


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

Lourenço Beirão da Veiga ^{*}

*Dipartimento di Matematica e Applicazioni,
Università degli Studi di Milano-Bicocca,*

Piazza dell'Ateneo Nuovo 1, 20126 Milano, Italy

Institute for Applied Mathematics and Information Technologies

“Enrico Magenes” (IMATI-PV), CNR, Via Ferrata 5/a, 27100 Pavia, Italy
lourenco.beirao@unimib.it

Daniele A. Di Pietro 

IMAG, Univ. Montpellier, CNRS, Montpellier, France,
daniele.di-pietro@umontpellier.fr

Kirubell B. Haile 

*Dipartimento di Matematica e Applicazioni,
Università degli Studi di Milano-Bicocca,*

Piazza dell'Ateneo Nuovo 1, 20126 Milano, Italy
k.haile@campus.unimib.it

Received 29 February 2024

Revised 8 April 2024

Accepted 29 April 2024

Published 29 June 2024

Communicated by F. Brezzi

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 the local Péclet numbers. A set of numerical tests support the theoretical derivations.

Keywords: Discontinuous Galerkin methods; diffusion–advection–reaction problems; p -Laplacian; Péclet-robust error estimates.

AMS Subject Classification 2020: 65N30, 65N08, 35K55

*Corresponding author.

This is an Open Access article published by World Scientific Publishing Company. It is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND) License, which permits use, distribution and reproduction, provided that the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

1. Introduction

Discontinuous Galerkin (DG) methods were introduced in the 70s,^{4, 29} 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. The specific focus of this work is on scalar advection with possibly nonlinear p -type diffusion. This type of problems arise in the modeling of several physical processes such as, for example, the flow of physical particles, chemical species or heat. They also serve as a conceptual model for the momentum conservation equation in the modeling of incompressible non-Newtonian fluid flows. When a polynomial degree $k \geq 1$ is used, classical error estimates for linear diffusion–advection(–reaction) problems (corresponding to the case $p = 2$ in this context) show that the error contribution stemming from diffusive terms is $\mathcal{O}(h^k)$ (with h denoting the mesh size), while the one stemming from advective terms is $\mathcal{O}(h^{k+\frac{1}{2}})$; see, for example, Ref. 3 as well as Ref. 22 and Sec. 4.6 of Ref. 21 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.

Moving to nonlinear problems, convergence analyses for various DG schemes applied to p -type diffusion can be found, for example, in Refs. 12, 16, 24 and 28. Specifically, in Refs. 12 and 16, the authors consider convergence by compactness of discretizations that are based on a full gradient incorporating the contribution from inter-element jumps. In Ref. 28, on the other hand, schemes where consistency is achieved through a separate boundary term are also considered, and error estimates are proved using a generalized Scott–Zhang interpolator valid for conforming simplicial meshes.

To our knowledge, a convergence study for model problems including both advection and nonlinear diffusion is missing in the literature of DG elements. Especially if one aims at developing sharp estimates which respect the local nature of diffusion and convection, the interaction between the (linear) advection and the nonlinear diffusion cannot be accounted for through a simple combination of known techniques, as the estimate of each term becomes dependent on the local regime.

The goal of this work is precisely to derive 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 Refs. 12 and 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 by interpreting the latter as a penalty contribution in the spirit of Ref. 11. The peculiarity of our error estimates is that they track the dependence of the local contributions to the error on local Péclet numbers, as mentioned above. To improve the estimates of certain terms, we provide a new extension to the nonconforming case of the techniques of Ref. 26, based in turn on the results of Ref. 23 (see also Ref. 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 diffusions). 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. We emphasize that our results hold for general polygonal and polyhedral meshes, which we believe is an important asset. This is obtained by avoiding to use globally continuous polynomial approximants, a choice which would have simplified certain proofs (by making many jump terms vanish) at the price of restricting the range of eligible meshes. This contribution furthermore sets the stage for future publications developing pressure-robust and advection-robust finite elements for time-dependent Navier–Stokes-type equations^{10, 25} modeling incompressible fluid flows with non-Newtonian rheology.

We finally remark that all theoretical results are supported by an extensive numerical validation.

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

2. The Continuous Problem

Let $\Omega \subset \mathbb{R}^d$, $d \geq 2$, denote a bounded, connected polyhedral domain. We develop a Péclet-robust Discontinuous Galerkin method for the following problem: Find $u : \Omega \rightarrow \mathbb{R}$ such that

$$\begin{aligned} -\nabla \cdot [\sigma(\nabla u) - \beta u] + \mu u &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{2.1}$$

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

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

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 the Leray–Lions-type operators and their generalizations; see, for example, Refs. 27 and 23.

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 mesh size (see below), the problem data and solution.

Lemma 2.1. (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*

$$\begin{aligned} & (\sigma_n(x) - \sigma_n(z)) \cdot (x - y) \\ & \leq \delta(\sigma_n(y) - \sigma_n(x)) \cdot (y - x) + c(\delta)(|x| + |z|)^{p-2}|x - z|^2, \end{aligned} \tag{2.2}$$

with $c(\delta)$ being a 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 $\varphi_a(t) := \int_0^t \varphi'(a+s) \frac{s}{a+s} ds$. By Eq. (6.28) in Ref. 23, it holds, for all $s, t \in \mathbb{R}^+$ and all $\delta > 0$,

$$\varphi'_a(s)t + \varphi'_a(t)s \leq \delta\varphi_a(s) + c(\delta)\varphi_a(t). \tag{2.3}$$

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

$$(\sigma_n(y) - \sigma_n(x)) \cdot (y - x) \simeq \varphi_{|x|}(|x - y|) \simeq (|x| + |y|)^{p-2}|x - y|^2 \tag{2.4}$$

and, as observed in Lemma 2.3 of Ref. 26, for all $x, z \in \mathbb{R}^n$,

$$|\sigma_n(x) - \sigma_n(z)| \lesssim \varphi'_{|x|}(|x - z|). \tag{2.5}$$

Let now $x, y, z \in \mathbb{R}^n$. Applying (2.3) with $a = |x|$, $s = |x - y|$ and $t = |x - z|$, and noticing that $\varphi'_a(s)t \geq 0$, we get

$$\varphi'_{|x|}(|x - z|)|x - y| \leq \delta\varphi_{|x|}(|x - y|) + c(\delta)\varphi_{|x|}(|x - z|). \tag{2.6}$$

The conclusion follows first observing that $(\sigma_n(x) - \sigma_n(z)) \cdot (x - y) \leq |\sigma_n(x) - \sigma_n(z)||x - y|$ and using (2.5) to estimate the left-hand side of (2.2) with the left-hand side of (2.6), then applying, respectively, the left-most equivalence in (2.4) to estimate $\varphi_{|x|}(|x - y|)$ and the right-most equivalence in (2.4) to estimate $\varphi_{|x|}(|x - z|)$ in the right-hand side of (2.6). \square

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 Sec. 1.4 of Ref. 21. 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^b and interfaces collected in \mathcal{F}_h^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 $h := \max_{T \in \mathcal{T}_h} h_T$. 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 $a \leq Cb$ with C being independent of the mesh size, 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 3.1. (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 instance, Refs. 9, 19, 8 and 18.

3.1. Local and broken spaces

Given a polynomial degree $k \geq 0$ 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

$$\mathcal{P}^k(\mathcal{T}_h) := \{v_h \in L^1(\Omega) \mid v_T := (v_h)|_T \in \mathcal{P}^k(T) \text{ for all } T \in \mathcal{T}_h\}.$$

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

$$W^{q,p}(\mathcal{T}_h) := \{v \in L^p(\Omega) \mid v|_T \in W^{q,p}(T) \text{ for all } T \in \mathcal{T}_h\}.$$

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

$$\frac{1}{p} + \frac{1}{p'} = 1 \Leftrightarrow p' = \frac{p}{p-1}. \tag{3.1}$$

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 $F \in \mathcal{F}_h^i$, 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 $\varphi \in W^{1,1}(\mathcal{T}_h)$,

$$[\varphi]_F := \varphi|_{T_1} - \varphi|_{T_2}, \quad \{\varphi\}_F := \frac{1}{2}(\varphi|_{T_1} + \varphi|_{T_2}).$$

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

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

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

$$\begin{aligned} \int_{\Omega} \tau \cdot \nabla_h v &= - \int_{\Omega} (\nabla_h \cdot \tau) v + \sum_{F \in \mathcal{F}_h^i} \int_F \{\tau\}_F \cdot n_F [v]_F \\ &+ \sum_{F \in \mathcal{F}_h^b} \int_F [\tau]_F \cdot n_F \{\varphi\}_F + \int_{\partial\Omega} (\tau \cdot n) v, \end{aligned} \tag{3.2}$$

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 $r_F^k : L^1(F) \rightarrow \mathcal{P}^k(\mathcal{T}_h)^d$ such that, for all $\psi \in L^1(F)$,

$$\int_{\Omega} r_F^k \psi \cdot \tau_h = \int_F \psi \{\tau_h\}_F \cdot n_F, \quad \forall \tau_h \in \mathcal{P}^k(\mathcal{T}_h)^d, \tag{3.3}$$

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

$$R_h^k \varphi := \sum_{F \in \mathcal{F}_h} r_F^k([\varphi]_F). \tag{3.4}$$

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

$$G_h^k \varphi := \nabla_h \varphi - R_h^k \varphi, \tag{3.5}$$

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

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

$$\begin{aligned} \|\varphi\|_{1,p,h} &:= \left(\|\nabla_h \varphi\|_{L^p(\Omega)^d}^p + |\varphi|_{1,p,h}^p \right)^{\frac{1}{p}} \quad \text{with} \\ |\varphi|_{1,p,h} &:= \left(\sum_{F \in \mathcal{F}_h} h_F^{1-p} \|[\varphi]_F\|_{L^p(F)}^p \right)^{\frac{1}{p}}, \end{aligned} \tag{3.6}$$

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

$$\|\varphi\|_{1,\infty,h} := \|\nabla_h \varphi\|_{L^\infty(\Omega)^d} + |\varphi|_{1,\infty,h} \quad \text{with} \quad |\varphi|_{1,\infty,h} := \max_{F \in \mathcal{F}_h} h_F^{-1} \|[\varphi]_F\|_{L^\infty(F)}. \tag{3.7}$$

In local estimates, we will also need the following local versions of the norms (3.6) and (3.7): For all $T \in \mathcal{T}_h$ and all $\varphi \in W^{1,p}(\mathcal{T}_T)$,

$$\begin{aligned} \|\varphi\|_{1,p,T} &:= \left(\|\nabla_h \varphi\|_{L^p(T)^d}^p + |\varphi|_{1,p,T}^p \right)^{\frac{1}{p}} \quad \text{with} \\ |\varphi|_{1,p,T} &:= \left(\sum_{F \in \mathcal{F}_T} h_F^{1-p} \|\llbracket \varphi \rrbracket_F\|_{L^p(F)}^p \right)^{\frac{1}{p}} \end{aligned} \tag{3.8}$$

and

$$\|\varphi\|_{1,\infty,T} := \|\nabla_h \varphi\|_{L^\infty(T)^d} + |\varphi|_{1,\infty,T} \quad \text{with} \quad |\varphi|_{1,\infty,T} := \max_{F \in \mathcal{F}_T} h_F^{-1} \|\llbracket \varphi \rrbracket_F\|_{L^\infty(F)}.$$

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

(1) Boundedness: For all $\varphi \in W^{1,p}(\mathcal{T}_h)$,

$$\|r_F^k([\varphi]_F)\|_{L^p(\Omega)^d} \lesssim h_F^{\frac{1-p}{p}} \|\llbracket \varphi \rrbracket_F\|_{L^p(F)}, \quad \forall F \in \mathcal{F}_h, \tag{3.9}$$

with the convention that $h_F^{\frac{1-\infty}{p}} := h_F^{-1}$ and

$$\|R_h^k \varphi\|_{L^p(T)^d} \lesssim |\varphi|_{1,p,T}, \quad \forall T \in \mathcal{T}_h. \tag{3.10}$$

(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 $r \in \{0, \dots, k\}$,

$$\|R_h^k \pi_h^k w\|_{L^p(T)^d} \lesssim h_T^r |w|_{W^{r+1,p}(\mathcal{T}_T)}, \quad \forall T \in \mathcal{T}_h. \tag{3.11}$$

Proof of (3.9)–(3.10). It holds, for all $F \in \mathcal{F}_h$,

$$\begin{aligned} \|r_F^k([\varphi]_F)\|_{L^p(\Omega)^d} &= \sup_{\tau \in L^{p'}(\Omega)^d \setminus \{0\}} \frac{\int_\Omega r_F^k([\varphi]_F) \cdot \tau}{\|\tau\|_{L^{p'}(\Omega)^d}} \\ &= \sup_{\tau \in L^{p'}(\Omega)^d \setminus \{0\}} \frac{\int_\Omega r_F^k([\varphi]_F) \cdot \pi_h^k \tau}{\|\tau\|_{L^{p'}(\Omega)^d}} \\ &\stackrel{(3.3)}{=} \sup_{\tau \in L^{p'}(\Omega)^d \setminus \{0\}} \frac{\int_F [\varphi]_F \{ \pi_h^k \tau \}_F \cdot n_F}{\|\tau\|_{L^{p'}(\Omega)^d}}, \end{aligned}$$

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 $h_F^{-\frac{1}{p'}} := 1$ if $p' = \infty$,

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

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 of Lemma 1.32 in Ref. 17 in the first passage, while the second passage is a consequence of the $L^{p'}$ -boundedness of the L^2 -orthogonal projector (cf. Lemma 1.44 in Ref. 17). Additionally noticing that, by (3.1), $-\frac{1}{p'} = \frac{1-p}{p}$ and that $\|\tau\|_{L^{p'}(\mathcal{T}_F)^d} \leq \|\tau\|_{L^{p'}(\Omega)^d}$, yields (3.9).

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 (3.9) and finally use $\text{card}(\mathcal{F}_T) \lesssim 1$,

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

which is the bound (3.10). □

Proof of (3.11). 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 $\text{card}(\mathcal{F}_T) \lesssim 1$, (3.11) follows. If $p = \infty$, we have

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

which concludes the proof. □

Lemma 3.2. (Approximation properties of the discrete gradient) *For all integers $k \geq 0$, 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*

$$\|G_h^k \pi_h^k w - \nabla w\|_{L^p(T)^d} \lesssim h_T^r |w|_{W^{r+1,p}(\mathcal{T}_T)}, \quad \forall T \in \mathcal{T}_h. \tag{3.12}$$

Proof. Using (3.5) and a triangle inequality, we obtain

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

The conclusion follows using the approximation properties of the L^2 -orthogonal projector for the first term and (3.11) for the second. □

Remark 3.2. (Local boundedness of $G_h^k \circ \pi_h^k$) For any $q \in [1, \infty]$, combining a triangle inequality with (3.12) written for $r = 0$, it is readily inferred that, for all $\varphi \in W^{1,q}(\mathcal{T}_h)$,

$$\|G_h^k \pi_h^k \varphi\|_{L^q(T)^d} \lesssim |\varphi|_{W^{1,q}(\mathcal{T}_T)}, \quad \forall T \in \mathcal{T}_h. \tag{3.13}$$

4. Discrete Problem and Main Results

4.1. Discrete problem

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

$$a_h(w_h, v_h) := \int_{\Omega} \sigma(G_h^k w_h) \cdot G_h^k v_h + s_h(w_h, v_h), \tag{4.1}$$

where

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

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

$$\begin{aligned} b_h(w_h, v_h) = & - \int_{\Omega} w_h (\beta \cdot \nabla_h v_h) + \int_{\Omega} \mu w_h v_h + \sum_{F \in \mathcal{F}_h} \int_F (\beta \cdot n_F) \{w_h\}_F [v_h]_F \\ & + \frac{1}{2} \sum_{F \in \mathcal{F}_h} \hat{\beta}_F \int_F [w_h]_F [v_h]_F, \end{aligned} \tag{4.2}$$

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

$$\hat{\beta}_F := \|\beta \cdot n_F\|_{L^\infty(F)}.$$

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

Remark 4.1. (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 the boundary conditions in the vanishing diffusion case.

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

$$a_h(u_h, v_h) + b_h(u_h, v_h) = \int_{\Omega} f v_h, \quad \forall v_h \in \mathcal{P}^k(\mathcal{T}_h). \tag{4.3}$$

4.2. Main results

In this subsection, we collect the main results of the analysis of problem (4.3). 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, Ref. 18).

4.2.1. *Dimensionless numbers and reference quantities*

In order to state these convergence results, we need to define here the 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 $\hat{\beta}_T := \|\beta\|_{L^\infty(T)^d}$ vanishes, we set $\text{Pe}_T(w) = 0$; otherwise,

$$\text{Pe}_T(w) := \frac{\hat{\beta}_T h_T}{\hat{K}_T(w)} \quad \text{with } \hat{K}_T(w) := \|\lvert \nabla w \rvert^{p-2}\|_{L^\infty(\mathcal{T}_T)}, \tag{4.4}$$

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

$$\hat{\tau}_T := \frac{1}{\max(\|\mu\|_{L^\infty(T)}, |\beta|_{W^{1,\infty}(T)^d})}. \tag{4.5}$$

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

$$\begin{aligned} \text{Pe}_F(w) &:= \frac{\hat{\beta}_F h_F}{\hat{K}_F(w)} \quad \text{with} \\ \hat{K}_F(w) &:= \max\left(\|\lvert \nabla w \rvert^{p-2}\|_{L^\infty(F)}, h_F^{2-p} \|\lvert [\pi_h^k w]_F \rvert^{p-2}\|_{L^\infty(F)}\right), \end{aligned} \tag{4.6}$$

where again $\hat{K}_F(w) = +\infty$ (and thus $\text{Pe}_F(w) = 0$) 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 $w : \Omega \rightarrow \mathbb{R}$, we set

$$\begin{aligned} \mathcal{T}_h^a(w) &:= \{T \in \mathcal{T}_h \mid \text{Pe}_T(w) > 1\}, & \mathcal{T}_h^d(w) &:= \mathcal{T}_h \setminus \mathcal{T}_h^a(w), \\ \mathcal{F}_h^a(w) &:= \{F \in \mathcal{F}_h \mid \text{Pe}_F(w) > 1\}, & \mathcal{F}_h^d(w) &:= \mathcal{F}_h \setminus \mathcal{F}_h^a(w). \end{aligned}$$

4.2.2. *Norms*

The relevant norm for the analysis of the diffusion terms is $\|\cdot\|_{1,p,h}$ (cf. (3.6)) as well as its restriction to an element $T \in \mathcal{T}_h$ (cf. (3.8)). The norm for the advective

and reactive terms is, on the other hand, given by

$$\|v_h\|_{\beta,\mu,h} := \left(\frac{1}{2} \sum_{F \in \mathcal{F}_h} \hat{\beta}_F \|[v_h]_F\|_{L^2(F)}^2 + \|\mu^{\frac{1}{2}} v_h\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}}, \quad \forall v_h \in \mathcal{P}^k(\mathcal{T}_h). \quad (4.7)$$

This choice of advection–reaction norm is justified as follows. By standard arguments (which essentially amount to applying the integration by parts formula (3.2) with $(\tau, v) = (\beta w_h, v_h)$ to the first term in the right-hand side of (4.2), using the continuity of $\beta \cdot n_F$ across interfaces and recalling that $\nabla \cdot \beta = 0$), it is easy to check that

$$b_h(v_h, v_h) = \|v_h\|_{\beta,\mu,h}^2, \quad \forall v_h \in \mathcal{P}^k(\mathcal{T}_h), \quad (4.8)$$

showing that b_h is coercive with respect to the norm defined by (4.7) with the 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 (4.3) 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 $u \in W^{1,p}(\Omega) \cap W^{r+1,\bar{p}}(\mathcal{T}_h)$ with $\bar{p} = \max\{2, 2p - 2, p'\}$.

Theorem 4.1. (Convergence) *Denote, respectively, by $u \in W_0^{1,p}(\Omega)$ and by $u_h \in \mathcal{P}^k(\mathcal{T}_h)$ the solutions of the weak formulation of problem (2.1) and of the discrete problem (4.3). Additionally assume that, for some $r \in \{0, \dots, k\}$,*

- $u|_{\omega_T} \in W^{r+1,p}(\mathcal{T}_T)$ for all $T \in \mathcal{T}_h^d(u)$;
- $u|_T \in H^{r+1}(T)$ 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 $F \in \mathcal{F}_h^d(u)$;
- $u|_{\omega_F} \in H^{r+1}(\mathcal{T}_F)$ and $\sigma(\nabla u)|_{\omega_F} \in H^{r+\frac{1}{2}}(\mathcal{T}_F)^d$ for all $F \in \mathcal{F}_h^a(u)$.

Then, letting

$$q := \begin{cases} 2 & \text{if } p < 2, \\ p & \text{if } p \geq 2, \end{cases} \quad (4.9)$$

it holds

$$\begin{aligned} & \|u_h - \pi_h^k u\|_{1,p,h}^q + \|u_h - \pi_h^k u\|_{\beta,\mu,h}^2 \\ & \lesssim \sum_{T \in \mathcal{T}_h} \hat{\tau}_T^{-2} \underline{\mu}_T^{-1} h_T^{2(r+1)} |u|_{H^{r+1}(T)}^2 \\ & \quad + \sum_{T \in \mathcal{T}_h^a(u)} \hat{\beta}_T h_T^{2r+1} |u|_{H^{r+1}(T)}^2 + \sum_{T \in \mathcal{T}_h^d(u)} \begin{cases} h_T^{rp} |u|_{W^{r+1,p}(\mathcal{T}_T)}^p & \text{if } p < 2 \\ h_T^{2r} |u|_{W^{r+1,p}(\mathcal{T}_T)}^2 & \text{if } p \geq 2 \end{cases} \end{aligned}$$

$$\begin{aligned}
 & + \sum_{F \in \mathcal{F}_h^a(u)} h_F^{2r+1} \left(\hat{K}_F(u)^{-1} |\sigma(\nabla u)|_{H^{r+\frac{1}{2}}(\mathcal{T}_F)^d}^2 + \hat{\beta}_F |u|_{H^{r+1}(\mathcal{T}_F)}^2 \right) \\
 & + \sum_{F \in \mathcal{F}_h^d(u)} h_F^{rp} |u|_{W^{r+1,p}(\mathcal{T}_F)}^p \\
 & + \left[\sum_{F \in \mathcal{F}_h^d(u)} h_F^{rp'} \left(|\sigma(\nabla u)|_{W^{r,p'}(\mathcal{T}_F)^d}^{p'} + \hat{K}_F(u)^{p'} |u|_{W^{r+1,p'}(\mathcal{T}_F)}^{p'} \right) \right]^{\frac{q'}{p'}} \cdot \quad (4.10)
 \end{aligned}$$

Proof. See Sec. 5.3. □

The above convergence result is fully local, being able to deliver sharp estimates also in situations where diffusion or advection dominates 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 the conforming Finite Element (FE) schemes (i.e. $\mathcal{O}(h^{\frac{rp}{2}})$), see Refs. 5 and 26. 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, notice that the negative power of \hat{K}_F appearing in the bound above is balanced by the associated σ term, see Remark 5.1. The case $p = 2$ corresponds to a linear diffusion–advection–reaction problem, for which the classical estimates are recovered (see, for example, Refs. 22 and 3 and also Sec. 4.6 of Ref. 21). 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 $\mathcal{O}(h^{rp'})$ term in the right-hand side (to be compared with $\mathcal{O}(h^{2r})$). 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 5.2.

5. Theoretical Analysis

5.1. Properties of the diffusion function

Lemma 5.1. (Stability of a_h) *For any $w_h, v_h \in \mathcal{P}^k(\mathcal{T}_h)$, recalling the definition (4.9) 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

$$C_a \|w_h - v_h\|_{1,p,h}^q \lesssim a_h(w_h, w_h - v_h) - a_h(v_h, w_h - v_h). \tag{5.1}$$

Proof. The proof is a straightforward adaptation of the monotonicity properties of σ and the arguments of Theorem 6.19(ii) in Ref. 17. \square

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

Lemma 5.2. (Estimate of the discrete diffusion error) *Let $w \in W_0^{1,p}(\Omega)$ 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 us define the diffusion error linear form $\mathcal{E}_{a,h}^k : \mathcal{P}^k(\mathcal{T}_h) \rightarrow \mathbb{R}$ such that, for all $v_h \in \mathcal{P}^k(\mathcal{T}_h)$,*

$$\mathcal{E}_{a,h}^k(w; v_h) := - \int_{\Omega} \nabla \cdot \sigma(\nabla w) v_h - a_h(\pi_h^k w, v_h). \tag{5.2}$$

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

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

Then, recalling (4.9), it holds, for any $w_h \in \mathcal{P}^k(\mathcal{T}_h)$ and any real number $\delta > 0$,

$$\begin{aligned} & \mathcal{E}_{a,h}^k(w; w_h - \pi_h^k w) \\ & \leq \delta \left(a_h(w_h, w_h - \pi_h^k w) - a_h(\pi_h^k w, w_h - \pi_h^k w) \right. \\ & \quad \left. + |w_h - \pi_h^k w|_{1,p,h}^q + \|w_h - \pi_h^k w\|_{\beta,\mu,h}^2 \right) \\ & \quad + c(\delta) \left(\sum_{T \in \mathcal{T}_h^a(w)} \hat{\beta}_T h_T^{2r+1} |w|_{H^{r+1}(T)}^2 + \sum_{T \in \mathcal{T}_h^d(w)} \begin{cases} h_T^{rp} |w|_{W^{r+1,p}(\mathcal{T}_T)}^p & \text{if } p < 2 \\ h_T^{2r} |w|_{W^{r+1,p}(\mathcal{T}_T)}^2 & \text{if } p \geq 2 \end{cases} \right) \\ & \quad + c(\delta) \sum_{F \in \mathcal{F}_h^a(w)} \hat{K}_F(w)^{-1} h_F^{2r+1} |\sigma(\nabla w)|_{H^{r+\frac{1}{2}}(\mathcal{T}_F)^d}^2 \\ & \quad + c(\delta) \left(\sum_{F \in \mathcal{F}_h^d(w)} h_F^{rp'} |\sigma(\nabla w)|_{W^{r,p'}(\mathcal{T}_F)^d}^{p'} \right)^{\frac{q'}{p'}} \\ & \quad + c(\delta) \left(\sum_{F \in \mathcal{F}_h^a(w)} \hat{\beta}_F h_F^{2r+1} |w|_{H^{r+1}(\mathcal{T}_F)}^2 + \sum_{F \in \mathcal{F}_h^d(w)} h_F^{rp} |w|_{W^{r+1,p}(\mathcal{T}_F)}^p \right), \tag{5.3} \end{aligned}$$

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

Proof. Let, for the sake of brevity,

$$v_h := w_h - \pi_h^k w.$$

Using the integration by parts formula (3.2) for the first term in the right-hand side of (5.2) along with the fact that $[\sigma(\nabla w)]_F \cdot n_F$ vanishes for all $F \in \mathcal{F}_h^i$ (which expresses the continuity of normal fluxes), expanding a_h according to its definition (4.1), 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. (3.4) and (3.3)) and adding and subtracting $\int_\Omega \sigma(\nabla w) \cdot R_h^k v_h$, we arrive at the following decomposition of the error:

$$\begin{aligned} \mathcal{E}_{a,h}^k(w; v_h) &= \underbrace{\int_\Omega [\sigma(G_h^k \pi_h^k w) - \sigma(\nabla w)] \cdot G_h^k (\pi_h^k w - w_h)}_{\mathfrak{I}_1} \\ &\quad - \underbrace{\sum_{F \in \mathcal{F}_h} \int_F \{\sigma(\nabla w) - \pi_h^k \sigma(\nabla w)\}_F \cdot n_F [v_h]_F}_{\mathfrak{I}_2} \\ &\quad + \int_\Omega \cancel{[\sigma(\nabla w) - \pi_h^k \sigma(\nabla w)] \cdot R_h^k v_h} + \underbrace{s_h(\pi_h^k w, \pi_h^k w - w_h)}_{\mathfrak{I}_3}, \end{aligned} \tag{5.4}$$

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{I}_1 . For the first term, we start by writing $\mathfrak{I}_1 = \sum_{T \in \mathcal{T}_h} \mathfrak{I}_1(T)$ and consider a single $T \in \mathcal{T}_h$. Using the bound (2.2) 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

$$\mathfrak{I}_1(T) \leq \delta \left(\int_T \sigma(G_h^k w_h) \cdot G_h^k v_h - \int_T \sigma(G_h^k \pi_h^k w) \cdot G_h^k v_h \right) + c(\delta) \mathfrak{I}_{1,\text{err}}(T), \tag{5.5}$$

where

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

We now distinguish between diffusion-dominated and advection-dominated elements.

Let first $T \in \mathcal{T}_h^d(w)$. In the case $p < 2$, we use the fact that $|\nabla w - G_h^k \pi_h^k w| \leq |\nabla w| + |G_h^k \pi_h^k w|$ 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

$$\mathfrak{I}_{1,\text{err}}(T) \leq \int_T |\nabla w - G_h^k \pi_h^k w|^p = \|\nabla w - G_h^k \pi_h^k w\|_{L^p(T)^d}^p \stackrel{(3.12)}{\lesssim} h_T^{rp} |w|_{W^{r+1,p}(\mathcal{T}_T)}^p. \tag{5.6}$$

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

$$\begin{aligned} \mathfrak{T}_{1,\text{err}}(T) &\lesssim \left(\|\nabla w\|_{L^p(T)^d} + \|G_h^k \pi_h^k w\|_{L^p(T)^d} \right)^{p-2} \|\nabla w - G_h^k \pi_h^k w\|_{L^p(T)^d}^2 \\ &\stackrel{(3.13), (3.12)}{\lesssim} \|\nabla w\|_{L^p(\mathcal{T}_T)^d}^{p-2} h_T^{2r} |w|_{W^{r+1,p}(\mathcal{T}_T)}^2 \lesssim h_T^{2r} |w|_{W^{r+1,p}(\mathcal{T}_T)}^2, \end{aligned} \tag{5.7}$$

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

Let now $T \in \mathcal{T}_h^a(w)$. We first consider the case $p < 2$. Using again the fact that $\mathbb{R}^+ \ni x \mapsto x^{p-2} \in \mathbb{R}$ is strictly decreasing, then applying a Hölder inequality with exponents $(\infty, 1)$ and using the approximation properties (3.12) of $G_h^k \circ \pi_h^k$, and finally recalling that $\text{Pe}_T(w) > 1$ (cf. (4.4) for its definition), we have

$$\begin{aligned} \mathfrak{T}_{1,\text{err}}(T) &\lesssim \int_T |\nabla w|^{p-2} |\nabla w - G_h^k \pi_h^k w|^2 \lesssim \| |\nabla w|^{p-2} \|_{L^\infty(T)} \|\nabla w - G_h^k \pi_h^k w\|_{L^2(T)^d}^2 \\ &\lesssim \hat{K}_T(w) h_T^{2r} |w|_{H^{r+1}(T)}^2 \leq \hat{\beta}_T h_T^{2r+1} |w|_{H^{r+1}(T)}^2. \end{aligned} \tag{5.8}$$

In the case $p \geq 2$, on the other hand, the local boundedness (3.13) of $G_h^k \circ \pi_h^k$ with $q = \infty$ along with the definition (4.4) of $\hat{K}_T(w)$ easily lead to

$$\mathfrak{T}_{1,\text{err}}(T) \lesssim \hat{K}_T(w) h_T^{2r} |w|_{H^{r+1}(T)}^2 \lesssim \hat{\beta}_T h_T^{2r+1} |w|_{H^{r+1}(T)}^2, \tag{5.9}$$

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

Plugging the estimates (5.6)–(5.9) into (5.5), we arrive at

$$\begin{aligned} \mathfrak{T}_1 &\leq \delta \left(\int_\Omega \sigma(G_h^k w_h) \cdot G_h^k v_h - \int_\Omega \sigma(G_h^k \pi_h^k w) \cdot G_h^k v_h \right) \\ &\quad + c(\delta) \sum_{T \in \mathcal{T}_h^a(w)} \hat{\beta}_T h_T^{2r+1} |w|_{H^{r+1}(T)}^2 \\ &\quad + c(\delta) \sum_{T \in \mathcal{T}_h^d(w)} \begin{cases} h_T^{rp} |w|_{W^{r+1,p}(\mathcal{T}_T)}^p & \text{if } p < 2, \\ h_T^{2r} |w|_{W^{r+1,p}(\mathcal{T}_T)}^2 & \text{if } p \geq 2. \end{cases} \end{aligned} \tag{5.10}$$

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 $F \in \mathcal{F}_h^d(w)$, we estimate $\mathfrak{T}_2(F)$ as follows:

$$\begin{aligned} \mathfrak{T}_2(F) &\leq \| \{ \sigma(\nabla w) - \pi_h^k \sigma(\nabla w) \}_F \|_{L^{p'}(F)^d} \| [v_h]_F \|_{L^p(F)} \\ &\lesssim h_F^{r - \frac{1}{p'} + \frac{p-1}{p}} | \sigma(\nabla w) |_{W^{r,p'}(\mathcal{T}_F)^d} h_F^{\frac{1-p}{p}} \| [v_h]_F \|_{L^p(F)} \\ &\stackrel{(3.1)}{=} h_F^r | \sigma(\nabla w) |_{W^{r,p'}(\mathcal{T}_F)^d} h_F^{\frac{1-p}{p}} \| [v_h]_F \|_{L^p(F)}, \end{aligned} \tag{5.11}$$

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 $F \in \mathcal{F}_h^a(w)$, on the other hand, we first notice that $\hat{\beta}_F \neq 0$ and write

$$\begin{aligned} \mathfrak{T}_2(F) &\leq \hat{\beta}_F^{-\frac{1}{2}} \left\| \{\sigma(\nabla w) - \pi_h^k \sigma(\nabla w)\}_F \right\|_{L^2(F)^d} \hat{\beta}_F^{\frac{1}{2}} \|[v_h]_F\|_{L^2(F)} \\ &\lesssim \hat{\beta}_F^{-\frac{1}{2}} h_F^r |\sigma(\nabla w)|_{H^{r+\frac{1}{2}}(\mathcal{T}_F)^d} \hat{\beta}_F^{\frac{1}{2}} \|[v_h]_F\|_{L^2(F)} \\ &\lesssim \hat{K}_F(w)^{-\frac{1}{2}} h_F^{r+\frac{1}{2}} |\sigma(\nabla w)|_{H^{r+\frac{1}{2}}(\mathcal{T}_F)^d} \hat{\beta}_F^{\frac{1}{2}} \|[v_h]_F\|_{L^2(F)}, \end{aligned} \tag{5.12}$$

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

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

$$\begin{aligned} \mathfrak{T}_2 &\leq \delta (|[v_h]_{1,p,h}^q + \|v_h\|_{\beta,\mu,h}^2) + c(\delta) \sum_{F \in \mathcal{F}_h^a(w)} \hat{K}_F(w)^{-1} h_F^{2r+1} |\sigma(\nabla w)|_{H^{r+\frac{1}{2}}(\mathcal{T}_F)^d}^2 \\ &\quad + c(\delta) \left(\sum_{F \in \mathcal{F}_h^d(w)} h_F^{rp'} |\sigma(\nabla w)|_{W^{r,p'}(\mathcal{T}_F)}^{p'} \right)^{\frac{q'}{p'}}. \end{aligned} \tag{5.13}$$

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 (2.2) with $n = 1$ and $(x, y, z) = ([\pi_h^k w]_F, [w_h]_F, [w]_F)$, additionally using the fact that $v_h = w_h - \pi_h^k w$; we obtain that, for all positive δ ,

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

with

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

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

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

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

$$\begin{aligned} \mathfrak{T}_{3,\text{err}}(F) &\leq h_F^{1-p} \|[[\pi_h^k w]_F]^{p-2}\|_{L^\infty(F)} \|[\pi_h^k w - w]_F\|_{L^2(F)}^2 \\ &\stackrel{(4.6)}{\leq} \frac{\hat{K}_F(w)}{h_F} \|[\pi_h^k w - w]_F\|_{L^2(F)}^2 \lesssim \hat{\beta}_F h_F^{2r+1} |w|_{H^{r+1}(\mathcal{T}_F)}^2, \end{aligned}$$

where the conclusion follows using the fact that $\text{Pe}_F^{-1}(w) < 1$ 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 :

$$\begin{aligned} \mathfrak{T}_3 &\leq \delta(s_h(w_h, v_h) - s_h(\pi_h^k w, v_h)) \\ &\quad + c(\delta) \sum_{F \in \mathcal{F}_h^a(w)} \hat{\beta}_F h_F^{2r+1} |w|_{H^{r+1}(\mathcal{T}_F)}^2 + c(\delta) \sum_{F \in \mathcal{F}_h^d(w)} h_F^{rp} |w|_{W^{r+1,p}(\mathcal{T}_F)}^p. \end{aligned} \tag{5.14}$$

Conclusion. Plugging (5.10), (5.13) and (5.14) into (5.4) and recalling that, in each of these estimates, $\delta > 0$ is arbitrary, the conclusion follows. \square

Remark 5.1. (Negative power of \hat{K}_F) As already mentioned, the negative power of \hat{K}_F appearing in bound (5.3) is balanced by the associated σ regularity term. The \hat{K}_F^{-1} stems from Eq. (5.12). The fact that the associated term $\mathfrak{T}_2(F)$, $F \in \mathcal{F}_h^a(w)$, cannot lead to an arbitrarily large contribution to the error becomes clear by bounding such term as in (5.11) instead of (5.12) (i.e. using the diffusive part of the norm instead of the advective one). Here, we decided to use (5.12) 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 5.3. (Estimate of the discrete advection–reaction error) *Let $w \in W_0^{1,p}(\Omega)$ and define the advection–reaction error linear form $\mathcal{E}_{b,h}^k(w; v_h) : \mathcal{P}^k(\mathcal{T}_h) \rightarrow \mathbb{R}$ such that, for all $v_h \in \mathcal{P}^k(\mathcal{T}_h)$,*

$$\mathcal{E}_{b,h}^k(w; v_h) := \int_{\Omega} \nabla \cdot (\beta w) v_h + \int_{\Omega} \mu w v_h - b_h(\pi_h^k w, v_h). \tag{5.15}$$

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 $F \in \mathcal{F}_h^d(w)$ for some $r \in \{0, \dots, k\}$. Then, with q as in (4.9), it holds, for any $w_h \in \mathcal{P}^k(\mathcal{T}_h)$ and any real number $\delta > 0$,

$$\begin{aligned} &\mathcal{E}_{b,h}^k(w; w_h - \pi_h^k w) \\ &\leq \delta(|w_h - \pi_h^k w|_{1,p,h}^q + \|w_h - \pi_h^k w\|_{\beta,\mu,h}^2) \\ &\quad + c(\delta) \left(\sum_{T \in \mathcal{T}_h} \hat{\tau}_T^{-2} \underline{\mu}_T^{-1} h_T^{2(r+1)} |w|_{H^{r+1}(T)}^2 + \sum_{F \in \mathcal{F}_h^a(w)} \hat{\beta}_F h_F^{2r+1} |w|_{H^{r+1}(\mathcal{T}_F)}^2 \right) \\ &\quad + c(\delta) \left(\sum_{F \in \mathcal{F}_h^d(w)} \hat{K}_F(w)^{p'} h_F^{rp'} |w|_{W^{r+1,p'}(\mathcal{T}_F)}^{p'} \right)^{\frac{q'}{p'}}, \end{aligned} \tag{5.16}$$

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

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

$$v_h := w_h - \pi_h^k w.$$

Using (3.2) with $(\tau, v) = (\beta w, v_h)$ to integrate by parts the first term in the right-hand side of (5.15) along with $\nabla \cdot \beta = 0$, recalling the single-valuedness of $\beta \cdot n_F$ and $(\beta \cdot n_F)w$ across any interface $F \in \mathcal{F}_h^i$ 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:

$$\begin{aligned} & \mathcal{E}_{b,h}^k(w; v_h) \\ &= - \int_{\Omega} (w - \pi_h^k w)(\beta \cdot \nabla_h v_h) + \int_{\Omega} \mu(w - \pi_h^k w)v_h \\ &+ \sum_{F \in \mathcal{F}_h} \int_F (\beta \cdot n_F) \{w - \pi_h^k w\}_F [v_h]_F + \frac{1}{2} \sum_{F \in \mathcal{F}_h} \hat{\beta}_F \int_F [w - \pi_h^k w]_F [v_h]_F. \\ &=: \mathfrak{T}_1 + \dots + \mathfrak{T}_4. \end{aligned} \tag{5.17}$$

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

$$\begin{aligned} \mathfrak{T}_1 &= \sum_{T \in \mathcal{T}_h} \int_T (w - \pi_T^k w)[(\beta - \pi_T^0 \beta) \cdot \nabla v_T] \\ &\leq \sum_{T \in \mathcal{T}_h} \|w - \pi_T^k w\|_{L^2(T)} \|\beta - \pi_T^0 \beta\|_{L^\infty(T)^d} \|\nabla v_T\|_{L^2(T)^d} \\ &\lesssim \sum_{T \in \mathcal{T}_h} h_T^{r+1} |w|_{H^{r+1}(T)} h_T |\beta|_{W^{1,\infty}(T)^d} h_T^{-1} \|v_T\|_{L^2(T)} \\ &\leq \left(\sum_{T \in \mathcal{T}_h} \hat{\tau}_T^{-2} \underline{\mu}_T^{-1} h_T^{2(r+1)} |w|_{H^{r+1}(T)}^2 \right)^{\frac{1}{2}} \|v_h\|_{\beta,\mu,h} \\ &\leq \delta \|v_h\|_{\beta,\mu,h}^2 + c(\delta) \sum_{T \in \mathcal{T}_h} \hat{\tau}_T^{-2} \underline{\mu}_T^{-1} h_T^{2(r+1)} |w|_{H^{r+1}(T)}^2, \end{aligned} \tag{5.18}$$

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 (4.5) 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

$$\begin{aligned}
 \mathfrak{T}_2 &\lesssim \sum_{T \in \mathcal{T}_h} \|\mu\|_{L^\infty(T)}^{\frac{1}{2}} h_T^{r+1} |w|_{H^{r+1}(T)} \left\| \mu^{\frac{1}{2}} v_T \right\|_{L^2(T)} \\
 &\stackrel{(4.5)}{\lesssim} \sum_{T \in \mathcal{T}_h} \hat{\tau}_T^{-\frac{1}{2}} h_T^{r+1} |w|_{H^{r+1}(T)} \left\| \mu^{\frac{1}{2}} v_T \right\|_{L^2(T)} \\
 &\leq \left(\sum_{T \in \mathcal{T}_h} \hat{\tau}_T^{-2} \underline{\mu}_T^{-1} h_T^{2(r+1)} |w|_{H^{r+1}(T)}^2 \right)^{\frac{1}{2}} \|v_h\|_{\beta, \mu, h} \\
 &\leq \delta \|v_h\|_{\beta, \mu, h}^2 + c(\delta) \sum_{T \in \mathcal{T}_h} \hat{\tau}_T^{-2} \underline{\mu}_T^{-1} h_T^{2(r+1)} |w|_{H^{r+1}(T)}^2, \tag{5.19}
 \end{aligned}$$

where we have used a discrete Cauchy–Schwarz inequality on the sum over $T \in \mathcal{T}_h$, noticed that $\hat{\tau}_T^{-1} \leq \hat{\tau}_T^{-2} \underline{\mu}_T^{-1}$, recalled the definition (4.7) 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 contributions on diffusion- and advection–dominated faces.

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

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

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 (3.1) to write $1 = h_F^{\frac{1}{p'}} h_F^{\frac{1-p}{p}}$ in the second inequality, the definition (4.6) of the local Péclet number in the equality and the fact that $\text{Pe}_F(w) \leq 1$ to conclude.

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

$$\begin{aligned}
 |\mathfrak{T}_{3+4}(F)| &\lesssim \hat{\beta}_F^{\frac{1}{2}} (\| \{w - \pi_h^k w\}_F \|_{L^2(F)} + \| [w - \pi_h^k w]_F \|_{L^2(F)}) \hat{\beta}_F^{\frac{1}{2}} \| [v_h]_F \|_{L^2(F)} \\
 &\lesssim \hat{\beta}_F^{\frac{1}{2}} h_F^{r+\frac{1}{2}} |w|_{H^{r+1}(\mathcal{T}_F)} \hat{\beta}_F^{\frac{1}{2}} \| [v_h]_F \|_{L^2(F)}, \tag{5.20}
 \end{aligned}$$

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

$$\begin{aligned}
 |\mathfrak{T}_3 + \mathfrak{T}_4| &\leq \delta(|v_h|_{1,p,h}^q + \|v_h\|_{\beta,\mu,h}^2) + c(\delta) \sum_{F \in \mathcal{F}_h^a(w)} \hat{\beta}_F h_F^{2r+1} |w|_{H^{r+1}(\mathcal{T}_F)}^2 \\
 &\quad + c(\delta) \left(\sum_{F \in \mathcal{F}_h^d(w)} \hat{K}_F(w)^{p'} h_F^{rp'} |w|_{W^{r+1,p'}(\mathcal{T}_F)}^{p'} \right)^{\frac{q'}{p'}}. \tag{5.21}
 \end{aligned}$$

Conclusion. Plugging (5.18), (5.19) and (5.21) into (5.17) and recalling that, in each of these estimates, $\delta > 0$ is arbitrary, the conclusion follows. \square

Remark 5.2. (Comparison with conforming finite elements) As already discussed at the end of Sec. 4.2, for $p > 2$ the bound (4.10) compares unfavorably with the conforming FE case in diffusion-dominated cases. The reason is that the terms \mathfrak{T}_2 for the diffusive part, cf. (5.13), and terms $\mathfrak{T}_3 + \mathfrak{T}_4$ for the advective part, cf. (5.21), behave as $\mathcal{O}(h^{rp'})$ for diffusion-dominated faces (instead of $\mathcal{O}(h^{2r})$). One could slightly improve such bounds by the following observations. The polynomial approximation estimate in (5.11) 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 (5.20), also in diffusion-dominated cases, thus avoiding the $\mathcal{O}(h^{rp'})$ term. Indeed note that, due to the presence of $\hat{\beta}_F$ in $\mathfrak{T}_{3+4}(F)$, the bound (5.20) does not need any assumption on dominant advection. The above modifications would lead to an $\mathcal{O}(h^{(r+1)p'})$ right-hand side for $p > 2$ in diffusion-dominated cases.

5.3. Proof of Theorem 4.1

We start by writing

$$\begin{aligned}
 C_a \|u_h - \pi_h^k u\|_{1,p,h}^q + \|u_h - \pi_h^k u\|_{\beta,\mu,h}^2 \\
 \stackrel{(5.1),(4.8)}{\leq} a_h(u_h, u_h - \pi_h^k u) - a_h(\pi_h^k u, u_h - \pi_h^k u) + b_h(u_h - \pi_h^k u, u_h - \pi_h^k u) \\
 \stackrel{(5.2),(5.15)}{=} \mathcal{E}_{a,h}^k(u; u_h - \pi_h^k u) + \mathcal{E}_{b,h}^k(u; u_h - \pi_h^k u) =: \mathcal{E}_h^k(u; u_h - \pi_h^k u), \tag{5.22}
 \end{aligned}$$

where we also used (2.1) and (4.3) in order to derive the last identity. For any real number $\delta > 0$, it holds

$$\begin{aligned}
 \mathcal{E}_h^k(u; u_h - \pi_h^k u) \\
 \stackrel{(5.3),(5.16),(4.8)}{\leq} \delta(a_h(u_h, u_h - \pi_h^k u) - a_h(\pi_h^k u, u_h - \pi_h^k u)) \\
 \quad + \delta b_h(u_h - \pi_h^k u, u_h - \pi_h^k u)
 \end{aligned}$$

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

where $c(\delta)$ denotes the largest value between (5.3) and (5.16), while E_h^k gathers all the terms multiplied by $c(\delta)$ in the sum of the right-hand sides of (5.3) and (5.16). For any $\delta < 1$, this gives

$$\mathcal{E}_h^k(u; u_h - \pi_h^k u) \leq \frac{2\delta}{1-\delta}|u_h - \pi_h^k u|_{1,p,h}^q + \frac{\delta}{1-\delta}\|u_h - \pi_h^k u\|_{\beta,\mu,h}^2 + \frac{c(\delta)}{1-\delta}E_h^k. \tag{5.23}$$

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 (5.23) with $\delta = \delta_\epsilon$ into (5.22), 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 the 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 4.1 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 mesh size. 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 $\epsilon = 10^{-10}$ and we take the initial guess as the discrete solution of the problem with $p = 2$.

6.1. Example 1

In this example, we consider the problem (2.1) with the following exact solution, velocity field and reaction terms:

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

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,

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

with q defined by (4.9).

In Figs. 1 and 2, we show, respectively, the convergence graphs for the diffusion-dominated and advection-dominated cases. The numbers appearing in the yellow boxes, directly on the graph segments in our plots, represent the reduction rates associated to two subsequent errors, i.e.,

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

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 $\mathcal{O}(h^k)$ instead of $\mathcal{O}(h^{\frac{kp}{2}})$, while, for $p > 2$, the error decreases at a rate of $\mathcal{O}(h^{\frac{kp}{2}})$ instead of $\mathcal{O}(h^{\frac{kp'}{2}})$. In both cases, the reduction rate corresponds to that obtained by the best approximation of 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 $\mathcal{O}(h^{k+\frac{1}{2}})$ 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 the problem (2.1) 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 set $\nu = 1$, $\beta = 0$ and $\mu = 0$ (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 Ref. 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 coordinate point $(0.5, 0.5)$ for the exponential solution, and in

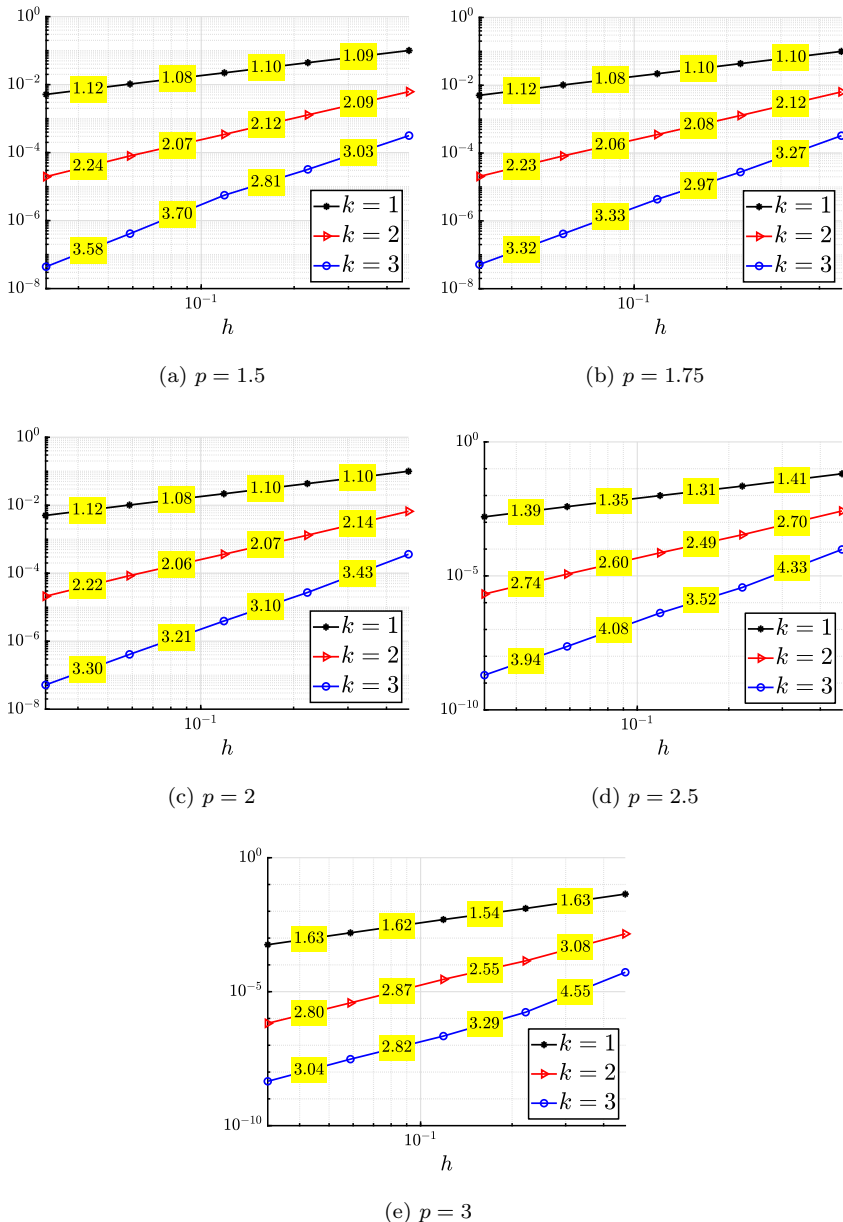


Fig. 1. Example of Sec. 6.1. Convergence rates of ERR_h in the diffusion-dominated regime. Theoretical convergence rates: $\frac{kp}{2}$ for $p \leq 2$ and $\frac{kp'}{2}$ for $p > 2$.

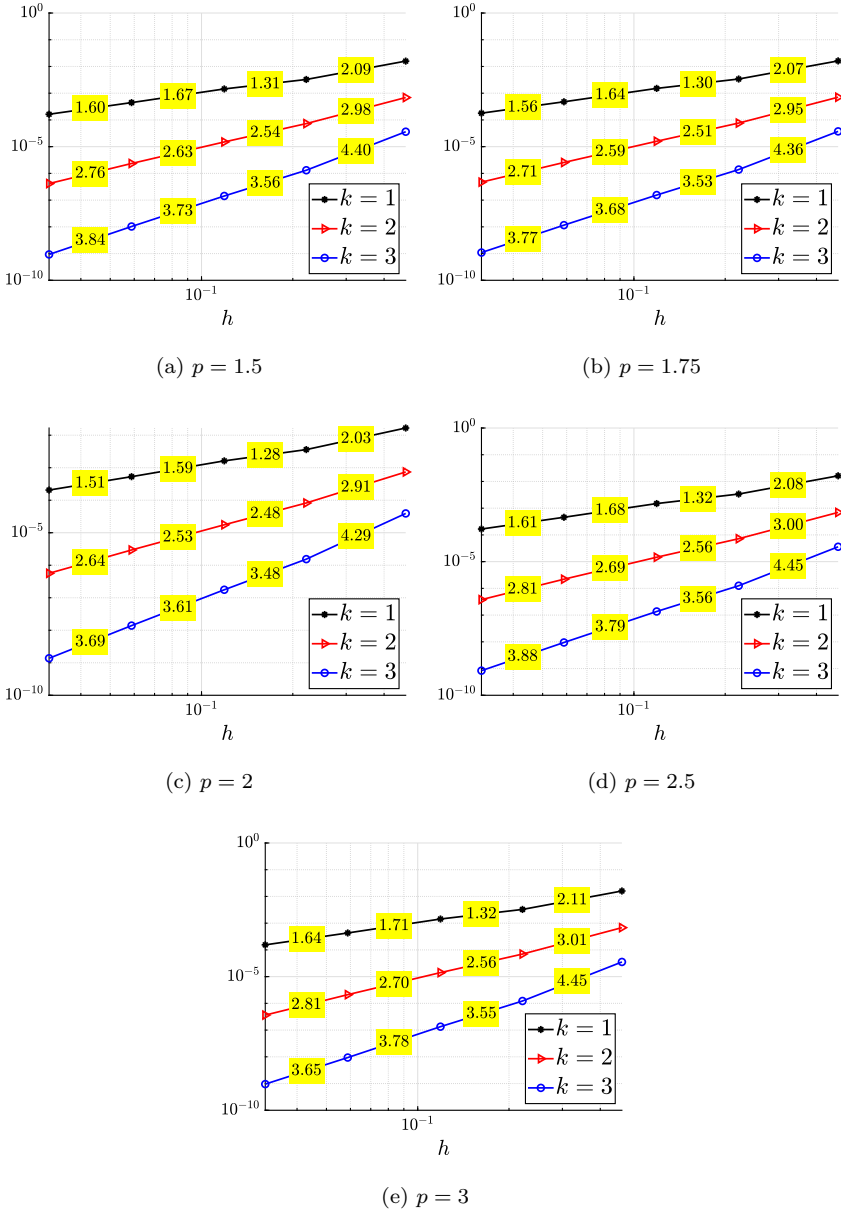


Fig. 2. Example of Sec. 6.1. Convergence rates of ERR_h in the advection-dominated regime. Theoretical convergence rate: $k + \frac{1}{2}$ for all p .

the region $\{(x, y) \in \Omega \mid x = \frac{1}{2} \text{ or } y = \frac{1}{2}\}$ for the polynomial solution, while, in the previous case, the solution had a nonzero gradient over the entire domain (which may determine a favorable situation for $p < 2$, see, for instance, Ref. 19). In this respect, the polynomial solution can be more challenging than the exponential one, as will be confirmed by the following results. Furthermore, the coarsest mesh, with

mesh size $h = 0.4714$, is removed and replaced by two new meshes obtained by halving subsequently the finest mesh: this results in two new mesh sizes 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 4.1 to hold.

As in the previous example, we compute the error term ERR_h (in this case with $\nu = 1$, $\beta = 0$ and $\mu = 0$) 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 Fig. 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 behavior is still pre-asymptotic, as is the case for the HHO method on meshes of similar size (cf., in particular, Table 4 of Ref. 19).

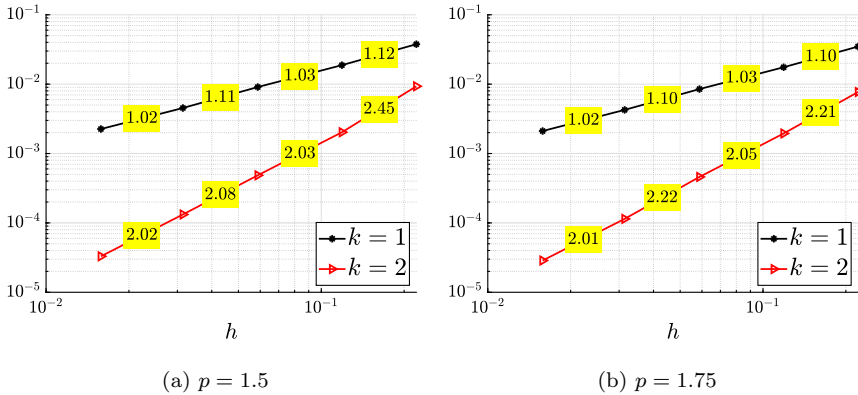


Fig. 3. Example of Sec. 6.2 with exponential solution. Convergence rates of ERR_h . Theoretical convergence rate: $\frac{kp}{2}$ for $p \in \{1.5, 1.75\}$.

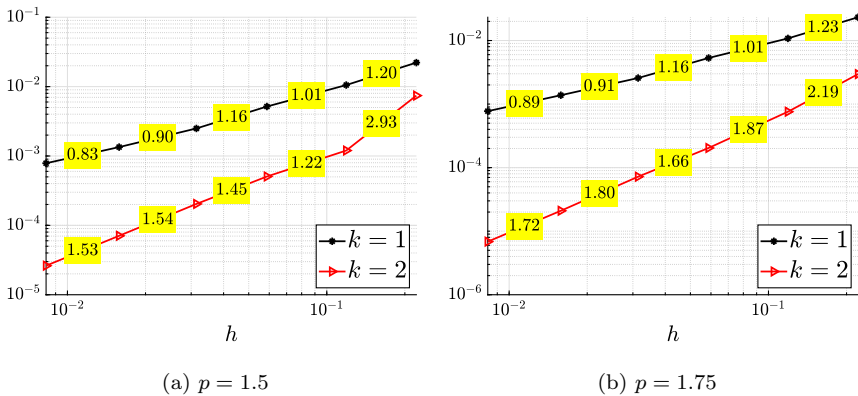



Fig. 4. Example of Sec. 6.2 with polynomial solution. Convergence rates of ERR_h . Theoretical convergence rate: $\frac{kp}{2}$ for $p \in \{1.5, 1.75\}$.


On the other hand, the results in Fig. 4, showing the convergence rates for the polynomial solution, are aligned with the expected convergence rate in the light of Theorem 4.1. Indeed, an $O(h^{\frac{hp}{2}})$ decay of ERR_h can be observed (especially for the finer meshes), which confirms from the practical side the sharpness of the theoretical results.


Acknowledgments

The present results were partially supported by the European Union (ERC Synergy, NEMESIS, Project No. 101115663). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency.

ORCID

Lourenço Beirão da Veiga  <https://orcid.org/0000-0001-5895-469X>

Daniele A. Di Pietro  <https://orcid.org/0000-0003-0959-8830>

Kirubell B. Haile  <https://orcid.org/0009-0007-2941-0926>

References

1. P. F. Antonietti, S. Giani and P. Houston, *hp*-version composite discontinuous Galerkin methods for elliptic problems on complicated domains, *SIAM J. Sci. Comput.* **35** (2013) A1417–A1439.
2. D. N. Arnold, F. Brezzi, B. Cockburn and L. D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems, *SIAM J. Numer. Anal.* **39** (2001) 1749–1779.
3. B. Ayuso and L. D. Marini, Discontinuous Galerkin methods for advection–diffusion–reaction problems, *SIAM J. Numer. Anal.* **47** (2009) 1391–1420.
4. G. A. Baker, Finite element methods for elliptic equations using nonconforming elements, *Math. Comput.* **31** (1977) 45–49.
5. J. Barrett and W. Liu, Finite element approximation of the *p*-Laplacian, *Math. Comput.* **61** (1993) 523–537.
6. F. Bassi, L. Botti and A. Colombo, Agglomeration-based physical frame dG discretizations: An attempt to be mesh free, *Math. Models Methods Appl. Sci.* **24** (2014) 1495–1539.
7. F. Bassi, L. Botti, A. Colombo, D. A. Di Pietro and P. Tesini, On the flexibility of agglomeration based physical space discontinuous Galerkin discretizations, *J. Comput. Phys.* **231** (2012) 45–65.
8. L. Beirão da Veiga, F. Dassi, C. Lovadina and G. Vacca, SUPG-stabilized virtual elements for diffusion-convection problems: A robustness analysis, *ESAIM: M2AN* **55** (2021) 2233–2258.
9. L. Beirão da Veiga, F. Dassi and G. Vacca, Vorticity-stabilized virtual elements for the Oseen equation, *Math. Models Methods Appl. Sci.* **31** (2021) 3009–3052.
10. L. Beirão da Veiga, F. Dassi and G. Vacca, Pressure robust SUPG-stabilized finite elements for the unsteady Navier–Stokes equation, *IMA J. Numer. Anal.* **44** (2023) 710–750.

11. F. Brezzi, L. D. Marini and E. Süli, Discontinuous Galerkin methods for first-order hyperbolic problems, *Math. Models Methods Appl. Sci.* **14** (2004) 1893–1903.
12. E. Burman and A. Ern, Discontinuous Galerkin approximation with discrete variational principle for the nonlinear Laplacian, *C.R. Math. Acad. Sci. Paris* **346** (2008) 1013–1016.
13. P. Castillo, B. Cockburn, I. Perugia and D. Schötzau, An *a priori* error analysis of the local discontinuous Galerkin method for elliptic problems, *SIAM J. Numer. Anal.* **38** (2000) 1676–1706.
14. P. Castillo, B. Cockburn, I. Perugia and D. Schötzau, Local discontinuous Galerkin methods for elliptic problems, *Commun. Numer. Methods Eng.* **18** (2002) 69–75.
15. B. Cockburn and C.-W. Shu, The Runge–Kutta local projection P^1 -discontinuous-Galerkin finite element method for scalar conservation laws, *ESAIM: Modél. Math. Anal. Numér.* **25** (1991) 337–361.
16. L. M. Del Pezzo, A. L. Lombardi and S. Martínez, Interior penalty discontinuous Galerkin FEM for the $p(x)$ -Laplacian, *SIAM J. Numer. Anal.* **50** (2012) 2497–2521.
17. D. A. Di Pietro and J. Droniou, *The Hybrid High-Order Method for Polytopal Meshes: Design, Analysis, and Applications*, Modeling, Simulation and Application, Vol. 19 (Springer International Publishing, 2020).
18. D. A. Di Pietro and J. Droniou, A polytopal method for the Brinkman problem robust in all regimes, *Comput. Methods Appl. Mech. Engg.* **409** (2023) 115981.
19. D. A. Di Pietro, J. Droniou and A. Harnist, Improved error estimates for Hybrid High-Order discretizations of Leray–Lions problems, *Calcolo* **58** (2021) 19.
20. D. A. Di Pietro and A. Ern, Discrete functional analysis tools for discontinuous Galerkin methods with application to the incompressible Navier–Stokes equations, *Math. Comput.* **79** (2010) 1303–1330.
21. D. A. Di Pietro and A. Ern, *Mathematical Aspects of Discontinuous Galerkin Methods*, Mathématiques et Applications, Vol. 69 (Springer, 2012).
22. D. A. Di Pietro, A. Ern and J.-L. Guermond, Discontinuous Galerkin methods for anisotropic semi-definite diffusion with advection, *SIAM J. Numer. Anal.* **46** (2008) 805–831.
23. L. Diening and F. ETTWEIN, Fractional estimates for non-differentiable elliptic systems with general growth, *Forum Math.* **20** (2008) 523–556.
24. L. Diening, D. Köner, M. Ružička and I. Touloupoulos, A local discontinuous Galerkin approximation for systems with p -structure, *IMA J. Numer. Anal.* **34** (2014) 1447–1488.
25. 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, *IMA J. Numer. Anal.* **42** (2021) 1568–1597.
26. A. Hirn, Approximation of the p -Stokes equations with equal-order finite elements, *J. Math. Fluid Mech.* **15** (2013) 65–88.
27. J. Leray and J.-L. Lions, Quelques résultats de Višik sur les problèmes elliptiques nonlinéaires par les méthodes de Minty-Browder, *Bull. Soc. Math. Fr.* **93** (1965) 97–107.
28. T. Malkmus, M. Ružička, S. Eckstein and I. Touloupoulos, Generalizations of SIP methods to systems with p -structure, *IMA J. Numer. Anal.* **38** (2018) 1420–1451.
29. W. H. Reed and T. R. Hill, Triangular mesh methods for the neutron transport equation, Technical Report No. LA-UR-73-0479/CONF-730414-2, Los Alamos Scientific Laboratory, Los Alamos, NM (1973).