a single $T \in \mathcal{T}_h$. Using the bound (2) with $n = d$ and $(x, y, z) = (G_h^k\pi_h^k w, G_h^k w_h, \nabla w)$ and recalling that $v_h = w_h - \pi_h^k w$, we obtain

$$\mathfrak{T}_1(T) \le \delta\left(\int_T \sigma(G_h^k w_h)\cdot G_h^k v_h - \int_T \sigma(G_h^k\pi_h^k w)\cdot G_h^k v_h\right) + c(\delta)\mathfrak{T}_{1,\mathrm{err}}(T), \tag{34}$$

where

$$\mathfrak{T}_{1,\mathrm{err}}(T) \coloneqq \int_T (|\nabla w| + |G_h^k\pi_h^k w|)^{p-2}|\nabla w - G_h^k\pi_h^k w|^2.$$

We now distinguish between diffusion-dominated and advection-dominated elements of the mesh to estimate $\mathfrak{T}_{1,\mathrm{err}}(T)$.

Let first $T \in \mathcal{T}_h^{\mathrm{d}}(w)$. In the case $p < 2$, we use the fact that $|\nabla w - G_h^k\pi_h^k w| \le |\nabla w| + |G_h^k\pi_h^k w|$ almost everywhere in T along with the fact that $\mathbb{R}^+ \ni x \mapsto x^{p-2} \in \mathbb{R}$ is strictly decreasing to write

$$\mathfrak{T}_{1,\mathrm{err}}(T) \le \int_T |\nabla w - G_h^k\pi_h^k w|^p = \|\nabla w - G_h^k\pi_h^k w\|_{L^p(T)^d}^p \overset{(18)}{\lesssim} h_T^{rp}|w|_{W^{r+1,p}(\mathcal{T}_T)}^p. \tag{35}$$

In the case $p \ge 2$, on the other hand, we apply a Hölder inequality with exponents $\left(\frac{p}{p-2}, \frac{p}{2}\right)$ and a triangle inequality to write

$$\begin{aligned}\mathfrak{T}_{1,\mathrm{err}}(T) &\lesssim \left(\|\nabla w\|_{L^p(T)^d} + \|G_h^k\pi_h^k w\|_{L^p(T)^d}\right)^{p-2}\|\nabla w - G_h^k\pi_h^k w\|_{L^p(T)^d}^2\\ &\overset{(19),(18)}{\lesssim} \|\nabla w\|_{L^p(\mathcal{T}_T)^d}^{p-2} h_T^{2r}|w|_{W^{r+1,p}(\mathcal{T}_T)}^2 \lesssim h_T^{2r}|w|_{W^{r+1,p}(\mathcal{T}_T)}^2,\end{aligned} \tag{36}$$

where the conclusion follows from the assumption $\|\nabla w\|_{L^p(\Omega)^d} \lesssim 1$.

Let now $T \in \mathcal{T}_h^{\mathrm{a}}(w)$. We first consider the case $p < 2$. Using again the fact that $\mathbb{R}^+ \ni x \mapsto x^{p-2} \in \mathbb{R}$ is stricly decreasing, then applying a Hölder inequality with exponents $(\infty, 1)$ and using the approximation properties (18) of $G_h^k \circ \pi_h^k$, and finally recalling that $\mathrm{Pe}_T(w) > 1$ (cf. (23) for its definition), we have

$$\begin{aligned}\mathfrak{T}_{1,\mathrm{err}}(T) &\lesssim \int_T |\nabla w|^{p-2}|\nabla w - G_h^k\pi_h^k w|^2 \lesssim \||\nabla w|^{p-2}\|_{L^\infty(T)}\,\|\nabla w - G_h^k\pi_h^k w\|_{L^2(T)^d}^2\\ &\lesssim \hat{K}_T(w)h_T^{2r}|w|_{H^{r+1}(T)}^2 \le \hat{\beta}_T h_T^{2r+1}|w|_{H^{r+1}(T)}^2.\end{aligned} \tag{37}$$

In the case $p \ge 2$, on the other hand, the local boundedness (19) of $G_h^k \circ \pi_h^k$ with $q = \infty$ along with the definition (23) of $\hat{K}_T(w)$ easily leads to

$$\mathfrak{T}_{1,\mathrm{err}}(T) \lesssim \hat{K}_T(w)h_T^{2r}|w|_{H^{r+1}(T)}^2 \lesssim \hat{\beta}_T h_T^{2r+1}|w|_{H^{r+1}(T)}^2, \tag{38}$$

where the conclusion follows again using $\mathrm{Pe}_T(w) > 1$.

Plugging the estimates (35), (36), (37), and (38) into (34), we arrive at

$$\begin{aligned}\mathfrak{T}_1 \le{}& \delta\left(\int_\Omega \sigma(G_h^k w_h)\cdot G_h^k v_h - \int_\Omega \sigma(G_h^k\pi_h^k w)\cdot G_h^k v_h\right)\\ &+ c(\delta)\sum_{T\in\mathcal{T}_h^{\mathrm{a}}(w)} \hat{\beta}_T h_T^{2r+1}|w|_{H^{r+1}(T)}^2 + c(\delta)\sum_{T\in\mathcal{T}_h^{\mathrm{d}}(w)} \begin{cases} h_T^{rp}|w|_{W^{r+1,p}(\mathcal{T}_T)}^p & \text{if } p < 2,\\ h_T^{2r}|w|_{W^{r+1,p}(\mathcal{T}_T)}^2 & \text{if } p \ge 2.\end{cases}\end{aligned} \tag{39}$$

Estimate of $\mathfrak{T}_2$. For the second term, we write $\mathfrak{T}_2 = \sum_{F\in\mathcal{F}_h} \mathfrak{T}_2(F)$ and, for all $F \in \mathcal{F}_h^{\mathrm{d}}(w)$, we estimate $\mathfrak{T}_2(F)$ as follows:

$$\begin{aligned}\mathfrak{T}_2(F) &\le \|\{\sigma(\nabla w) - \pi_h^k\sigma(\nabla w)\}_F\|_{L^{p'}(F)^d}\, \|[v_h]_F\|_{L^p(F)}\\ &\lesssim h_F^{r-\frac{1}{p'}+\frac{p-1}{p}} |\sigma(\nabla w)|_{W^{r,p'}(\mathcal{T}_F)}\, h_F^{\frac{1-p}{p}} \|[v_h]_F\|_{L^p(F)}\\ &\overset{(7)}{=} h_F^r |\sigma(\nabla w)|_{W^{r,p'}(\mathcal{T}_F)}\, h_F^{\frac{1-p}{p}} \|[v_h]_F\|_{L^p(F)},\end{aligned} \tag{40}$$

where we have used a triangle inequality along with the approximation properties of the L^2-orthogonal projector to treat the first factor in the passage to the second line.

For $F \in \mathcal{F}_h^{\mathrm{a}}(w)$, on the other hand, we first notice that $\hat\beta_F \neq 0$ and then use a Cauchy–Schwarz inequality to write

$$\begin{aligned}\mathfrak{T}_2(F) &\le \hat\beta_F^{-\frac12}\|\{\sigma(\nabla w) - \pi_h^k\sigma(\nabla w)\}_F\|_{L^2(F)^d}\, \hat\beta_F^{\frac12}\|[v_h]_F\|_{L^2(F)}\\ &\lesssim \hat\beta_F^{-\frac12} h_F^r |\sigma(\nabla w)|_{H^{r+\frac12}(\mathcal{T}_F)^d}\, \hat\beta_F^{\frac12}\|[v_h]_F\|_{L^2(F)}\\ &\lesssim \hat K_F^{-\frac12}(w) h_F^{r+\frac12} |\sigma(\nabla w)|_{H^{r+\frac12}(\mathcal{T}_F)^d}\, \hat\beta_F^{\frac12}\|[v_h]_F\|_{L^2(F)},\end{aligned} \tag{41}$$

where we have used the fact that $\mathrm{Pe}_F(w) > 1$ to conclude.

Gathering the above bounds and applying a Hölder inequality with exponents (p', p) on the sum over $F \in \mathcal{F}_h^{\mathrm{d}}(w)$, a Cauchy–Schwarz inequality on the sum over $F \in \mathcal{F}_h^{\mathrm{a}}(w)$, and using a generalized Young inequality with exponents (q', q), we get

$$\begin{aligned}\mathfrak{T}_2 \le{}& \delta\left(|v_h|_{1,p,h}^q + \|v_h\|_{\beta,\mu,h}^2\right) + c(\delta)\sum_{F\in\mathcal{F}_h^{\mathrm{a}}(w)} \hat K_F^{-1}(w) h_F^{2r+1} |\sigma(\nabla w)|^2_{H^{r+\frac12}(\mathcal{T}_F)^d}\\ &+ c(\delta)\left(\sum_{F\in\mathcal{F}_h^{\mathrm{d}}(w)} h_F^{rp'} |\sigma(\nabla w)|^{p'}_{W^{r,p'}(\mathcal{T}_F)}\right)^{\frac{q'}{p'}}.\end{aligned} \tag{42}$$

Estimate of $\mathfrak{T}_3$. Finally, for the third term, we write again $\mathfrak{T}_3 = \sum_{F\in\mathcal{F}_h} \mathfrak{T}_3(F)$. We then first recall that $[w]_F = 0$ and then apply (2) with $n = 1$ and $(x, y, z) = ([\pi_k^k w]_F, [w_h]_F, [w]_F)$, additionally using the fact that $v_h = w_h - \pi_h^k w$; we obtain that, for all positive δ,

$$\begin{aligned}\mathfrak{T}_3(F) &= h_F^{1-p}\int_F (\sigma_1([\pi_h^k w]_F) - \sigma_1([w]_F))[\pi_h^k w - w_h]_F\\ &\le \delta\left(h_F^{1-p}\int_F \sigma_1([w_h]_F)[v_h]_F - h_F^{1-p}\int_F \sigma_1([\pi_h^k w]_F)[v_h]_F\right) + c(\delta)\mathfrak{T}_{3,\mathrm{err}}(F),\end{aligned}$$

with

$$\mathfrak{T}_{3,\mathrm{err}}(F) \coloneqq h_F^{1-p}\int_F (|[\pi_h^k w]_F| + |[w]_F|)^{p-2}|[w - \pi_h^k w]_F|^2.$$

For $F \in \mathcal{F}_h^{\mathrm{d}}(w)$, this term is bounded trivially recalling that $[w]_F = 0$:

$$\mathfrak{T}_{3,\mathrm{err}}(F) = h_F^{1-p}\int_F |[\pi_h^k w - w]_F|^p = h_F^{1-p}\|[\pi_h^k w - w]_F\|^p_{L^p(F)} \lesssim h_F^{rp}|w|^p_{W^{r+1,p}(\mathcal{T}_F)}.$$

For $F \in \mathcal{F}_h^{\mathrm{a}}(w)$, on the other hand, recalling again that $[w]_F = 0$ and using a Hölder inequality with exponents $(\infty, 1)$, we have

$$\begin{aligned}\mathfrak{T}_{3,\mathrm{err}}(F) &\le h_F^{1-p} \| |[\pi_h^k w]_F|^{p-2} \|_{L^\infty(F)} \| [\pi_h^k w - w]_F \|_{L^2(F)}^2 \\ &\overset{(25)}{\le} \frac{\hat{K}_F(w)}{h_F} \| [\pi_h^k w - w]_F \|_{L^2(F)}^2 \lesssim \hat{\beta}_F h_F^{2r+1} |w|_{H^{r+1}(\mathcal{T}_F)}^2,\end{aligned}$$

where the conclusion follows using the fact that $\mathrm{Pe}_F^{-1}(w) < 1$ for the first factor and a triangle inequality followed by the approximation properties of π_h^k for the second.

Gathering the above estimates, we arrive at the following bound for $\mathfrak{T}_3$:

$$\begin{aligned}\mathfrak{T}_3 \le{}& \delta \left(s_h(w_h, v_h) - s_h(\pi_h^k w, v_h) \right) \\ &+ c(\delta) \sum_{F \in \mathcal{F}_h^{\mathrm{a}}(w)} \hat{\beta}_F h_F^{2r+1} |w|_{H^{r+1}(\mathcal{T}_F)}^2 + c(\delta) \sum_{F \in \mathcal{F}_h^{\mathrm{d}}(w)} h_F^{rp} |w|_{W^{r+1,p}(\mathcal{T}_F)}^p.\end{aligned} \tag{43}$$

Conclusion. Plugging (39), (42), and (43) into (33) and recalling that, in each of these estimates, $\delta > 0$ is arbitrary, the conclusion follows. □

Remark 10 (Negative power of $\hat{K}_F$). As already mentioned, the negative power of $\hat{K}_F$ appearing in bound (32) is balanced by the associated σ regularity term. The $\hat{K}_F^{-1}$ stems from equation (41). The fact that the associated term $\mathfrak{T}_2(F)$, $F \in \mathcal{F}_h^{\mathrm{a}}(w)$, cannot lead to an arbitrarily large contribution to the error becomes clear by bounding such term as in (40) instead of (41) (that is, using the diffusive part of the norm instead of the advective one). Here, we decided to use (41) in order to clearly underline the faster pre-asymptotic reduction rate occurring in advection dominated cases.

5.2 Properties of the advection-reaction bilinear form

We now estimate the error stemming from the advection component of the equation.

Lemma 11 (Estimate of the discrete advection-reaction error)**.** *Let $w \in W^{1,p}(\Omega)$ and define the advection-reaction error linear form $\mathcal{E}_{b,h}^k(w; v_h) : \mathcal{P}^k(\mathcal{T}_h) \to \mathbb{R}$ such that, for all $v_h \in \mathcal{P}^k(\mathcal{T}_h)$,*

$$\mathcal{E}_{b,h}^k(w; v_h) \coloneqq \int_\Omega \nabla \cdot (\beta w) v_h + \int_\Omega \mu w v_h - b_h(\pi_h^k w, v_h). \tag{44}$$

Additionally assume that $w \in H^{r+1}(\mathcal{T}_h)$ and $w_{|\omega_F} \in W^{r+1,p'}(\mathcal{T}_F)$ for all $F \in \mathcal{F}_h^{\mathrm{d}}(w)$ for some $r \in \{0, \dots, k\}$. Then, with q as in (28)*, it holds, for any $w_h \in \mathcal{P}^k(\mathcal{T}_h)$ and any real number $\delta > 0$,*

$$\begin{aligned}\mathcal{E}_{b,h}^k(w; w_h - \pi_h^k w) \le{}& \delta \left(|w_h - \pi_h^k w|_{1,p,h}^q + \|w_h - \pi_h^k w\|_{\beta,\mu,h}^2 \right) \\ &+ c(\delta) \left(\sum_{T \in \mathcal{T}_h} \hat{\tau}_T^{-2} \underline{\mu}_T^{-1} h_T^{2(r+1)} |w|_{H^{r+1}(T)}^2 + \sum_{F \in \mathcal{F}_h^{\mathrm{a}}(w)} \hat{\beta}_F h_F^{2r+1} |w|_{H^{r+1}(\mathcal{T}_F)}^2 \right) \\ &+ c(\delta) \left(\sum_{F \in \mathcal{F}_h^{\mathrm{d}}(w)} \hat{K}_F(w)^{p'} h_F^{rp'} |w|_{W^{r+1,p'}(\mathcal{T}_F)}^{p'} \right)^{\frac{q'}{p'}},\end{aligned} \tag{45}$$

with $c(\delta)$ independent of the particular mesh in $\{\mathcal{T}_h\}_h$ and the function w.

Proof. We set again, for the sake of brevity,

$$v_h \coloneqq w_h - \pi_h^k w.$$

Using (8) with $(\tau, v) = (\beta w, v_h)$ to integrate by parts the first term in the right-hand side of (44) along with $\nabla \cdot \beta = 0$, recalling the single-valuedness of $\beta \cdot n_F$ and $(\beta \cdot n_F)w$ across any interface $F \in \mathcal{F}_h^{\mathrm{i}}$, and inserting w into the jump operator after noticing that this quantity is single-valued across interfaces and it vanishes on boundary faces, we arrive at the following decomposition of the error:

$$\begin{aligned}\mathcal{E}_{b,h}^k(w; v_h) &= -\int_\Omega (w - \pi_h^k w)(\beta \cdot \nabla_h v_h) + \int_\Omega \mu (w - \pi_h^k w) v_h \\ &\quad + \sum_{F \in \mathcal{F}_h} \int_F (\beta \cdot n_F)\{w - \pi_h^k w\}_F [v_h]_F + \frac{1}{2} \sum_{F \in \mathcal{F}_h} \hat{\beta}_F \int_F [w - \pi_h^k w]_F [v_h]_F. \\ &=: \mathfrak{T}_1 + \cdots + \mathfrak{T}_4.\end{aligned} \tag{46}$$

We proceed to estimate the terms in the right-hand side.

Estimate of $\mathfrak{T}_1$. For the first term, we use the definition of π_T^k along with the fact that $\pi_T^0 \beta \cdot \nabla v_T \in \mathcal{P}^{k-1}(T) \subset \mathcal{P}^k(T)$ to write

$$\begin{aligned}\mathfrak{T}_1 &= \sum_{T \in \mathcal{T}_h} \int_T (w - \pi_T^k w)[(\beta - \pi_T^0 \beta) \cdot \nabla v_T] \\ &\le \sum_{T \in \mathcal{T}_h} \|w - \pi_T^k w\|_{L^2(T)} \|\beta - \pi_T^0 \beta\|_{L^\infty(T)^d} \|\nabla v_T\|_{L^2(T)^d} \\ &\lesssim \sum_{T \in \mathcal{T}_h} h_T^{r+1} |w|_{H^{r+1}(T)}\ h_T |\beta|_{W^{1,\infty}(T)^d}\ h_T^{-1} \|v_T\|_{L^2(T)} \\ &\le \left(\sum_{T \in \mathcal{T}_h} \hat{\tau}_T^{-2} \underline{\mu}_T^{-1} h_T^{2(r+1)} |w|_{H^{r+1}(T)}^2 \right)^{\frac{1}{2}} \|v_h\|_{\beta,\mu,h} \\ &\le \delta \|v_h\|_{\beta,\mu,h}^2 + c(\delta) \sum_{T \in \mathcal{T}_h} \hat{\tau}_T^{-2} \underline{\mu}_T^{-1} h_T^{2(r+1)} |w|_{H^{r+1}(T)}^2,\end{aligned} \tag{47}$$

where we have used a Hölder inequality with exponents $(2, \infty, 2)$ to pass to the second line, the approximation properties of the L^2-orthogonal projector along with a discrete inverse inequality to pass to the third line, a discrete Cauchy–Schwarz inequality on the sum over $T \in \mathcal{T}_h$ along with the definition (24) of the reference time to pass to the fourth line, and a generalized Young inequality to conclude.

Estimate of $\mathfrak{T}_2$. For the second term, a Hölder inequality with exponents $(\infty, 2, 2)$ and the approximation properties of the L^2-orthogonal projector readily give

$$\begin{aligned}\mathfrak{T}_2 &\lesssim \sum_{T \in \mathcal{T}_h} \|\mu\|_{L^\infty(T)}^{\frac{1}{2}} h_T^{r+1} |w|_{H^{r+1}(T)}\ \|\mu^{\frac{1}{2}} v_T\|_{L^2(T)} \\ &\overset{(24)}{\le} \sum_{T \in \mathcal{T}_h} \hat{\tau}_T^{-\frac{1}{2}} h_T^{r+1} |w|_{H^{r+1}(T)}\ \|\mu^{\frac{1}{2}} v_T\|_{L^2(T)} \\ &\le \left(\sum_{T \in \mathcal{T}_h} \hat{\tau}_T^{-2} \underline{\mu}_T^{-1} h_T^{2(r+1)} |w|_{H^{r+1}(T)}^2 \right)^{\frac{1}{2}} \|v_h\|_{\beta,\mu,h} \\ &\le \delta \|v_h\|_{\beta,\mu,h}^2 + c(\delta) \sum_{T \in \mathcal{T}_h} \hat{\tau}_T^{-2} \underline{\mu}_T^{-1} h_T^{2(r+1)} |w|_{H^{r+1}(T)}^2,\end{aligned} \tag{48}$$

where we have used a discrete Cauchy–Schwarz inequality on the sum over $T \in \mathcal{T}_h$, noticed that $\hat{\tau}_T^{-1} \le \hat{\tau}_T^{-2}\underline{\mu}_T^{-1}$, recalled the definition (26) of the advection-reaction norm in the third inequality, and used a generalized Young inequality to conclude.

Estimate of $\mathfrak{T}_3+\mathfrak{T}_4$. We next write $\mathfrak{T}_3+\mathfrak{T}_4 = \sum_{F\in\mathcal{F}_h}\mathfrak{T}_{3+4}(F)$ and estimate separately the local contribution on diffusion- and advection-dominated faces.

For all $F \in \mathcal{F}_h^{\mathrm{d}}(w)$, we write

$$\begin{aligned}
|\mathfrak{T}_{3+4}(F)| &\lesssim \hat{\beta}_F\left(\|\{w-\pi_h^k w\}_F\|_{L^{p'}(F)} + \|[w-\pi_h^k w]_F\|_{L^{p'}(F)}\right)\|[v_h]_F\|_{L^p(F)}\\
&\lesssim \hat{\beta}_F h_F^{r+1-\frac{1}{p'}}|w|_{W^{r+1,p'}(\mathcal{T}_F)}\, h_F^{\frac{1}{p'}} h_F^{\frac{1-p}{p}}\|[v_h]_F\|_{L^p(F)}\\
&= \hat{K}_F(w)\mathrm{Pe}_F(w)h_F^r|w|_{W^{r+1,p'}(\mathcal{T}_F)}\, h_F^{\frac{1-p}{p}}\|[v_h]_F\|_{L^p(F)}\\
&\le \hat{K}_F(w)h_F^r|w|_{W^{r+1,p'}(\mathcal{T}_F)}\, h_F^{\frac{1-p}{p}}\|[v_h]_F\|_{L^p(F)},
\end{aligned}$$

where we have used a Hölder inequality with exponents (∞,p',p) in the first inequality, triangle inequalities followed by the trace approximation properties of the L^2-orthogonal projector along with (7) to write $1 = h_F^{\frac{1}{p'}}h_F^{\frac{1-p}{p}}$ in the second inequality, the definition (25) of the local Péclet number in the equality, and the fact that $\mathrm{Pe}_F(w)\le 1$ to conclude.

For all $F\in\mathcal{F}_h^{\mathrm{a}}(w)$, on the other hand, the estimate is

$$\begin{aligned}
|\mathfrak{T}_{3+4}(F)| &\lesssim \hat{\beta}_F^{\frac12}\left(\|\{w-\pi_h^k w\}_F\|_{L^2(F)} + \|[w-\pi_h^k w]_F\|_{L^2(F)}\right)\hat{\beta}_F^{\frac12}\|[v_h]_F\|_{L^2(F)}\\
&\lesssim \hat{\beta}_F^{\frac12}h_F^{r+\frac12}|w|_{H^{r+1}(\mathcal{T}_F)}\,\hat{\beta}_F^{\frac12}\|[v_h]_F\|_{L^2(F)},
\end{aligned}\tag{49}$$

where we have used a Hölder inequality with exponents $(\infty,2,2)$ in the first inequality and triangle inequalities followed by the approximation properties of the L^2-orthogonal projector in the second inequality. After applying a discrete Hölder inequality with exponents (p',p) on the sum over diffusive faces, a discrete Cauchy–Schwarz inequality on the sum over advective faces, and using generalized Young inequalities, we arrive at

$$\begin{aligned}
|\mathfrak{T}_3+\mathfrak{T}_4| \le{}& \delta\left(|v_h|_{1,p,h}^q + \|v_h\|_{\beta,\mu,h}^2\right) + c(\delta)\sum_{F\in\mathcal{F}_h^{\mathrm{a}}(w)}\hat{\beta}_F h_F^{2r+1}|w|^2_{H^{r+1}(\mathcal{T}_F)}\\
&+ c(\delta)\left(\sum_{F\in\mathcal{F}_h^{\mathrm{d}}(w)}\hat{K}_F(w)^{p'}h_F^{rp'}|w|^{p'}_{W^{r+1,p'}(\mathcal{T}_F)}\right)^{\frac{q'}{p'}}.
\end{aligned}\tag{50}$$

Conclusion. Plugging (47), (48), and (50) into (46) and recalling that, in each of these estimates, $\delta>0$ is arbitrary, the conclusion follows. □

Remark 12 (Comparison with conforming finite elements). As already discussed at the end of Section 4.2, for $p>2$ the bound (29) compares unfavorably with the conforming FE case in diffusion dominated cases. The reason are the terms $\mathfrak{T}_2$ for the diffusive part, c.f. (42), and $\mathfrak{T}_3+\mathfrak{T}_4$ for the advective part, c.f. (50), which behave as $O(h^{rp'})$ for diffusion dominated faces (instead of $O(h^{2r})$). One could slightly improve such bounds by the following observations. The polynomial approximation estimate in (40) can be pushed further, requiring a higher regularity $\sigma(\nabla w)\in W^{r+1,p'}(\mathcal{T}_F)$ but yielding a bound of order h_F^{r+1}. Furthermore, $\mathfrak{T}_3+\mathfrak{T}_4$ could be bounded using advection, as in (49), also in diffusion dominated cases, thus avoiding the $O(h^{rp'})$ term. Indeed note that, due to the presence of $\hat{\beta}_F$ in $\mathfrak{T}_{3+4}(F)$, the bound (49) does not need any assumption on dominant advection. The above modifications would lead to an $O(h^{(r+1)p'})$ right hand side for $p>2$ in diffusion dominated cases.

5.3 Proof of Theorem 7

We start by writing

$$
\begin{aligned}
C_a\|u_h-\pi_h^k u\|_{1,p,h}^q+\|u_h-\pi_h^k u\|_{\beta,\mu,h}^2 & \\
\overset{(30),(27)}{\leq} & a_h(u_h,u_h-\pi_h^k u)-a_h(\pi_h^k u,u_h-\pi_h^k u)+b_h(u_h-\pi_h^k u,u_h-\pi_h^k u) \quad (51)\\
\overset{(31),(44)}{=} & \mathcal{E}_{a,h}^k(u;u_h-\pi_h^k u)+\mathcal{E}_{b,h}^k(u;u_h-\pi_h^k u)=:\mathcal{E}_h^k(u;u_h-\pi_h^k u)\,,
\end{aligned}
$$

where we also used (1) and (22) in order to derive the last identity. For any real number $\delta>0$, it holds

$$
\begin{aligned}
\mathcal{E}_h^k(u;u_h-\pi_h^k u) \overset{(32),(45),(27)}{\leq} & \delta\left(a_h(u_h,u_h-\pi_h^k u)-a_h(\pi_h^k u,u_h-\pi_h^k u)+b_h(u_h-\pi_h^k u,u_h-\pi_h^k u)\right)\\
& +\delta\left(2|u_h-\pi_h^k u|_{1,p,h}^q+\|u_h-\pi_h^k u\|_{\beta,\mu,h}^2\right)+c(\delta)E_h^k\\
\leq & \,\delta\mathcal{E}_h^k(u;u_h-\pi_h^k u)+\delta\left(2|u_h-\pi_h^k u|_{1,p,h}^q+\|u_h-\pi_h^k u\|_{\beta,\mu,h}^2\right)+c(\delta)E_h^k,
\end{aligned}
$$

where $c(\delta)$ denotes the largest value between (32) and (45), while E_h^k gathers all the terms multiplied by $c(\delta)$ in the sum of the right-hand sides of (32) and (45). For any $\delta<1$, this gives

$$
\mathcal{E}_h^k(u;u_h-\pi_h^k u)\leq\frac{2\delta}{1-\delta}|u_h-\pi_h^k u|_{1,p,h}^q+\frac{\delta}{1-\delta}\|u_h-\pi_h^k u\|_{\beta,\mu,h}^2+\frac{c(\delta)}{1-\delta}E_h^k. \quad (52)
$$

Let now ϵ and δ_ϵ denote two real numbers such that $0<\epsilon<\frac{2}{2+C_a}$ and $0<\delta_\epsilon<\frac{1}{2}\min(1,\epsilon C_a)$. Plugging (52) with $\delta=\delta_\epsilon$ into (51), noticing that, by definition, $\frac{2\delta_\epsilon}{1-\delta_\epsilon}<\frac{\epsilon}{1-\delta_\epsilon}C_a$, rearranging, and multiplying the resulting inequality by $(1-\delta_\epsilon)$, we get

$$
(1-\delta_\epsilon-\epsilon)C_a\|u_h-\pi_h^k u\|_{1,p,h}^q+(1-2\delta_\epsilon)\|u_h-\pi_h^k u\|_{\beta,\mu,h}^2\leq c(\delta_\epsilon)E_h^k.
$$

We conclude noticing that, by definition of ϵ and δ_ϵ, $1-\delta_\epsilon-\epsilon>1-\frac{\epsilon}{2}C_a-\epsilon>0$, and $1-2\delta_\epsilon>0$.

6 Numerical tests

In this section, we investigate from the practical standpoint the error estimates derived in Theorem 7 through some numerical experiments. The computational domain for all the tests developed in this section is the standard unit square $\Omega=(0,1)^2$. In order to analyze the numerical convergence rate, we consider a family of five triangular meshes $\mathcal{T}_h$ with decreasing diameters, namely

$$
h\in\{0.4714,\ 0.2215,\ 0.1189,\ 0.0588,\ 0.0314\}.
$$

Starting from the coarsest mesh, the subsequent meshes are obtained by (approximately) halving the meshsize. Due to the nonlinearity of the problem for $p\neq 2$, we use a fixed-point strategy to compute the discrete solution. We set the maximum number of iterations to $N_{max}=500$, the tolerance for the relative residual error to $\varepsilon=10^{-10}$, and we take the initial guess as the discrete solution of the problem with $p=2$.

6.1 Example 1

In the present example we consider problem (1) with the following exact solution, velocity field and reaction terms

$$
u(x,y):=\sin(x+0.1)\cos(y+0.1),\quad \beta(x,y):=\begin{bmatrix}\sin(x)\cos(y)\\-\sin(y)\cos(x)\end{bmatrix},\quad \mu(x,y):=1\,.
$$

The problem is investigated for different values of the Sobolev index p, specifically for the following choices

$$p \in \{1.5, 1.75, 2, 2.5, 3\}.$$

The source term f and the nonhomogeneous Dirichlet boundary condition are taken in accordance with p, the above analytical solution, and the remaining terms in the equation.

Furthermore, we introduce a coefficient ν which multiplies the diffusive term $-\nabla \cdot \sigma(\nabla u)$ and allows to control the relative magnitude of the diffusion and advection terms. Specifically, we set $\nu = 1$ for a diffusion-dominated regime and $\nu = 10^{-4}$ for an advection-dominated regime.

We compute the discrete solution $u_h \in \mathcal{P}^k(\mathcal{T}_h)$ for $k \in \{1, 2, 3\}$, in both the diffusion-dominated and advection-dominated regimes, with the aim of analyzing the numerical behavior of the error quantity

$$\mathrm{ERR}_h := \left(\nu \|u - u_h\|_{1,p,h}^q + \|u - u_h\|_{\beta,\mu,h}^2\right)^{\frac{1}{2}}$$

with q defined by (28).
In Figure 1 and Figure 2 we show, respectively, convergence graphs for the diffusion-dominated and advection-dominated case. The numbers appearing in the yellow boxes, directly on the graph segments in our plots, represent the reduction rate associated to two subsequent errors, that is

$$m_{h_1,h_2} = \frac{\log(\mathrm{ERR}_{h_2} - \mathrm{ERR}_{h_1})}{\log(h_2 - h_1)}$$

where h_1, h_2 here denote the two mesh sizes associated to the segment endpoints.

In the first setting, the results are in agreement with the theoretical estimates, but exhibit a higher error reduction rate with respect to the theoretical prediction. Indeed, for $p < 2$ we observe a reduction of the error behaving as $O(h^k)$ instead of $O(h^{\frac{kp}{2}})$ while, for $p > 2$, the error decreases at a rate of $O(h^{\frac{kp}{2}})$ instead of $O(h^{\frac{kp'}{2}})$. In both cases, the reduction rate corresponds to that obtained by the best approximant to the solution u in $\mathcal{P}^k(\mathcal{T}_h)$; we better investigate this aspect in the next example.

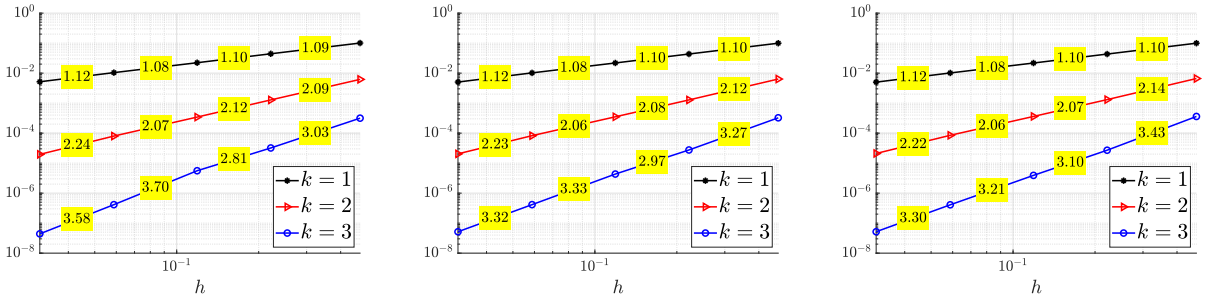
In the advection-dominated regime, on the other hand, the observed convergence rates closely match the theoretical estimates, i.e. ERR_h exhibits an $O(h^{k+\frac{1}{2}})$ decay. In particular, we can observe the additional $h^{\frac{1}{2}}$ factor which is gained due to the convection robustness of the method.

6.2 Example 2

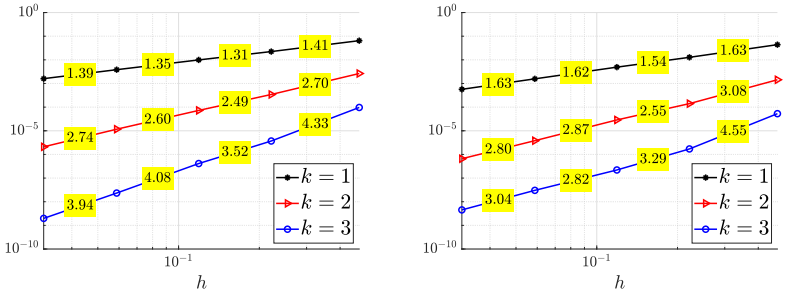
In the second example, we consider problem (1) without the presence of advection and reaction phenomena. The motivation of this second example is to better investigate the "higher than expected" reduction rate for the diffusion dominated case in Example 1. We therefore directly set $\nu = 1$, $\beta = 0$, $\mu = 0$ (pure diffusion) and choose the right-hand side and the Dirichlet boundary condition in accordance with two distinct solutions (the exponential (p, k)-dependent solution was originally proposed in [19]) :

- $u(x, y) := \frac{1}{10}\exp\left[-10\left(|-x + 0.5|^{p+\frac{k+2}{4}} + |-y + 0.5|^{p+\frac{k+2}{4}}\right)\right]$;
- $u(x, y) := \left(x - \frac{1}{2}\right)^2 \left(y - \frac{1}{2}\right)^2$

Here, the difference with respect to the preceding example is that the gradient ∇u vanishes at the point of coordinates $(0.5, 0.5)$ for the exponential solution, and in the region $\left\{(x, y) \in \Omega \mid x = \frac{1}{2} \text{ or } y = \frac{1}{2}\right\}$ for the polynomial solution, while, in the previous case, the solution had a non-zero gradient over the entire domain (which may determine a favorable situation for $p < 2$, see for instance [19]). In this respect, the the polynomial solution can be more challenging than the exponential one, as will be confirmed by the following results. Furthermore, the coarsest mesh, with meshsize $h = 0.4714$, is removed and replaced

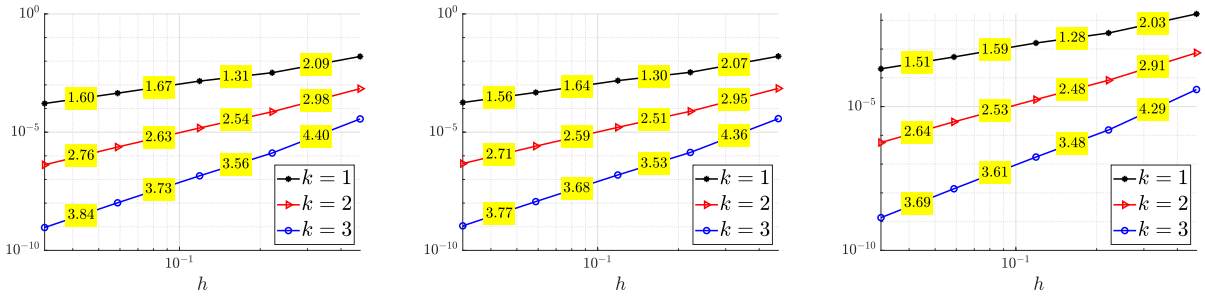


(a) $p = 1.5$ (b) $p = 1.75$ (c) $p = 2$

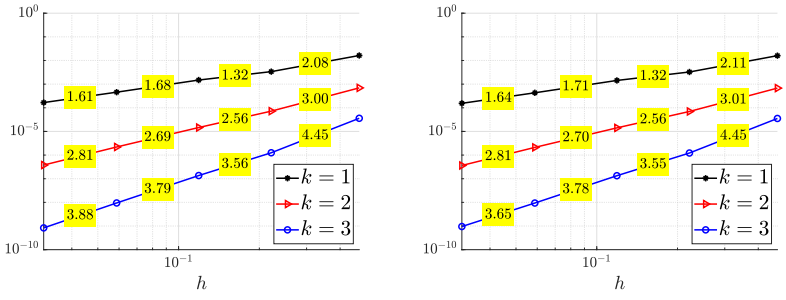


(d) $p = 2.5$ (e) $p = 3$

Figure 1: Example of Section 6.1. Convergence rate of ERR_h in the diffusion-dominated regime. Theoretical convergence rate: $\frac{kp}{2}$ for $p \le 2$ and $\frac{kp'}{2}$ for $p > 2$.

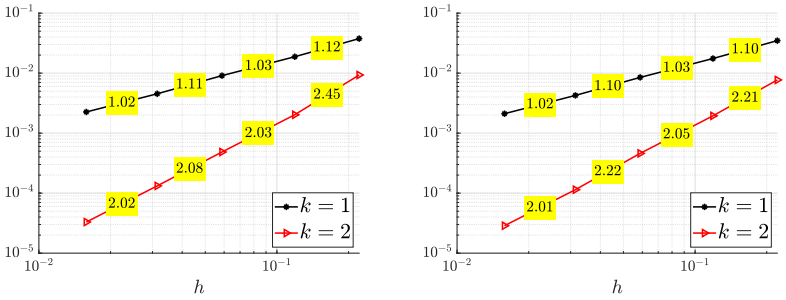


(a) $p = 1.5$ (b) $p = 1.75$ (c) $p = 2$



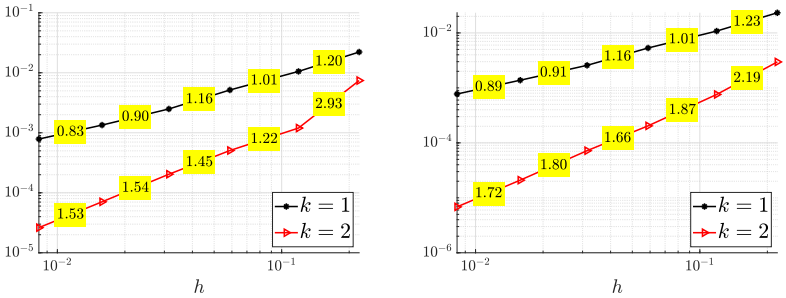
(d) $p = 2.5$ (e) $p = 3$

Figure 2: Example of Section 6.1. Convergence rate of ERR_h in the advection-dominated regime. Theoretical convergence rate: $k + \frac{1}{2}$ for all p.



(a) $p = 1.5$ (b) $p = 1.75$

Figure 3: Example of Section 6.2 with exponential solution. Convergence rate of ERR_h. Theoretical convergence rate: $\frac{kp}{2}$ for $p \in \{1.5, 1.75\}$.



(a) $p = 1.5$ (b) $p = 1.75$

Figure 4: Example of Section 6.2 with polynomial solution. Convergence rate of ERR_h. Theoretical convergence rate: $\frac{kp}{2}$ for $p \in \{1.5, 1.75\}$.

by two new meshes obtained by halving subsequently the finest mesh: this results in two new meshsizes with $h = 0.0158$ and $h = 0.0082$. For both solutions we have checked, by direct computation and/or numerically, that the flux σ is sufficiently regular for the estimates of Theorem 7 to hold.

As in the previous example, we compute the error term ERR_h (in this case with $\nu = 1$, $\beta = 0$, $\mu = 0$) considering $u_h \in \mathcal{P}^k(\mathcal{T}_h)$ for all combinations (p, k) with $p \in \{1.5, 1.75\}$ and $k \in \{1, 2\}$. The outcome in Figure 3, where ERR_h is plotted for the exponential solution, is similar to the previous example despite the different solution (now with vanishing gradient in a point of the domain) and the finer meshes adopted. Our current conclusions are that, probably, such behaviour is still pre-asymptotic, as is the case for the HHO method on meshes of similar size (cf., in particular, [19, Table 4]).

On the other hand, the results in Figure 4, showing the convergence rates for the polynomial solution, are aligned with the expected convergence rate on the light of Theorem 7. Indeed, an $O(h^{\frac{kp}{2}})$ decay of ERR_h can be observed (especially for the finer meshes), which confirms from the practical side the sharpness of the theoretical results.

Acknowledgements

The present results where partially supported by the European Union (ERC Synergy, NEMESIS, project number 101115663). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency.

References

[1] P. F. Antonietti, S. Giani, and P. Houston. "hp-version composite discontinuous Galerkin methods for elliptic problems on complicated domains". In: *SIAM J. Sci. Comput.* 35.3 (2013), A1417–A1439. DOI: 10.1137/120877246.

[2] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini. "Unified analysis of discontinuous Galerkin methods for elliptic problems". In: *SIAM J. Numer. Anal.* 39.5 (2001), pp. 1749–1779. DOI: 10.1137/S0036142901384162.

[3] B. Ayuso and L. D. Marini. "Discontinuous Galerkin methods for advection-diffusion-reaction problems". In: *SIAM J. Numer. Anal.* 47.2 (2009), pp. 1391–1420. DOI: 10.1137/080719583.

[4] G. A. Baker. "Finite element methods for elliptic equations using nonconforming elements". In: *Math. Comp.* 31.137 (1977), pp. 45–49. DOI: 10.2307/2005779.

[5] J. Barrett and W. Liu. "Finite element approximation of the p-Laplacian". In: *Math. Comp.* 61 (1993), pp. 523–537. DOI: 10.2307/2153239.

[6] F. Bassi, L. Botti, A. Colombo, D. A. Di Pietro, and P. Tesini. "On the flexibility of agglomeration based physical space discontinuous Galerkin discretizations". In: *J. Comput. Phys.* 231.1 (2012), pp. 45–65. DOI: 10.1016/j.jcp.2011.08.018.

[7] F. Bassi, L. Botti, and A. Colombo. "Agglomeration-based physical frame dG discretizations: an attempt to be mesh free". In: *Math. Models Methods Appl. Sci.* 24.8 (2014), pp. 1495–1539. DOI: 10.1142/S0218202514400028.

[8] L. Beirao da Veiga, F. Dassi, C. Lovadina, and G. Vacca. "SUPG-stabilized virtual elements for diffusion-convection problems: a robustness analysis". In: *ESAIM: M2AN* 55.5 (2021), pp. 2233–2258. DOI: 10.1051/m2an/2021050.

[9] L. Beirão da Veiga, F. Dassi, and G. Vacca. "Pressure robust SUPG-stabilized finite elements for the unsteady Navier–Stokes equation". In: *IMA J. Numer. Anal.* (2023). DOI: 10.1093/imanum/drad021.

[10] L. Beirão da Veiga, F. Dassi, and G. Vacca. "Vorticity-stabilized virtual elements for the Oseen equation". In: *Math. Models Methods Appl. Sci.* 31.14 (2021), pp. 3009–3052. DOI: 10.1142/S0218202521500688.

[11] F. Brezzi, L. D. Marini, and E. Süli. "Discontinuous Galerkin methods for first-order hyperbolic problems". In: *Math. Models Methods Appl. Sci.* 14.12 (2004), pp. 1893–1903. DOI: 10.1142/S0218202504003866.

[12] E. Burman and A. Ern. "Discontinuous Galerkin approximation with discrete variational principle for the nonlinear Laplacian". In: *C. R. Math. Acad. Sci. Paris* 346.17-18 (2008), pp. 1013–1016. DOI: 10.1016/j.crma.2008.07.005.

[13] P. Castillo, B. Cockburn, I. Perugia, and D. Schötzau. "Local discontinuous Galerkin methods for elliptic problems". In: *Comm. Numer. Methods Engrg.* 18.1 (2002), pp. 69–75. DOI: 10.1002/cnm.471.

[14] P. Castillo, B. Cockburn, I. Perugia, and D. Schötzau. "An A Priori Error Analysis of the Local Discontinuous Galerkin Method for Elliptic Problems". In: *SIAM J. Numer. Anal.* 38.5 (2000), pp. 1676–1706. DOI: 10.1137/S0036142900371003.

[15] B. Cockburn and C.-W. Shu. "The Runge-Kutta local projection P^1-discontinuous-Galerkin finite element method for scalar conservation laws". In: *RAIRO Modél. Math. Anal. Numér.* 25.3 (1991), pp. 337–361. DOI: 10.1051/m2an/1991250303371.

[16] L. M. Del Pezzo, A. L. Lombardi, and S. Martínez. "Interior penalty discontinuous Galerkin FEM for the $p(x)$-Laplacian". In: *SIAM J. Numer. Anal.* 50.5 (2012), pp. 2497–2521. DOI: 10.1137/110820324.

[17] D. A. Di Pietro and J. Droniou. "A polytopal method for the Brinkman problem robust in all regimes". In: *Comput. Meth. Appl. Mech. Engrg.* 409.115981 (2023). DOI: 10.1016/j.cma.2023.115981.

[18] D. A. Di Pietro and J. Droniou. *The Hybrid High-Order method for polytopal meshes. Design, analysis, and applications*. Vol. 19. Modeling, Simulation and Application. Springer International Publishing, 2020. DOI: 10.1007/978-3-030-37203-3.

[19] D. A. Di Pietro, J. Droniou, and A. Harnist. "Improved error estimates for Hybrid High-Order discretizations of Leray–Lions problems". In: *Calcolo* 58.19 (2021). DOI: 10.1007/s10092-021-00410-z.

[20] D. A. Di Pietro and A. Ern. "Discrete functional analysis tools for discontinuous Galerkin methods with application to the incompressible Navier–Stokes equations". In: *Math. Comp.* 79 (2010), pp. 1303–1330. DOI: 10.1090/S0025-5718-10-02333-1.

[21] D. A. Di Pietro and A. Ern. *Mathematical aspects of discontinuous Galerkin methods*. Vol. 69. Mathématiques & Applications (Berlin) [Mathematics & Applications]. Springer, Heidelberg, 2012. DOI: 10.1007/978-3-642-22980-0.

[22] D. A. Di Pietro, A. Ern, and J.-L. Guermond. "Discontinuous Galerkin methods for anisotropic semi-definite diffusion with advection". In: *SIAM J. Numer. Anal.* 46.2 (2008), pp. 805–831. DOI: 10.1137/060676106.

[23] L. Diening and F. Ettwein. "Fractional estimates for non-differentiable elliptic systems with general growth". In: *Forum Math.* 20.3 (2008), pp. 523–556. DOI: 10.1515/FORUM.2008.027.

[24] Y. Han and Y. Hou. "Semirobust analysis of an H(div)-conforming DG method with semi-implicit time-marching for the evolutionary incompressible Navier–Stokes equations". In: *IMA J. Numer. Anal.* 42.2 (2021), pp. 1568–1597. DOI: 10.1093/imanum/draa104.

[25] A. Hirn. "Approximation of the p-Stokes Equations with Equal-Order Finite Elements". In: *J. Math. Fluid Mech.* 15 (2013), pp. 65–88. DOI: 10.1007/s00021-012-0095-0.

[26] J. Leray and J.-L. Lions. "Quelques résulatats de Višik sur les problèmes elliptiques nonlinéaires par les méthodes de Minty-Browder". In: *Bull. Soc. Math. France* 93 (1965), pp. 97–107. URL: http://www.numdam.org/item?id=BSMF_1965__93__97_0.

[27] W. H. Reed and T. R. Hill. *Triangular mesh methods for the neutron transport equation*. Tech. rep. LA-UR-73-0479. Los Alamos, NM: Los Alamos Scientific Laboratory, 1973. URL: http://lib-www.lanl.gov/cgi-bin/getfile%7B?%7D00354107.pdf.