Improved error estimates for Hybrid High-Order discretizations of Leray-Lions problems

Daniele Antonio Di Pietro, Jérôme Droniou, André Harnist

Improved error estimates for Hybrid High-Order discretizations of Leray–Lions problems

Daniele A. Di Pietro [*1], Jérôme Droniou [†2], and André Harnist [‡1]

[1]IMAG, Univ Montpellier, CNRS, Montpellier, France
[2]School of Mathematics, Monash University, Melbourne, Australia

June 4, 2021

Abstract

We derive novel error estimates for Hybrid High-Order (HHO) discretizations of Leray–Lions problems set in $W^{1,p}$ with $p \in (1,2]$. Specifically, we prove that, depending on the degeneracy of the problem, the convergence rate may vary between $(k+1)(p-1)$ and $(k+1)$, with k denoting the degree of the HHO approximation. These regime-dependent error estimates are illustrated by a complete panel of numerical experiments.

Keywords: Hybrid High-Order methods, degenerate Leray–Lions problems, regime-dependent error estimates

MSC2018 classification: 65N08, 65N30, 65N12

1 Introduction

We consider Hybrid High-Order (HHO) approximations of Leray–Lions problems set in $W^{1,p}$ with $p \in (1,2]$. For this class of problems, negative powers of the gradient of the solution can appear in the flux. Depending on the expression of the latter, this can lead to a (local) degeneracy of the problem when the gradient of the solution vanishes or becomes large.

In this work, we prove novel error estimates that highlight the dependence of the convergence rate on the two possible cases of degeneracy simultaneously, and we do not differentiate these two degeneracies any longer. Specifically, we show that, for the globally non-degenerate case, the energy-norm of the error converges as h^{k+1}, with h denoting the mesh size and k the degree of the HHO approximation. In the globally degenerate case, on the other hand, the energy-norm of the error converges as $h^{(k+1)(p-1)}$, coherently with the estimate originally proved in [8, Theorem 3.2]. We additionally introduce, for each mesh element T of diameter h_T, a dimensionless number η_T that captures the local degeneracy of the model in T and identifies the contribution of the element to the global error: from the fully degenerate regime, corresponding to a contribution in $O(h_T^{(k+1)(p-1)})$, to the non-degenerate regime, corresponding to a contribution in $O(h_T^{k+1})$, through all intermediate regimes. Estimates depending on local regimes have been established in linear settings (see, e.g., [10, 11] for advection–diffusion–reaction models and [5] for the Brinkman problem). Such work has also been done for the Leray–Lions problems with Adaptative Finite Element methods in [4]; to the best of our knowledge, their extension to HHO methods is entirely new.

Error estimates for the lowest-order conforming finite element approximation of the pure p-Laplacian have been known for quite some time; see, e.g., the founding work [12], in which $O(h^{1/(3-p)})$ error

[*]daniele.di-pietro@umontpellier.fr
[†]jerome.droniou@monash.edu
[‡]andre.harnist@umontpellier.fr, corresponding author

estimates are obtained in the case $p \leq 2$ considered here. These estimates were later improved in [2] to $O(h)$, for solutions of high (and global) regularity – in the space $W^{3,1}(\Omega) \cap C^{2,(2-p)/p}(\overline{\Omega})$. The above results have been extended to non-conforming finite elements in [16]. A glaciology model is considered in [13], corresponding to a non-degenerate p-Laplace equation (with flux satisfying (2) below with $\delta = 1$), and error estimates for the conforming finite element approximation have been obtained: $O(h)$ if the solution is in $H^2(\Omega)$, and $O(h^{p/2})$ if it belongs to $W^{2,p}(\Omega)$. A common feature of all these studies, in which sharp error estimates are derived (which do not degrade too much as p gets far from 2), is that they only consider low-order schemes on 2D triangular meshes and with continuity properties – either all along the edges for the conforming method, or at the edges midpoints for the non-conforming method. To our knowledge, for higher-order methods that may involve fully discontinuous functions and are applicable to generic polytopal meshes, such as HHO, no sharp error estimates are known and only convergence in $h^{(k+1)(p-1)}$ has been established so far. This paper therefore bridges a gap between the results available for the low-order finite element methods and HHO methods. Notice that, very recently, $O(h^{(k+1)/(3-p)})$ error estimates have been obtained in [**carstensen:20**] for an HHO method on standard simplicial meshes based on a stable gradient inspired by [10]. In passing, even though we focus on the HHO method, our approach could in all likelihood be extended to other polytopal methods such as, e.g., the Mimetic Finite Difference method [1] or the Virtual Element method [3]; see the preface of [9] for an up-to-date literature review on this subject.

The rest of the paper is organized as follows. In Section 2 we establish the continuous setting, including novel assumptions on the flux function weaker than the ones considered in [8, Section 3.1]. In Section 3 we briefly recall the discrete setting upon which rests the HHO scheme described in Section 4. The main result of this paper is contained Section 5. Finally, Section 6 contains a complete panel of numerical tests illustrating the effect of local degeneracy on the convergence rate.

2 Continuous setting

2.1 Flux function

Let $\Omega \subset \mathbb{R}^d$, $d \in \mathbb{N}^*$, denote a bounded, connected, polytopal open set with Lipschitz boundary $\partial\Omega$. We consider the Leray–Lions problem, which consist in finding $u : \Omega \to \mathbb{R}$ such that

$$-\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla}u) = f \qquad \text{in } \Omega, \tag{1a}$$

$$u = 0 \qquad \text{on } \partial\Omega, \tag{1b}$$

where $f : \Omega \to \mathbb{R}$ represents a volumetric force term, while $\boldsymbol{\sigma} : \Omega \times \mathbb{R}^d \to \mathbb{R}^d$ is the *flux function*. The flux function is possibly variable in space and depends on the *potential* $u : \Omega \to \mathbb{R}$ only through its gradient. The following assumptions characterize $\boldsymbol{\sigma}$.

Assumption 1 (Flux function)**.** Let a real number $p \in (1, 2]$ be fixed and denote by

$$p' \coloneqq \frac{p}{p-1} \in [2, +\infty)$$

the conjugate exponent of p. The flux function satisfies

$$\boldsymbol{\sigma}(\boldsymbol{x}, \mathbf{0}) = \mathbf{0} \text{ for almost every } \boldsymbol{x} \in \Omega, \tag{2a}$$

$$\boldsymbol{\sigma}(\cdot, \boldsymbol{\xi}) : \Omega \to \mathbb{R}^d \text{ is measurable for all } \boldsymbol{\xi} \in \mathbb{R}^d. \tag{2b}$$

Moreover, there exist a *degeneracy function* $\delta \in L^p(\Omega, [0, +\infty))$ and two real numbers $\sigma_{\mathrm{hc}}, \sigma_{\mathrm{sm}} \in (0, +\infty)$ such that, for all $\boldsymbol{\tau}, \boldsymbol{\eta} \in \mathbb{R}^d$ and almost every $\boldsymbol{x} \in \Omega$, we have the *continuity* property

$$|\boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{\tau}) - \boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{\eta})| \leq \sigma_{\mathrm{hc}} \left(\delta(\boldsymbol{x})^p + |\boldsymbol{\tau}|^p + |\boldsymbol{\eta}|^p\right)^{\frac{p-2}{p}} |\boldsymbol{\tau} - \boldsymbol{\eta}|, \tag{2c}$$

and the *strong monotonicity* property

$$(\boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{\tau}) - \boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{\eta})) \cdot (\boldsymbol{\tau} - \boldsymbol{\eta}) \geq \sigma_{\mathrm{sm}} \left(\delta(\boldsymbol{x})^p + |\boldsymbol{\tau}|^p + |\boldsymbol{\eta}|^p\right)^{\frac{p-2}{p}} |\boldsymbol{\tau} - \boldsymbol{\eta}|^2. \tag{2d}$$

Some remarks are in order.

Remark 1 (Flux at rest). Assumption (2a) expresses the fact that the flux at rest is zero, and can be relaxed taking $\boldsymbol{\sigma}(\cdot,\mathbf{0}) \in L^{p'}(\Omega)^d$. This modification requires only minor changes in the analysis, not detailed for the sake of conciseness.

Remark 2 (Relations between the continuity and monotonicity constants). Inequalities (2c) and (2d) give

$$\sigma_{\mathrm{sm}} \le \sigma_{\mathrm{hc}}. \tag{3}$$

Indeed, let $\boldsymbol{\tau} \in \mathbb{R}^d$ be such that $|\boldsymbol{\tau}| > 0$. Using the strong monotonicity (2d) (with $\boldsymbol{\eta} = \mathbf{0}$) along with (2a), the Cauchy–Schwarz inequality, and the continuity (2c) (again with $\boldsymbol{\eta} = \mathbf{0}$) and (2a), we infer that

$$\sigma_{\mathrm{sm}} \left(\delta^p + |\boldsymbol{\tau}|^p\right)^{\frac{p-2}{p}} |\boldsymbol{\tau}|^2 \le \boldsymbol{\sigma}(\cdot,\boldsymbol{\tau}) \cdot \boldsymbol{\tau} \le |\boldsymbol{\sigma}(\cdot,\boldsymbol{\tau})||\boldsymbol{\tau}| \le \sigma_{\mathrm{hc}} \left(\delta^p + |\boldsymbol{\tau}|^p\right)^{\frac{p-2}{p}} |\boldsymbol{\tau}|^2$$

almost everywhere in Ω, hence (3).

Remark 3 (Degenerate case). We note the following inequality: For all $x, y \in \mathbb{R}^n$, $n \in \mathbb{N}^*$, and all $\alpha \in [0, +\infty)$,

$$(\alpha + |x| + |y|)^{p-2}|x - y| \le |x - y|^{p-1}. \tag{4}$$

To prove (4), notice that, if $\alpha + |x| + |y| > 0$, using a triangle inequality to write $|x| + |y| \ge |x - y|$ together with the fact that $\mathbb{R} \ni t \mapsto t^{p-2} \in \mathbb{R}$ is non-increasing (since $p < 2$) and $\alpha \ge 0$, we infer that $(\alpha + |x| + |y|)^{p-2} \le |x - y|^{p-2}$, which, multiplying by $|x - y|$, gives (4). Since (4) is valid when $\alpha + |x| + |y| > 0$, we can extend the left-hand side by continuity (with value 0) in the singular case $\alpha + |x| + |y| = 0$ and this estimate remains valid.

Inequality (29) below together with (4) ensures that properties (2c)–(2d) are well-formulated by extension also when $\delta(\boldsymbol{x})^p + |\boldsymbol{\tau}|^p + |\boldsymbol{\eta}|^p$ vanishes. As a consequence, $\boldsymbol{\sigma}(\boldsymbol{x},\cdot) : \mathbb{R}^d \to \mathbb{R}^d$ is continuous for a.e. $\boldsymbol{x} \in \Omega$. The relation (4) will also play a key role in the proof of Theorem 11 below.

Remark 4 (Non-degenerate case). In [8], an error estimate is given for broader versions of inequalities (2c)–(2d). The novelty here lies in the introduction of the degeneracy function δ, since it directly affects the convergence rate of the method. The crux of its intervention is located in the proof of Theorem 11, and more precisely at (43) where it prevents singularities. See Remark 12 for more details, see also Figure 1 for a set of numerical results illustrating this phenomenon.

Example 5 (p-Laplace flux function)**.** A typical example of flux function is $\boldsymbol{\sigma}(\boldsymbol{x},\boldsymbol{\tau}) = |\boldsymbol{\tau}|^{p-2}\boldsymbol{\tau}$, for which (1) is the p-Laplace equation $-\boldsymbol{\nabla}\cdot(|\boldsymbol{\nabla} u|^{p-2}\boldsymbol{\nabla} u) = f$. This flux function satisfies Assumption 1 with degeneracy function $\delta = 0$, see e.g. [9, Lemma 6.26].

Example 6 (Carreau–Yasuda flux function)**.** Another example of function $\boldsymbol{\sigma}$ which satisfies Assumption 1, inspired by the rheology of Carreau–Yosida fluids, is obtained setting, for almost every $\boldsymbol{x} \in \Omega$ and all $\boldsymbol{\tau} \in \mathbb{R}^d$,

$$\boldsymbol{\sigma}(\boldsymbol{x},\boldsymbol{\tau}) = \mu(\boldsymbol{x}) \left(\delta(\boldsymbol{x})^{a(\boldsymbol{x})} + |\boldsymbol{\tau}|^{a(\boldsymbol{x})}\right)^{\frac{p-2}{a(\boldsymbol{x})}} \boldsymbol{\tau}, \tag{5}$$

where $\mu : \Omega \to [\mu_-, \mu_+]$ is a measurable function with $\mu_-, \mu_+ \in (0, +\infty)$ corresponding to the *local flow consistency index*, $\delta \in L^p(\Omega, [0, +\infty))$ is the *degeneracy parameter*, $a : \Omega \to [a_-, a_+]$ is a measurable function with $a_-, a_+ \in (0, +\infty)$ expressing the *local transition flow behavior index*, and $p \in (1, 2]$ is the *flow behavior index*. It was proved in [6, Appendix A] that $\boldsymbol{\sigma}$ is a p-power-framed function (with a straightforward analogy to replace the degeneracy constant σ_{de} therein by the degeneracy function δ) with

$$\sigma_{\mathrm{hc}} = \frac{\mu_+}{p-1} 2^{\left[-\left(\frac{1}{a_+}-\frac{1}{p}\right)^{\ominus}-1\right](p-2)+\frac{1}{p}} \quad \text{and} \quad \sigma_{\mathrm{sm}} = \mu_-(p-1)2^{\left(\frac{1}{a_-}-\frac{1}{p}\right)^{\oplus}(p-2)},$$

where $\xi^{\oplus} := \max(0;\xi)$ and $\xi^{\ominus} := -\min(0;\xi)$ denote, respectively, the positive and negative parts of a real number ξ. As a consequence, the flux function (5) matches Assumption 1.

2.2 Weak formulation

We define the following space for the potential embedding the homogeneous boundary condition:

$$U \coloneqq W_0^{1,p}(\Omega).$$

Assuming $f \in L^{p'}(\Omega)$, the weak formulation of problem (1) reads: Find $u \in U$ such that

$$a(u,v) = \int_\Omega fv \qquad \forall v \in U, \tag{6}$$

where the *diffusion function* $a : U \times U \to \mathbb{R}$ is defined such that, for all $v, w \in U$,

$$a(w,v) \coloneqq \int_\Omega \boldsymbol{\sigma}(\cdot, \boldsymbol{\nabla} w) \cdot \boldsymbol{\nabla} v. \tag{7}$$

Proposition 7 (Well-posedness and a priori estimate)**.** *Under Assumption 1, the continuous problem (6) admits a unique solution $u \in U$ that satisfies the following a priori bound:*

$$\|\boldsymbol{\nabla} u\|_{L^p(\Omega)^d} \le \left(2^{\frac{2-p}{p}} C_{\mathrm{P}} \sigma_{\mathrm{sm}}^{-1} \|f\|_{L^{p'}(\Omega)}\right)^{\frac{1}{p-1}} + \min\left(\|\delta\|_{L^p(\Omega)}; 2^{\frac{2-p}{p}} C_{\mathrm{P}} \sigma_{\mathrm{sm}}^{-1} \|\delta\|_{L^p(\Omega)}^{2-p} \|f\|_{L^{p'}(\Omega)}\right). \tag{8}$$

where the real number $C_{\mathrm{P}} > 0$, only depending on Ω and on p, is such that, for all $v \in W_0^{1,p}(\Omega)$, the Poincaré inequality $\|v\|_{L^p(\Omega)} \le C_{\mathrm{P}} \|\boldsymbol{\nabla} v\|_{L^p(\Omega)^d}$ holds.

Proof. For the existence and uniqueness of a solution to (6) see, e.g., [14, Section 2.4]. To prove the a priori bound (8), use the strong monotonicity (2d) of $\boldsymbol{\sigma}$, (6) written for $v = u$, and invoke the Hölder and Poincaré inequalities to write

$$\sigma_{\mathrm{sm}} \left(\|\delta\|_{L^p(\Omega)}^p + \|\boldsymbol{\nabla} u\|_{L^p(\Omega)^d}^p\right)^{\frac{p-2}{p}} \|\boldsymbol{\nabla} u\|_{L^p(\Omega)^d}^2 \le a(u,u) = \int_\Omega fu \le C_{\mathrm{P}} \|f\|_{L^{p'}(\Omega)} \|\boldsymbol{\nabla} u\|_{L^p(\Omega)^d},$$

that is,

$$\mathcal{N} \coloneqq \left(\|\delta\|_{L^p(\Omega)}^p + \|\boldsymbol{\nabla} u\|_{L^p(\Omega)^d}^p\right)^{\frac{p-2}{p}} \|\boldsymbol{\nabla} u\|_{L^p(\Omega)^d} \le C_{\mathrm{P}} \sigma_{\mathrm{sm}}^{-1} \|f\|_{L^{p'}(\Omega)}. \tag{9}$$

Observing that $\|\boldsymbol{\nabla} u\|_{L^p(\Omega)^d} \le 2^{\frac{2-p}{p}} \max\left(\|\boldsymbol{\nabla} u\|_{L^p(\Omega)^d}; \|\delta\|_{L^p(\Omega)}\right)^{2-p} \mathcal{N}$, we obtain, enumerating the cases for the maximum and summing the corresponding bounds,

$$\|\boldsymbol{\nabla} u\|_{L^p(\Omega)^d} \le (2^{\frac{2-p}{p}} \mathcal{N})^{\frac{1}{p-1}} + 2^{\frac{2-p}{p}} \|\delta\|_{L^p(\Omega)}^{2-p} \mathcal{N}. \tag{10}$$

On the other hand, we have $\mathcal{N} \ge 2^{\frac{p-2}{p}} \|\boldsymbol{\nabla} u\|_{L^p(\Omega)^d}^{p-1}$ if $\|\boldsymbol{\nabla} u\|_{L^p(\Omega)^d} \ge \|\delta\|_{L^p(\Omega)}$. Thus we have, for any value of $\|\boldsymbol{\nabla} u\|_{L^p(\Omega)^d}$,

$$\|\boldsymbol{\nabla} u\|_{L^p(\Omega)^d} \le (2^{\frac{2-p}{p}} \mathcal{N})^{\frac{1}{p-1}} + \|\delta\|_{L^p(\Omega)}. \tag{11}$$

Combining (9) with the minimum of inequalities (10) and (11) gives (8). □

3 Discrete setting

3.1 Mesh

For any set $X \subset \mathbb{R}^d$, denote by h_X its diameter. A polytopal mesh is defined as a couple $\mathcal{M}_h \coloneqq (\mathcal{T}_h, \mathcal{F}_h)$, where $\mathcal{T}_h$ is a finite collection of polytopal elements $T \in \mathcal{T}_h$ such that $h = \max_{T\in\mathcal{T}_h} h_T$, while $\mathcal{F}_h$ is a finite collection of hyperplanar faces. It is assumed henceforth that the mesh $\mathcal{M}_h$ matches the geometrical requirements detailed in [9, Definition 1.7]. Boundary faces lying on $\partial\Omega$ and internal faces contained in Ω are collected in the sets $\mathcal{F}_h^{\mathrm{b}}$ and $\mathcal{F}_h^{\mathrm{i}}$, respectively. For every mesh element $T \in \mathcal{T}_h$, we denote by $\mathcal{F}_T$ the subset of $\mathcal{F}_h$ collecting the faces that lie on the boundary ∂T of T. For every face $F \in \mathcal{F}_h$, we denote by $\mathcal{T}_F$ the subset of $\mathcal{T}_h$ containing the one (if $F \in \mathcal{F}_h^{\mathrm{b}}$) or two (if $F \in \mathcal{F}_h^{\mathrm{i}}$) elements on whose boundary F lies. For each mesh element $T \in \mathcal{T}_h$ and face $F \in \mathcal{F}_T$, $\boldsymbol{n}_{TF}$ denotes the (constant) unit vector normal to F pointing out of T.

Our focus is on the h-convergence analysis, so we consider a sequence of refined meshes that is regular in the sense of [9, Definition 1.9], with regularity parameter uniformly bounded away from zero. The mesh regularity assumption implies, in particular, that the diameter of a mesh element and those of its faces are comparable uniformly in h, and that the number of faces of one element is bounded above by an integer independent of h.

3.2 Notation for inequalities up to a multiplicative constant

To avoid the proliferation of generic constants, we write henceforth $a \lesssim b$ (resp., $a \gtrsim b$) for the inequality $a \le Cb$ (resp., $a \ge Cb$) with real number $C > 0$ independent of h, of the parameters $\delta, \sigma_{\mathrm{hc}}, \sigma_{\mathrm{sm}}$ in Assumption 1, and, for local inequalities, of the mesh element or face on which the inequality holds. We also write $a \simeq b$ to mean $a \lesssim b$ and $b \lesssim a$. The dependencies of the hidden constants are further specified when needed.

3.3 Projectors and broken spaces

Given $X \in \mathcal{T}_h \cup \mathcal{F}_h$ and $l \in \mathbb{N}$, we denote by $\mathbb{P}^l(X)$ the space spanned by the restriction to X of scalar-valued, d-variate polynomials of total degree $\le l$. The local L^2-orthogonal projector $\pi_X^l : L^1(X) \to \mathbb{P}^l(X)$ is defined such that, for all $v \in L^1(X)$,

$$\int_X (\pi_X^l v - v) w = 0 \qquad \forall w \in \mathbb{P}^l(X). \tag{12}$$

When applied to vector-valued functions in $L^1(X)^d$, the L^2-orthogonal projector mapping on $\mathbb{P}^l(X)^d$ acts component-wise and is denoted in boldface font as $\boldsymbol{\pi}_X^l$. Let $T \in \mathcal{T}_h$, $n \in [0, l+1]$, and $m \in [0, n]$. The following (n, p, m)-approximation properties of π_T^l hold: For any $v \in W^{n,p}(T)$,

$$|v - \pi_T^l v|_{W^{m,p}(T)} \lesssim h_T^{n-m} |v|_{W^{n,p}(T)}. \tag{13a}$$

The above property will also be used in what follows with p replaced by its conjugate exponent p'. If, additionally, $n \ge 1$, we have the following (n, p')-trace approximation property:

$$\|v - \pi_T^l v\|_{L^{p'}(\partial T)} \lesssim h_T^{n-\frac{1}{p'}} |v|_{W^{n,p'}(T)}. \tag{13b}$$

The hidden constants in (13) are independent of h and T, but possibly depend on d, the mesh regularity parameter, l, n, and p. The approximation properties (13) are proved for integer n and m in [7, Appendix A.2] (see also [9, Theorem 1.45]), and can be extended to non-integer values using standard interpolation techniques (see, e.g., [15, Theorem 5.1]).

The additional regularity on the exact solution in the error estimates will be expressed in terms of the broken Sobolev spaces

$$W^{n,p}(\mathcal{T}_h) \coloneqq \left\{ v \in L^p(\Omega) \ : \ v_{|T} \in W^{n,p}(T) \quad \forall T \in \mathcal{T}_h \right\}.$$

The corresponding seminorm is such that $|v|_{W^{n,p}(\mathcal{T}_h)} \coloneqq \left(\sum_{T\in\mathcal{T}_h} |v|_{W^{n,p}(T)}^p \right)^{\frac{1}{p}}$ for all $v \in W^{n,p}(\mathcal{T}_h)$.

4 HHO discretization

4.1 Hybrid space and norms

Let an integer $k \geq 0$ be fixed. The HHO space is, with usual notation,

$$\underline{U}_h^k \coloneqq \left\{ \underline{v}_h = ((v_T)_{T\in\mathcal{T}_h}, (v_F)_{F\in\mathcal{F}_h}) \;:\; v_T \in \mathbb{P}^k(T) \;\; \forall T \in \mathcal{T}_h \;\text{ and }\; v_F \in \mathbb{P}^k(F) \;\; \forall F \in \mathcal{F}_h \right\}.$$

The interpolation operator $\underline{I}_h^k : W^{1,1}(\Omega) \to \underline{U}_h^k$ maps a function $v \in W^{1,1}(\Omega)$ on the vector of discrete unknowns $\underline{I}_h^k v$ defined as follows:

$$\underline{I}_h^k v \coloneqq ((\pi_T^k v_{|T})_{T\in\mathcal{T}_h}, (\pi_F^k v_{|F})_{F\in\mathcal{F}_h}).$$

For all $T \in \mathcal{T}_h$, we denote by $\underline{U}_T^k$ and $\underline{I}_T^k$ the restrictions of $\underline{U}_h^k$ and $\underline{I}_h^k$ to T, respectively, and, for all $\underline{v}_h \in \underline{U}_h^k$, we let $\underline{v}_T \coloneqq (v_T, (v_F)_{F\in\mathcal{F}_T}) \in \underline{U}_T^k$ denote the vector collecting the discrete unknowns attached to T and its faces. Furthermore, for all $\underline{v}_h \in \underline{U}_h^k$, we define the broken polynomial field $v_h \in \mathbb{P}^k(\mathcal{T}_h)$ obtained patching element unknowns, that is,

$$(v_h)_{|T} \coloneqq v_T \qquad \forall T \in \mathcal{T}_h.$$

For all $q \in (1, +\infty)$, we define on $\underline{U}_h^k$ the $W^{1,q}(\Omega)$-like seminorm $\|\cdot\|_{1,q,h}$ such that, for all $\underline{v}_h \in \underline{U}_h^k$,

$$\|\underline{v}_h\|_{1,q,h} \coloneqq \left(\sum_{T\in\mathcal{T}_h} \|\underline{v}_T\|_{1,q,T}^q \right)^{\frac{1}{q}} \tag{14a}$$

$$\text{with } \|\underline{v}_T\|_{1,q,T} \coloneqq \left(\|\boldsymbol{\nabla} v_T\|_{L^q(T)^d}^q + \sum_{F\in\mathcal{F}_T} h_F^{1-q} \|v_F - v_T\|_{L^q(F)}^q \right)^{\frac{1}{q}} \text{ for all } T \in \mathcal{T}_h. \tag{14b}$$

The following boundedness property for $\underline{I}_T^k$ is proved in [9, Proposition 6.24]: For all $T \in \mathcal{T}_h$ and all $v \in W^{1,p}(T)$,

$$\|\underline{I}_T^k v\|_{1,p,T} \lesssim |v|_{W^{1,p}(T)}, \tag{15}$$

where the hidden constant depends only on d, the mesh regularity parameter, p, and k.

The discrete potential is sought in the subspace of $\underline{U}_h^k$ embedding the homogeneous boundary condition:

$$\underline{U}_{h,0}^k \coloneqq \left\{ \underline{v}_h = ((v_T)_{T\in\mathcal{T}_h}, (v_F)_{F\in\mathcal{F}_h}) \in \underline{U}_h^k \;:\; v_F = 0 \quad \forall F \in \mathcal{F}_h^{\mathrm{b}} \right\}.$$

The following discrete Poincaré inequality descends from [7, Proposition 5.4] (cf. Remark 5.5 therein): For all $\underline{v}_h \in \underline{U}_{h,0}^k$,

$$\|v_h\|_{L^p(\Omega)} \lesssim \|\underline{v}_h\|_{1,p,h}. \tag{16}$$

By virtue of this inequality, $\|\cdot\|_{1,p,h}$ is a norm on $\underline{U}_{h,0}^k$ (reason as in [9, Corollary 2.16]).

4.2 Reconstructions

For all $T \in \mathcal{T}_h$, we define the *local gradient reconstruction* $\mathbf{G}_T^k : \underline{U}_T^k \to \mathbb{P}^k(T)^d$ such that, for all $\underline{v}_T \in \underline{U}_T^k$,

$$\int_T \mathbf{G}_T^k \underline{v}_T \cdot \boldsymbol{\tau} = \int_T \boldsymbol{\nabla} v_T \cdot \boldsymbol{\tau} + \sum_{F\in\mathcal{F}_T} \int_F (v_F - v_T)\, (\boldsymbol{\tau} \cdot \boldsymbol{n}_{TF}) \qquad \forall \boldsymbol{\tau} \in \mathbb{P}^k(T)^d. \tag{17}$$

By design, the following relation holds (see [9, Section 7.2.5]): For all $v \in W^{1,1}(T)$,

$$\mathbf{G}_T^k(\underline{I}_T^k v) = \boldsymbol{\pi}_T^k(\boldsymbol{\nabla} v). \tag{18}$$

The *local potential reconstruction* $\mathrm{r}_T^{k+1} : \underline{U}_T^k \to \mathbb{P}^{k+1}(T)$ is such that, for all $\underline{v}_T \in \underline{U}_T^k$,

$$\int_T (\boldsymbol{\nabla} \mathrm{r}_T^{k+1} \underline{v}_T - \mathbf{G}_T^k \underline{v}_T) \cdot \boldsymbol{\nabla} w = 0 \text{ for all } w \in \mathbb{P}^{k+1}(T) \text{ and } \int_T \mathrm{r}_T^{k+1} \underline{v}_T = \int_T v_T. \tag{19}$$

Composed with the local interpolator, this reconstruction commutes with the elliptic projector; see [9, Sections 1.3 and 2.1.1–2.1.3].

4.3 Discrete diffusion function

The *discrete diffusion function* $\mathrm{a}_h : \underline{U}_h^k \times \underline{U}_h^k \to \mathbb{R}$, discretizing the function a defined by (7), is such that, for all $\underline{v}_h, \underline{w}_h \in \underline{U}_h^k$,

$$\mathrm{a}_h(\underline{w}_h, \underline{v}_h) \coloneqq \sum_{T \in \mathcal{T}_h} \left(\int_T \boldsymbol{\sigma}(\cdot, \mathbf{G}_T^k \underline{w}_T) \cdot \mathbf{G}_T^k \underline{v}_T + \mathrm{s}_T(\underline{w}_T, \underline{v}_T) \right). \tag{20}$$

Above, for all $T \in \mathcal{T}_h$, $\mathrm{s}_T : \underline{U}_T^k \times \underline{U}_T^k \to \mathbb{R}$ is a local stabilization function. To state the assumptions on this function, we introduce the mesh skeleton $\partial\mathcal{M}_h \coloneqq \bigcup_{F \in \mathcal{F}_h} \overline{F}$ and set

$$\begin{aligned} L^p(\partial\mathcal{M}_h) &\coloneqq \left\{ \mu : \partial\mathcal{M}_h \to \mathbb{R} : \mu_{|F} \in L^p(F) \quad \forall F \in \mathcal{F}_h \right\}, \\ \|\mu\|_{L^p(\partial\mathcal{M}_h)} &\coloneqq \left(\sum_{T \in \mathcal{T}_h} h_T \sum_{F \in \mathcal{F}_T} \|\mu_{|F}\|_{L^p(F)}^p \right)^{\frac{1}{p}}. \end{aligned} \tag{21}$$

Assumption 2 (Local stabilization functions)**.** There exists $\zeta \in L^p(\partial\mathcal{M}_h; [0, +\infty))$ such that, for all $T \in \mathcal{T}_h$ and all $\underline{v}_T, \underline{w}_T \in \underline{U}_T^k$,

$$\mathrm{s}_T(\underline{w}_T, \underline{v}_T) \coloneqq h_T \int_{\partial T} \mathrm{S}_T(\cdot, \Delta_{\partial T}^k \underline{w}_T)\, \Delta_{\partial T}^k \underline{v}_T, \tag{22}$$

where $\mathrm{S}_T : \partial T \times \mathbb{R} \to \mathbb{R}$ is a measurable function satisfying, for all $v, w \in \mathbb{R}$ and almost every $\boldsymbol{x} \in \partial T$,

$$|\mathrm{S}_T(\boldsymbol{x}, w) - \mathrm{S}_T(\boldsymbol{x}, v)| \lesssim \sigma_{\mathrm{hc}} \left(\zeta(\boldsymbol{x})^p + |w|^p + |v|^p\right)^{\frac{p-2}{p}} |w - v|, \tag{23a}$$

$$\left(\mathrm{S}_T(\boldsymbol{x}, w) - \mathrm{S}_T(\boldsymbol{x}, v)\right)(w - v) \gtrsim \sigma_{\mathrm{sm}} \left(\zeta(\boldsymbol{x})^p + |w|^p + |v|^p\right)^{\frac{p-2}{p}} |w - v|^2, \tag{23b}$$

$$\mathrm{S}_T(\boldsymbol{x}, 0) = 0, \tag{23c}$$

while the boundary residual operator $\Delta_{\partial T}^k : \underline{U}_T^k \to L^p(\partial T)$ is such that, for all $\underline{v}_T \in \underline{U}_T^k$,

$$(\Delta_{\partial T}^k \underline{v}_T)_{|F} \coloneqq \frac{1}{h_T} \left[\pi_F^k(\mathrm{r}_T^{k+1} \underline{v}_T - v_F) - \pi_T^k(\mathrm{r}_T^{k+1} \underline{v}_T - v_T) \right] \qquad \forall F \in \mathcal{F}_T \tag{24}$$

with potential reconstruction r_T^{k+1} defined by (19).

Example 8 (Stabilization function)**.** Local stabilization functions that match Assumption 2 can be obtained setting, for all $T \in \mathcal{T}_h$, all $w \in \mathbb{R}$, and all $\boldsymbol{x} \in \partial T$,

$$\mathrm{S}_T(\boldsymbol{x}, w) \coloneqq \gamma_T \left(\zeta(\boldsymbol{x})^p + |w|^p\right)^{\frac{p-2}{p}} w, \tag{25}$$

with $\gamma_T \in [\sigma_{\mathrm{sm}}, \sigma_{\mathrm{hc}}]$ (see (3)). It can be checked that S_T is a non-degenerate p-power-framed function satisfying (23); see [6, Appendix A] for a proof.

Leveraging the results in [8], and additionally using $h_F \simeq h_T$ for all $F \in \mathcal{F}_T$, it can be checked that, for all $q \in (1, +\infty)$,

$$\|\mathbf{G}_T^k \underline{v}_T\|_{L^q(T)^d}^q + h_T \|\Delta_{\partial T}^k \underline{v}_T\|_{L^q(\partial T)}^q \simeq \|\underline{v}_T\|_{1,q,T}^q \qquad \forall \underline{v}_T \in \underline{U}_T^k. \tag{26}$$

Additionally, $\Delta_{\partial T}^k$ is polynomially consistent, i.e.,

$$\Delta_{\partial T}^k(\underline{I}_T^k w) = 0 \qquad \forall w \in \mathbb{P}^{k+1}(T)^d. \tag{27}$$

4.4 Discrete problem

The discrete problem reads: Find $\underline{u}_h \in \underline{U}_{h,0}^k$ such that

$$\mathrm{a}_h(\underline{u}_h, \underline{v}_h) = \int_\Omega f v_h \qquad \forall \underline{v}_h \in \underline{U}_{h,0}^k. \tag{28}$$

5 Error analysis

In this section, after establishing a stability result for the discrete function a_h, we prove the error estimate that constitutes the main result of this paper.

5.1 Strong monotonicity of the discrete diffusion function

We recall the following inequality between sums of power (see [6, Eq. (15)]): Let an integer $n \geq 1$ and a real number $m \in (0, +\infty)$ be given. Then, for all $a_1, \ldots, a_n \in (0, +\infty)$, we have

$$n^{-(m-1)^\ominus} \sum_{i=1}^n a_i^m \leq \left(\sum_{i=1}^n a_i\right)^m \leq n^{(m-1)^\oplus} \sum_{i=1}^n a_i^m. \tag{29}$$

Lemma 9 (Strong monotonicity of a_h)**.** *For all $\underline{v}_h, \underline{w}_h \in \underline{U}_h^k$, setting $\underline{e}_h := \underline{v}_h - \underline{w}_h$, it holds*

$$\begin{aligned}\|\underline{e}_h\|_{1,p,h}^2 \lesssim \sigma_{\mathrm{sm}}^{-1} &\left(\|\delta\|_{L^p(\Omega)}^p + \|\zeta\|_{L^p(\partial\mathcal{M}_h)}^p + \|\underline{v}_h\|_{1,p,h}^p + \|\underline{w}_h\|_{1,p,h}^p\right)^{\frac{2-p}{p}} \\ &\times \left(\mathrm{a}_h(\underline{v}_h, \underline{e}_h) - \mathrm{a}_h(\underline{w}_h, \underline{e}_h)\right).\end{aligned} \tag{30}$$

Proof. Let $T \in \mathcal{T}_h$. Using the strong monotonicity (2d) of $\boldsymbol{\sigma}$ and the $\left(\frac{2}{2-p}, \frac{2}{p}\right)$-Hölder inequality, we get

$$\begin{aligned}
&\sigma_{\mathrm{sm}}^{\frac{p}{2}} \|\mathbf{G}_T^k \underline{e}_T\|_{L^p(T)^d}^p \\
&\leq \int_T \left(\delta^p + |\mathbf{G}_T^k \underline{v}_T|^p + |\mathbf{G}_T^k \underline{w}_T|^p\right)^{\frac{2-p}{2}} \left[\left(\boldsymbol{\sigma}(\cdot, \mathbf{G}_T^k \underline{v}_T) - \boldsymbol{\sigma}(\cdot, \mathbf{G}_T^k \underline{w}_T)\right) \cdot \mathbf{G}_T^k \underline{e}_T\right]^{\frac{p}{2}} \\
&\leq \left(\|\delta\|_{L^p(T)}^p + \|\mathbf{G}_T^k \underline{v}_T\|_{L^p(T)^d}^p + \|\mathbf{G}_T^k \underline{w}_T\|_{L^p(T)^d}^p\right)^{\frac{2-p}{2}} \\
&\quad \times \left[\int_T \left(\boldsymbol{\sigma}(\cdot, \mathbf{G}_T^k \underline{v}_T) - \boldsymbol{\sigma}(\cdot, \mathbf{G}_T^k \underline{w}_T)\right) \cdot \mathbf{G}_T^k \underline{e}_T\right]^{\frac{p}{2}} \\
&\lesssim \left(\|\delta\|_{L^p(T)}^p + \|\underline{v}_T\|_{1,p,T}^p + \|\underline{w}_T\|_{1,p,T}^p\right)^{\frac{2-p}{2}} \\
&\quad \times \left[\int_T \left(\boldsymbol{\sigma}(\cdot, \mathbf{G}_T^k \underline{v}_T) - \boldsymbol{\sigma}(\cdot, \mathbf{G}_T^k \underline{w}_T)\right) \cdot \mathbf{G}_T^k \underline{e}_T\right]^{\frac{p}{2}},
\end{aligned} \tag{31}$$

where the conclusion follows from the seminorm equivalence (26). Similarly, the strong monotonicity (23b) of S_T followed by the same reasoning as above yields,

$$\begin{aligned}\sigma_{\mathrm{sm}}^{\frac{p}{2}} h_T \|\Delta_{\partial T}^k \underline{e}_T\|_{L^p(\mathcal{F}_T)}^p \lesssim &\left(h_T \|\zeta\|_{L^p(\partial T)}^p + \|\underline{v}_T\|_{1,p,T}^p + \|\underline{w}_T\|_{1,p,T}^p\right)^{\frac{2-p}{2}} \\ &\times \left(\mathrm{s}_T(\underline{v}_T, \underline{e}_T) - \mathrm{s}_T(\underline{w}_T, \underline{e}_T)\right)^{\frac{p}{2}}.\end{aligned} \tag{32}$$

Combining the norm equivalence (26) with (31) and (32) and using (29) yields

$$\begin{aligned}\sigma_{\mathrm{sm}}^{\frac{p}{2}} \|\underline{e}_T\|_{1,p,T}^p \lesssim &\left(\|\delta\|_{L^p(T)}^p + h_T \|\zeta\|_{L^p(\partial T)}^p + \|\underline{v}_T\|_{1,p,T}^p + \|\underline{w}_T\|_{1,p,T}^p\right)^{\frac{2-p}{2}} \\ &\times \left(\mathrm{a}_T(\underline{v}_T, \underline{e}_T) - \mathrm{a}_T(\underline{w}_T, \underline{e}_T)\right)^{\frac{p}{2}}.\end{aligned}$$

Summing over $T \in \mathcal{T}_h$, applying the discrete $\left(\frac{2}{2-p}, \frac{2}{p}\right)$-Hölder inequality, and raising to the power $\frac{2}{p}$ yields (30). □

Remark 10 (Well-posedness and a priori estimate). Using standard techniques (cf. [7, Theorem 4.5], see also [6, Theorem 17]), it can be proved that there exists a unique solution $\underline{u}_h \in \underline{U}_{h,0}^k$ to the discrete problem (28). Additionally, it can be shown in a similar way as for the continuous case (cf. Proposition 7) that the following a priori bound holds:

$$\|\underline{u}_h\|_{1,p,h} \lesssim \left(\sigma_{\mathrm{sm}}^{-1}\|f\|_{L^{p'}(\Omega)}\right)^{\frac{1}{p-1}} + \min\left(\left(\|\delta\|_{L^p(\mathcal{T}_h)}^p + \|\zeta\|_{L^p(\partial\mathcal{M}_h)}^p\right)^{\frac{1}{p}}; \sigma_{\mathrm{sm}}^{-1}\left(\|\delta\|_{L^p(\mathcal{T}_h)}^p + \|\zeta\|_{L^p(\partial\mathcal{M}_h)}^p\right)^{\frac{2-p}{p}}\|f\|_{L^{p'}(\Omega)}\right). \quad (33)$$

5.2 Error estimate

Theorem 11 (Error estimate)**.** *Let $u \in U$ and $\underline{u}_h \in \underline{U}_{h,0}^k$ solve (6) and (28), respectively. Assume $u \in W^{k+2,p}(\mathcal{T}_h)$ and $\boldsymbol{\sigma}(\cdot, \boldsymbol{\nabla} u) \in W^{1,p'}(\Omega)^d \cap W^{k+1,p'}(\mathcal{T}_h)^d$. Then, under Assumptions 1 and 2,*

$$\|\underline{u}_h - \underline{I}_h^k u\|_{1,p,h} \lesssim \mathcal{N}_f\, h^{k+1}|\boldsymbol{\sigma}(\cdot, \boldsymbol{\nabla} u)|_{W^{k+1,p'}(\mathcal{T}_h)^d} + \mathcal{N}_f\, \sigma_{\mathrm{hc}}\left[\sum_{T\in\mathcal{T}_h}\left(\min\left(\eta_T; 1\right)^{2-p} h_T^{(k+1)(p-1)}|u|_{W^{k+2,p}(T)}^{p-1}\right)^{p'}\right]^{\frac{1}{p'}}, \quad (34)$$

where, for all $T \in \mathcal{T}_h$, defining $\mathfrak{D}_T := \min\left(\operatorname{ess\,inf}_{\boldsymbol{x}\in T}\left(\delta(\boldsymbol{x}) + |\boldsymbol{\nabla} u(\boldsymbol{x})|\right); \operatorname{ess\,inf}_{\boldsymbol{x}\in\partial T}\zeta(\boldsymbol{x})\right)$, we have set

$$\eta_T := \frac{h_T^{k+1}|u|_{W^{k+2,p}(T)}}{|T|^{\frac{1}{p}}\mathfrak{D}_T} \quad (35)$$

with the convention that $\eta_T = +\infty$ if $\mathfrak{D}_T = 0 < |u|_{W^{k+2,p}(\mathcal{T}_h)}$, and $\eta_T = 0$ if $\mathfrak{D}_T = |u|_{W^{k+2,p}(\mathcal{T}_h)} = 0$, and where

$$\mathcal{N}_f := \sigma_{\mathrm{sm}}^{-1}\left[\|\delta\|_{L^p(\Omega)} + \|\zeta\|_{L^p(\partial\mathcal{M}_h)} + \left(\sigma_{\mathrm{sm}}^{-1}\|f\|_{L^{p'}(\Omega)}\right)^{\frac{1}{p-1}}\right]^{2-p}.$$

Remark 12 (Convergence rates). For any $T \in \mathcal{T}_h$, the local *flux degeneracy parameter* $\mathfrak{D}_T$ which appears in (35) is a measure of the local degeneracy of the flux and the stabilization function: the closer it is to zero, the more degenerate the model is. The dimensionless number η_T defined in (35) determines the convergence rate of the contribution to the approximation error stemming from T. If $\eta_T \geq 1$ (*locally degenerate case*), then the element T contributes to the error with a term in $h_T^{(k+1)(p-1)}$. If $\eta_T \leq h_T^{k+1}|u|_{W^{k+2,p}(T)}|T|^{-\frac{1}{p}}$, i.e. $\mathfrak{D}_T \geq 1$ (*locally non-degenerate case*), the contribution to the error is in h_T^{k+1}. The case $\eta_T \in (h_T^{k+1}|u|_{W^{k+2,p}(T)}|T|^{-\frac{1}{p}}, 1)$ corresponds to intermediate rates of convergence.

At the global level, defining the number $\eta_h := \max_{T\in\mathcal{T}_h}\eta_T$, the bound $h_T \leq h$ together with the error estimate (34) yields

$$\|\underline{u}_h - \underline{I}_h^k u\|_{1,p,h} \lesssim \mathcal{N}_f\left(h^{k+1}|\boldsymbol{\sigma}(\cdot, \boldsymbol{\nabla} u)|_{W^{k+1,p'}(\mathcal{T}_h)^d} + \sigma_{\mathrm{hc}}\min\left(\eta_h; 1\right)^{2-p} h^{(k+1)(p-1)}|u|_{W^{k+2,p}(\mathcal{T}_h)}^{p-1}\right). \quad (36)$$

As a consequence, if $\eta_h \geq 1$ (*globally degenerate case*), then the convergence rate is $(k+1)(p-1)$. If $\eta_h \leq h^{k+1}|u|_{W^{k+2,p}(\mathcal{T}_h)}$ (*globally non-degenerate case*), the convergence rate is $k+1$. Finally, the case

$\eta_h \in (h^{k+1}|u|_{W^{k+2,p}(\mathcal{T}_h)}, 1)$ corresponds to intermediate rates of convergence. This is the finest global estimate that can be obtained from the local one. However, for practical purposes, if $u \in W^{k+2,\infty}(\mathcal{T}_h)$ then we can replace η_h in (36) by the larger number

$$\tilde{\eta}_h \coloneqq \frac{|u|_{W^{k+2,\infty}(\mathcal{T}_h)} h^{k+1}}{\min_{T\in\mathcal{T}_h} \mathfrak{D}_T} = \frac{|u|_{W^{k+2,\infty}(\mathcal{T}_h)} h^{k+1}}{\min\left(\operatorname{ess\,inf}_\Omega (\delta + |\boldsymbol{\nabla} u|)\,; \operatorname{ess\,inf}_{\partial\mathcal{M}_h} \zeta\right)}, \tag{37}$$

with the same convention as above regarding fractions $C/0$ and $0/0$. The convergence rate will result from the position of $\tilde{\eta}_h$ with respect to $h^{k+1}|u|_{W^{k+2,\infty}(\mathcal{T}_h)}$ and 1, and $\tilde{\eta}_h \le h^{k+1}|u|_{W^{k+2,\infty}(\mathcal{T}_h)}$ (non-degenerate case) is equivalent to $\min\left(\operatorname{ess\,inf}_\Omega (\delta + |\boldsymbol{\nabla} u|)\,; \operatorname{ess\,inf}_{\partial\mathcal{M}_h} \zeta\right) \ge 1$, which is consistent with the local requirement stated above.

Proof of Theorem 11. Define the consistency error as the linear form $\mathcal{E}_h : \underline{U}_h^k \to \mathbb{R}$ such that, for all $\underline{v}_h \in \underline{U}_h^k$,

$$\mathcal{E}_h(\underline{v}_h) \coloneqq \int_\Omega \boldsymbol{\nabla}\cdot\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla} u)\; v_h + \mathrm{a}_h(\underline{I}_h^k u, \underline{v}_h). \tag{38}$$

Let, for the sake of brevity, $\underline{\hat{u}}_h \coloneqq \underline{I}_h^k u$ and $\underline{e}_h \coloneqq \underline{u}_h - \underline{\hat{u}}_h \in \underline{U}_{h,0}^k$.

(i) *Estimate of the consistency error.* Expanding a_h according to its definition (20) in the expression (38) of $\mathcal{E}_h$, inserting $\sum_{T\in\mathcal{T}_h}\left(\int_T \boldsymbol{\pi}_T^k\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla} u)\cdot\mathbf{G}_T^k\underline{e}_T - \int_T \boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla} u)\cdot\mathbf{G}_T^k\underline{e}_T\right) = 0$ (the equality is a consequence of the definition of $\boldsymbol{\pi}_T^k$), and rearranging, we obtain

$$\begin{aligned}\mathcal{E}_h(\underline{e}_h) = \underbrace{\int_\Omega \boldsymbol{\nabla}\cdot\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla} u)\; e_h + \sum_{T\in\mathcal{T}_h}\int_T \boldsymbol{\pi}_T^k\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla} u)\cdot\mathbf{G}_T^k\underline{e}_T}_{\mathcal{T}_1}\\ + \underbrace{\sum_{T\in\mathcal{T}_h}\int_T \left[\boldsymbol{\sigma}(\cdot,\mathbf{G}_T^k\underline{\hat{u}}_T) - \boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla} u)\right]\cdot\mathbf{G}_T^k\underline{e}_T}_{\mathcal{T}_2} + \underbrace{\sum_{T\in\mathcal{T}_h}\mathrm{s}_T(\underline{\hat{u}}_T,\underline{e}_T)}_{\mathcal{T}_3}.\end{aligned} \tag{39}$$

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

For the first term, we start by noticing that

$$\sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T}\int_F e_F\,(\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla} u)\cdot\boldsymbol{n}_{TF}) = 0 \tag{40}$$

as a consequence of the continuity of the normal trace of $\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla} u)$ together with the single-valuedness of e_F across each interface $F \in \mathcal{F}_h^{\mathrm{i}}$ and the fact that $e_F = 0$ for every boundary face $F \in \mathcal{F}_h^{\mathrm{b}}$ (see [9, Corollary 1.19]). Using an element by element integration by parts on the first term of $\mathcal{T}_1$ along with the definition (17) of $\mathbf{G}_T^k$, we can write

$$\begin{aligned}\mathcal{T}_1 &= \cancel{\sum_{T\in\mathcal{T}_h}\int_T \left[\boldsymbol{\pi}_T^k\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla} u) - \boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla} u)\right]\cdot\boldsymbol{\nabla} e_T}\\ &\quad + \sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T}\left[\int_F (e_F - e_T)\left(\boldsymbol{\pi}_T^k\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla} u)\cdot\boldsymbol{n}_{TF}\right) + \int_F e_T\,(\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla} u)\cdot\boldsymbol{n}_{TF})\right]\\ &= \sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T}\int_F (e_F - e_T)\left[\boldsymbol{\pi}_T^k\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla} u) - \boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla} u)\right]\cdot\boldsymbol{n}_{TF},\end{aligned}$$

where we have used the definition of $\boldsymbol{\pi}_T^k$ together with the fact that $\boldsymbol{\nabla} e_T \in \mathbb{P}^{k-1}(T)^d \subset \mathbb{P}^k(T)^d$ to cancel the term in the first line, and we have inserted (40) and rearranged to conclude. Hölder inequalities give

$$\begin{aligned}
|\mathcal{T}_1| &\lesssim \left(\sum_{T\in\mathcal{T}_h} h_T \|\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla} u) - \boldsymbol{\pi}_T^k \boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla} u)\|_{L^{p'}(\partial T)^d}^{p'}\right)^{\frac{1}{p'}} \\
&\quad \times \left(\sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T} h_F^{1-p} \|e_F - e_T\|_{L^p(F)}^p\right)^{\frac{1}{p}} \\
&\lesssim h^{k+1} |\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla} u)|_{W^{k+1,p'}(\mathcal{T}_h)^d} \|\underline{e}_h\|_{1,p,h},
\end{aligned} \tag{41}$$

where we have used the $(k+1,p')$-trace approximation properties (13b) of $\boldsymbol{\pi}_T^k$ along with $h_T \le h$ for the first factor, and the definition (14) of $\|\cdot\|_{1,p,h}$ for the second.

We move on to the next term $\mathcal{T}_2$. Let an element $T \in \mathcal{T}_h$ be fixed. If $\eta_T \ge 1$, using the (p',p)-Hölder inequality together with the equivalence (26), we obtain

$$\begin{aligned}
&\left|\int_T \left[\boldsymbol{\sigma}(\cdot,\mathbf{G}_T^k \underline{\hat{u}}_T) - \boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla} u)\right]\cdot \mathbf{G}_T^k \underline{e}_T\right| \\
&\quad \le \|\boldsymbol{\sigma}(\cdot,\mathbf{G}_T^k \underline{\hat{u}}_T) - \boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla} u)\|_{L^{p'}(T)^d} \|\underline{e}_T\|_{1,p,T} \\
&\quad \le \sigma_{\mathrm{hc}} \left\|\left(\delta + |\boldsymbol{\pi}_T^k(\boldsymbol{\nabla} u)| + |\boldsymbol{\nabla} u|\right)^{p-2} |\boldsymbol{\pi}_T^k(\boldsymbol{\nabla} u) - \boldsymbol{\nabla} u|\right\|_{L^{p'}(T)} \|\underline{e}_T\|_{1,p,T} \\
&\quad \le \sigma_{\mathrm{hc}} \|\boldsymbol{\pi}_T^k(\boldsymbol{\nabla} u) - \boldsymbol{\nabla} u\|_{L^p(T)^d}^{p-1} \|\underline{e}_T\|_{1,p,T} \\
&\quad \lesssim \sigma_{\mathrm{hc}} h_T^{(k+1)(p-1)} |u|_{W^{k+2,p}(T)}^{p-1} \|\underline{e}_T\|_{1,p,T} \\
&\quad = \sigma_{\mathrm{hc}} \min(\eta_T;1)^{2-p}\, h_T^{(k+1)(p-1)} |u|_{W^{k+2,p}(T)}^{p-1} \|\underline{e}_T\|_{1,p,T},
\end{aligned} \tag{42}$$

where we have used the continuity (2c) of $\boldsymbol{\sigma}$ together with the commutation property (18) of the discrete gradient and (29) in the second bound, inequality (4) with $x = \boldsymbol{\pi}_T^k(\boldsymbol{\nabla} u)$, $y = \boldsymbol{\nabla} u$, and $\alpha = \delta$ in the third bound, and the $(k+1,p,0)$-approximation properties of π_T^k to conclude.

On the other hand, if $\eta_T < 1$ then, using the (p,p')-Hölder inequality together with the boundedness (26), we infer

$$\begin{aligned}
&\left|\int_T \left[\boldsymbol{\sigma}(\cdot,\mathbf{G}_T^k \underline{\hat{u}}_T) - \boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla} u)\right]\cdot \mathbf{G}_T^k \underline{e}_T\right| \le \|\boldsymbol{\sigma}(\cdot,\mathbf{G}_T^k \underline{\hat{u}}_T) - \boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla} u)\|_{L^p(T)^d} \|\underline{e}_T\|_{1,p',T} \\
&\quad \lesssim \sigma_{\mathrm{hc}} \left\|\left(\delta + |\boldsymbol{\pi}_T^k(\boldsymbol{\nabla} u)| + |\boldsymbol{\nabla} u|\right)^{p-2} |\boldsymbol{\pi}_T^k(\boldsymbol{\nabla} u) - \boldsymbol{\nabla} u|\right\|_{L^p(T)} |T|^{\frac{p-2}{p}} \|\underline{e}_T\|_{1,p,T} \\
&\quad \le \sigma_{\mathrm{hc}} \mathfrak{D}_T^{p-2} |T|^{\frac{p-2}{p}} \|\boldsymbol{\pi}_T^k(\boldsymbol{\nabla} u) - \boldsymbol{\nabla} u\|_{L^p(T)^d} \|\underline{e}_T\|_{1,p,T} \\
&\quad \lesssim \sigma_{\mathrm{hc}} \mathfrak{D}_T^{p-2} |T|^{\frac{p-2}{p}} h_T^{k+1} |u|_{W^{k+2,p}(T)} \|\underline{e}_T\|_{1,p,T} \\
&\quad = \sigma_{\mathrm{hc}} \min(\eta_T;1)^{2-p}\, h_T^{(k+1)(p-1)} |u|_{W^{k+2,p}(T)}^{p-1} \|\underline{e}_T\|_{1,p,T},
\end{aligned} \tag{43}$$

where we passed to the second line as in (42) additionally using the bound $\|\underline{e}_T\|_{1,p',T} \lesssim |T|^{\frac{p-2}{p}} \|\underline{e}_T\|_{1,p,T}$ (see [7, Lemmas 5.1 and 5.2]), used in the third line the fact that $\mathbb{R} \ni x \mapsto x^{p-2} \in \mathbb{R}$ is non-increasing to infer that $\left(\delta + |\boldsymbol{\pi}_T^k(\boldsymbol{\nabla} u)| + |\boldsymbol{\nabla} u|\right)^{p-2} \le (\delta + |\boldsymbol{\nabla} u|)^{p-2} \le \mathfrak{D}_T^{p-2}$ almost everywhere in T, and concluded as above.

Gathering the estimates (42) and (43) and using a discrete Hölder inequality yields

$$|\mathcal{T}_2| \lesssim \sigma_{\mathrm{hc}} \left[\sum_{T\in\mathcal{T}_h} \left(\min(\eta_T;1)^{2-p}\, h_T^{(k+1)(p-1)} |u|_{W^{k+2,p}(T)}^{p-1}\right)^{p'}\right]^{\frac{1}{p'}} \|\underline{e}_h\|_{1,p,h}. \tag{44}$$

Let us finally consider $\mathcal{T}_3$. Let $T \in \mathcal{T}_h$, set $\underline{\check{u}}_T := \underline{I}_T^k(\pi_T^{k+1}u)$ for the sake of brevity, and observe that $\mathrm{S}_T\big(\cdot, \Delta_{\partial T}^k \underline{\check{u}}_T\big) = 0$ thanks to the polynomial consistency (27) of $\Delta_{\partial T}^k$ and the property (23c) of S_T.

If $\eta_T \geq 1$, using the (p', p)-Hölder inequality together with the boundedness property (26) (with $q = p$), we infer

$$
\begin{aligned}
|\mathrm{s}_T(\underline{\hat{u}}_T, \underline{e}_T)| &\lesssim h_T^{\frac{1}{p'}} \big\|\mathrm{S}_T\big(\cdot, \Delta_{\partial T}^k \underline{\hat{u}}_T\big) - \mathrm{S}_T\big(\cdot, \Delta_{\partial T}^k \underline{\check{u}}_T\big)\big\|_{L^{p'}(\partial T)} \|\underline{e}_T\|_{1,p,T} \\
&\lesssim \sigma_{\mathrm{hc}} h_T^{\frac{1}{p'}} \left\|\Big(\zeta + |\Delta_{\partial T}^k \underline{\hat{u}}_T| + |\Delta_{\partial T}^k \underline{\check{u}}_T|\Big)^{p-2} \Delta_{\partial T}^k(\underline{\hat{u}}_T - \underline{\check{u}}_T)\right\|_{L^{p'}(\partial T)} \|\underline{e}_T\|_{1,p,T} \\
&\leq \sigma_{\mathrm{hc}} h_T^{\frac{1}{p'}} \|\Delta_{\partial T}^k(\underline{\hat{u}}_T - \underline{\check{u}}_T)\|_{L^p(\partial T)}^{p-1} \|\underline{e}_T\|_{1,p,T} \\
&\lesssim \sigma_{\mathrm{hc}} h_T^{(k+1)(p-1)} |u|_{W^{k+2,p}(T)}^{p-1} \|\underline{e}_T\|_{1,p,T} \\
&= \sigma_{\mathrm{hc}} \min(\eta_T; 1)^{2-p}\, h_T^{(k+1)(p-1)} |u|_{W^{k+2,p}(T)}^{p-1} \|\underline{e}_T\|_{1,p,T},
\end{aligned}
\tag{45}
$$

where we have used the continuity (23a) of S_T together with (29) to pass to the second line, inequality (4) with $x = \Delta_{\partial T}^k \underline{\hat{u}}_T$, $y = \Delta_{\partial T}^k \underline{\check{u}}_T$ and $\alpha = \zeta$ in the third line, and the boundedness (26) of $\Delta_{\partial T}^k$ and (15) of $\underline{I}_T^k$ together with the $(k+1, p, 1)$-approximation properties of π_T^{k+1} to conclude by writing $h_T^{\frac{1}{p}} \|\Delta_{\partial T}^k(\underline{\hat{u}}_T - \underline{\check{u}}_T)\|_{L^p(\partial T)} \lesssim \|\underline{I}_T^k(u - \pi_T^{k+1}u)\|_{1,p,T} \lesssim |u - \pi_T^{k+1}u|_{W^{1,p}(T)} \lesssim h_T^{k+1}|u|_{W^{k+2,p}(T)}$.

Otherwise, $\eta_T < 1$ and using the (p, p')-Hölder inequality together with boundedness property (26) (with $q = p'$), we infer as above that

$$
\begin{aligned}
&|\mathrm{s}_T(\underline{\hat{u}}_T, \underline{e}_T)| \\
&\quad\lesssim h_T^{\frac{1}{p}} \big\|\mathrm{S}_T\big(\cdot, \Delta_{\partial T}^k \underline{\hat{u}}_T\big) - \mathrm{S}_T\big(\cdot, \Delta_{\partial T}^k \underline{\check{u}}_T\big)\big\|_{L^p(\partial T)} \|\underline{e}_T\|_{1,p',T} \\
&\quad\lesssim \sigma_{\mathrm{hc}} h_T^{\frac{1}{p}} \left\|\Big(\zeta + |\Delta_{\partial T}^k \underline{\hat{u}}_T| + |\Delta_{\partial T}^k \underline{\check{u}}_T|\Big)^{p-2} \Delta_{\partial T}^k(\underline{\hat{u}}_T - \underline{\check{u}}_T)\right\|_{L^p(\partial T)} |T|^{\frac{p-2}{p}} \|\underline{e}_T\|_{1,p,T} \\
&\quad\leq \sigma_{\mathrm{hc}} \mathfrak{D}_T^{p-2} |T|^{\frac{p-2}{p}} h_T^{\frac{1}{p}} \|\Delta_{\partial T}^k(\underline{\hat{u}}_T - \underline{\check{u}}_T)\|_{L^p(\partial T)} \|\underline{e}_T\|_{1,p,T} \\
&\quad\lesssim \sigma_{\mathrm{hc}} \mathfrak{D}_T^{p-2} |T|^{\frac{p-2}{p}} h_T^{k+1} |u|_{W^{k+2,p}(T)} \|\underline{e}_T\|_{1,p,T} \\
&\quad= \sigma_{\mathrm{hc}} \min(\eta_T; 1)^{2-p}\, h_T^{(k+1)(p-1)} |u|_{W^{k+2,p}(T)}^{p-1} \|\underline{e}_T\|_{1,p,T},
\end{aligned}
\tag{46}
$$

where the second inequality follows as before from the bound $\|\underline{e}_T\|_{1,p',T} \lesssim |T|^{\frac{p-2}{p}} \|\underline{e}_T\|_{1,p,T}$ (see [7, Lemmas 5.1 and 5.2]), the third inequality is a consequence of the monotonicity of $\mathbb{R} \ni x \mapsto x^{p-2} \in \mathbb{R}$ that yields $\big(\zeta + |\Delta_{\partial T}^k \underline{\hat{u}}_T| + |\Delta_{\partial T}^k \underline{\check{u}}_T|\big)^{p-2} \leq \zeta^{p-2} \leq \mathfrak{D}_T^{p-2}$ almost everywhere in ∂T, and the conclusion is obtained as in (45).

Following then the same reasoning that lead to (44), we obtain for the third term

$$
|\mathcal{T}_3| \lesssim \sigma_{\mathrm{hc}} \left[\sum_{T \in \mathcal{T}_h} \Big(\min(\eta_T; 1)^{2-p}\, h_T^{(k+1)(p-1)} |u|_{W^{k+2,p}(T)}^{p-1}\Big)^{p'}\right]^{\frac{1}{p'}} \|\underline{e}_h\|_{1,p,h}. \tag{47}
$$

Plugging the bounds (41), (44), and (47) into (39) yields

$$
\begin{aligned}
|\mathcal{E}_h(\underline{e}_h)| &\lesssim h^{k+1} |\boldsymbol{\sigma}(\cdot, \boldsymbol{\nabla} u)|_{W^{k+1,p'}(\mathcal{T}_h)^d} \|\underline{e}_h\|_{1,p,h} \\
&\quad + \sigma_{\mathrm{hc}} \left[\sum_{T \in \mathcal{T}_h} \Big(\min(\eta_T; 1)^{2-p}\, h_T^{(k+1)(p-1)} |u|_{W^{k+2,p}(T)}^{p-1}\Big)^{p'}\right]^{\frac{1}{p'}} \|\underline{e}_h\|_{1,p,h}.
\end{aligned}
\tag{48}
$$

(ii) *Error estimate.* Using the strong monotonicity (30) of a_h, we get

$$\|\underline{e}_h\|_{1,p,h}^2 \lesssim \sigma_{\mathrm{sm}}^{-1} \left(\|\delta\|_{L^p(\Omega)}^p + \|\zeta\|_{L^p(\partial\mathcal{M}_h)}^p + \|\underline{u}_h\|_{1,p,h}^p + \|\hat{\underline{u}}_h\|_{1,p,h}^p \right)^{\frac{2-p}{p}} \tag{49}$$

$$\times \left[\mathrm{a}_h(\underline{u}_h, \underline{e}_h) - \mathrm{a}_h(\hat{\underline{u}}_h, \underline{e}_h) \right]$$

$$\lesssim \mathcal{N}_f \left[\mathrm{a}_h(\underline{u}_h, \underline{e}_h) - \mathrm{a}_h(\hat{\underline{u}}_h, \underline{e}_h) \right], \tag{50}$$

where we have used the a priori bound (33) on the discrete solution along with the boundedness (15) of the global interpolator, the a priori bound (8) on the continuous solution, and (29) to conclude. Furthermore, using the equation (28) (with $\underline{v}_h = \underline{e}_h$), and the fact that $f = -\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}(\cdot, \boldsymbol{\nabla} u)$ almost everywhere in Ω, we see that

$$\mathrm{a}_h(\underline{u}_h, \underline{e}_h) - \mathrm{a}_h(\hat{\underline{u}}_h, \underline{e}_h) = \int_\Omega f e_h - \mathrm{a}_h(\hat{\underline{u}}_h, \underline{e}_h) = -\mathcal{E}_h(\underline{e}_h). \tag{51}$$

Hence, plugging (51) into (50), recalling the bound (48) on the consistency error, and simplifying, (34) follows. □

6 Numerical examples

In this section, we give some numerical results to confirm Theorem 11. We consider the domain $\Omega = (0,1)^2$ and define $\boldsymbol{\sigma}$ as the Carreau–Yasuda law of Example 6 with $p \in \{1.25, 1.5, 1.75\}$ and $\mu = a = 1$. The degeneracy parameter δ and the exact solution u will depend on the considered case. The stabilization functions are defined by (25) with ζ such that the local flux degeneracy number $\mathfrak{D}_T$ introduced in (35) is equal to the first argument of its min for all $T \in \mathcal{T}_h$, so that ζ does not influence the error estimates. The function f and the Dirichlet boundary condition are inferred from the exact solution. In all cases, except for the non-homogeneous boundary condition, these solutions match the assumptions required in Theorem 11. We consider the HHO scheme for $k \in \{1, 2, 3\}$ on a triangular mesh family.

6.1 Non-degenerate flux

We consider nonzero constant degeneracy parameters $\delta \in \{1, 0.1, 10^{-2}, 5 \cdot 10^{-4}\}$, and the potential u is given by

$$u(x, y) = \sin(\pi x) \sin(\pi y) \quad \forall (x, y) \in \Omega.$$

Thus, the dimensionless number $\tilde{\eta}_h$ defined in (37) satisfies

$$\tilde{\eta}_h = \mu_h(\delta) \coloneqq \frac{2^{\frac{k-1}{2}} \pi^k h^{k+1}}{\delta}. \tag{52}$$

Therefore, we should observe a $(k+1)(p-1)$ pre-asymptotic order of convergence until the size of the mesh is small enough compared to δ so that the convergence rate switches to $k+1$ (see Remark 12).

Indeed, in Figure 1 and Table 1, we can see for $k \in \{1, 2\}$ in the first row of results (corresponding to the case $\delta = 1$) a constant convergence rate of $k+1$, which is in agreement with Remark 12 since $\tilde{\eta}_h \leq h^{k+1} |u|_{W^{k+2,\infty}(\mathcal{T}_h)} \Leftrightarrow \delta \geq 1$. From row to row, we can observe a lower pre-asymptotic convergence rate (still above $(k+1)(p-1)$, but close to it in certain cases) which becomes worse as δ decreases; this is expected since $\tilde{\eta}_h$ is proportional to $1/\delta$. We note that, as k increases, the asymptotic rates are lower than the expected h^{k+1}, which could be due to the asymptotic regime not being achieved yet for these high-order schemes (due to the constants involving higher derivatives of u in (12)). The saturation of convergence rate, for $k = 2, 3$, at a lower rate than $(k+1)(p-1)$ when δ is very small could also be explained by the fact that the $W^{k+1,p'}$ norm of $\boldsymbol{\sigma}(\cdot, \boldsymbol{\nabla} u)$ explodes as $\delta \to 0$.

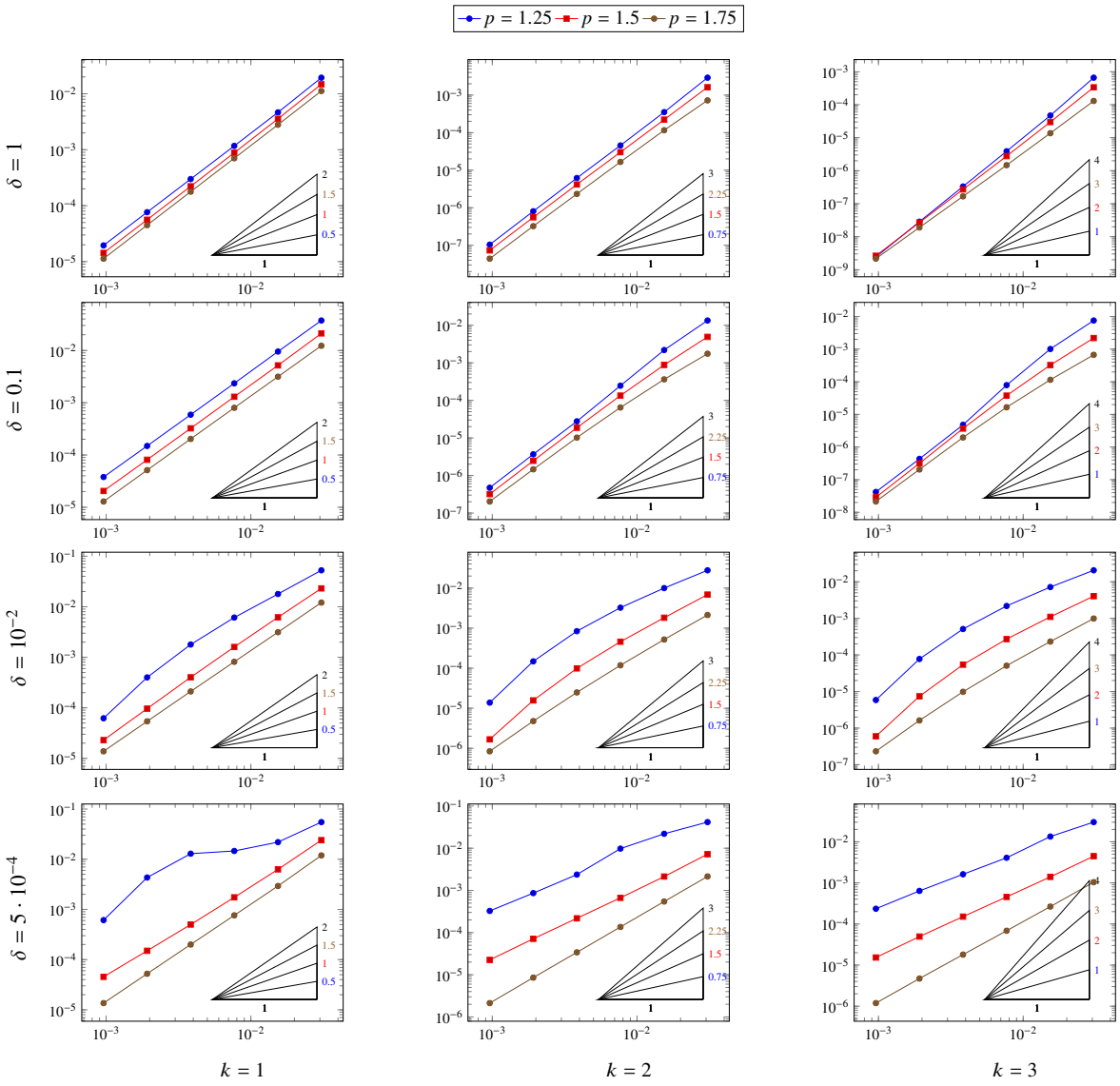

Figure 1: Numerical results for the test case of Subsection 6.1. The steeper slope (in black) indicates the $k+1$ convergence rate expected from Theorem 11 when the number δ is large enough compared to the mesh size. Otherwise the other slopes indicate the $(k+1)(p-1)$ convergence rate according to p.

Table 1: Convergence rates for the test case of Subsection 6.1. The bold numbers in each column correspond to the $(k+1)(p-1) \sim (k+1)$ convergence rates.

$\delta = 1$									
k	1			2			3		
h \ p	1.25	1.5	1.75	1.25	1.5	1.75	1.25	1.5	1.75
3.07e-02	**0.5 ~ 2**	**1 ~ 2**	**1.5 ~ 2**	**0.75 ~ 3**	**1.5 ~ 3**	**2.25 ~ 3**	**1 ~ 4**	**2 ~ 4**	**3 ~ 4**
1.54e-02	2.07	2.06	2.01	3.06	2.88	2.65	3.83	3.51	3.26
7.68e-03	1.98	2.01	1.98	2.94	2.87	2.79	3.59	3.41	3.21
3.84e-03	1.97	1.99	1.98	2.88	2.86	2.83	3.55	3.35	3.14
1.92e-03	1.97	1.98	1.99	2.94	2.90	2.85	3.54	3.33	3.13
9.60e-04	1.97	1.97	1.99	2.95	2.93	2.88	3.56	3.34	3.14

$\delta = 0.1$									
k	1			2			3		
h \ p	1.25	1.5	1.75	1.25	1.5	1.75	1.25	1.5	1.75
3.07e-02	**0.5 ~ 2**	**1 ~ 2**	**1.5 ~ 2**	**0.75 ~ 3**	**1.5 ~ 3**	**2.25 ~ 3**	**1 ~ 4**	**2 ~ 4**	**3 ~ 4**
1.54e-02	1.98	2.04	1.97	2.63	2.49	2.28	2.92	2.77	2.55
7.68e-03	2.01	1.99	1.97	3.14	2.70	2.48	3.67	3.10	2.79
3.84e-03	1.99	2.00	1.98	3.15	2.84	2.66	4.04	3.35	3.07
1.92e-03	1.99	2.00	1.98	2.92	2.93	2.80	3.47	3.54	3.26
9.60e-04	1.98	1.98	1.99	2.96	2.95	2.85	3.37	3.46	3.26

$\delta = 10^{-2}$									
k	1			2			3		
h \ p	1.25	1.5	1.75	1.25	1.5	1.75	1.25	1.5	1.75
3.07e-02	**0.5 ~ 2**	**1 ~ 2**	**1.5 ~ 2**	**0.75 ~ 3**	**1.5 ~ 3**	**2.25 ~ 3**	**1 ~ 4**	**2 ~ 4**	**3 ~ 4**
1.54e-02	1.57	1.91	1.96	1.48	1.93	2.04	1.53	1.89	2.10
7.68e-03	1.54	1.94	1.94	1.62	1.98	2.13	1.71	2.01	2.18
3.84e-03	1.77	1.99	1.95	1.95	2.21	2.25	2.09	2.32	2.37
1.92e-03	2.16	2.06	1.97	2.50	2.65	2.38	2.72	2.87	2.60
9.60e-04	2.69	2.07	1.97	3.42	3.23	2.50	3.73	3.63	2.80

$\delta = 5 \cdot 10^{-4}$									
k	1			2			3		
h \ p	1.25	1.5	1.75	1.25	1.5	1.75	1.25	1.5	1.75
3.07e-02	**0.5 ~ 2**	**1 ~ 2**	**1.5 ~ 2**	**0.75 ~ 3**	**1.5 ~ 3**	**2.25 ~ 3**	**1 ~ 4**	**2 ~ 4**	**3 ~ 4**
1.54e-02	1.33	1.94	2.03	0.92	1.77	1.98	1.17	1.68	1.98
7.68e-03	0.58	1.84	1.94	1.17	1.67	2.00	1.71	1.61	1.94
3.84e-03	0.18	1.80	1.93	2.04	1.61	2.00	1.33	1.59	1.93
1.92e-03	1.58	1.74	1.94	1.46	1.61	1.99	1.34	1.61	1.93
9.60e-04	2.82	1.72	1.94	1.40	1.66	2.01	1.44	1.69	1.98

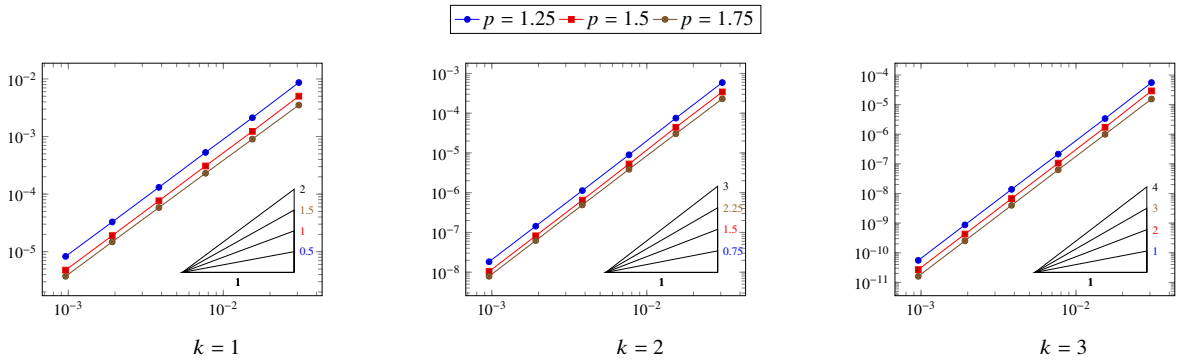

Figure 2: Numerical results for the test case of Subsection 6.2. The steeper slope (in black) indicates the $k+1$ convergence rate. The other slopes indicate the $(k+1)(p-1)$ convergence rate according to p.

Table 2: Convergence rates for the test case of Subsection 6.2. The bold numbers in each column correspond to the $(k+1)(p-1) \sim (k+1)$ convergence rates.

k	1			2			3		
h \ p	1.25	1.5	1.75	1.25	1.5	1.75	1.25	1.5	1.75
3.07e-02	**0.5 ~ 2**	**1 ~ 2**	**1.5 ~ 2**	**0.75 ~ 3**	**1.5 ~ 3**	**2.25 ~ 3**	**1 ~ 4**	**2 ~ 4**	**3 ~ 4**
1.54e-02	2.04	2.04	1.98	2.98	2.97	2.93	4.06	4.10	3.98
7.68e-03	2.00	1.99	1.96	3.05	3.06	2.96	3.97	4.01	3.96
3.84e-03	2.01	2.01	1.98	2.99	3.02	2.98	3.95	3.98	3.98
1.92e-03	2.00	2.01	1.98	2.97	2.98	2.98	3.98	3.98	3.98
9.60e-04	1.99	1.99	1.99	2.98	2.98	2.99	3.98	3.99	3.98

6.2 Non-degenerate potential

We consider $\delta = 0$ (the p-Laplacian case), and the potential u is given by

$$u(x,y) = \sin(\pi x)\sin(\pi y) + (\pi+1)(x+y) \quad \forall (x,y) \in \Omega.$$

Since $|\nabla u| \geq 1$ on Ω, $\tilde{\eta}_h = \mu_h(1)$ and we should observe a constant convergence rate of $k+1$, which is indeed the case (see Figure 2 and Table 2).

6.3 Non-degenerate flux-potential couple

The exact solution u is given by

$$u(x,y) = \sin(\pi x)\sin(\pi y) \quad \forall (x,y) \in \Omega.$$

Since ∇u vanishes at points $(x_i, y_i)_{1 \leq i \leq 5} := \{(0,0), (1,0), (0,1), (1,1), (0.5, 0.5)\}$, we consider a degeneracy parameter function δ as the sum of bump functions centered at these points with 0.2 radius. Specifically,

$$\delta(x,y) = \sum_{i=1}^{5} \begin{cases} \exp\left(1 - \dfrac{1}{1 - 25((x-x_i)^2 + (y-y_i)^2)}\right) & \text{if } \sqrt{(x-x_i)^2 + (y-y_i)^2} < 0.2, \\ 0 & \text{otherwise.} \end{cases} \tag{53}$$

As a consequence, δ vanishes on about three quarters of Ω, however, $\tilde{\eta}_h = \mu_h(1)$ and we should observe a constant convergence rate of $k+1$. This is confirmed by the results presented in Figure 3 and Table 3.

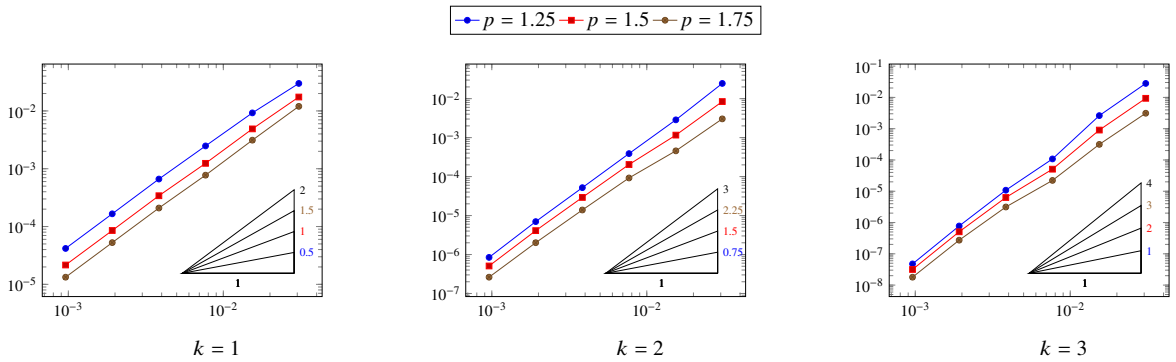

Figure 3: Numerical results for the test case of Subsection 6.3. The steeper slope (in black) indicates the $k+1$ convergence rate. The other slopes indicate the $(k+1)(p-1)$ convergence rate according to p.

Table 3: Convergence rates for the test case of Subsection 6.3. The bold numbers in each column correspond to the $(k+1)(p-1) \sim (k+1)$ convergence rates.

k	1			2			3		
h \ p	1.25	1.5	1.75	1.25	1.5	1.75	1.25	1.5	1.75
3.07e-02	**0.5 ~ 2**	**1 ~ 2**	**1.5 ~ 2**	**0.75 ~ 3**	**1.5 ~ 3**	**2.25 ~ 3**	**1 ~ 4**	**2 ~ 4**	**3 ~ 4**
1.54e-02	1.70	1.84	1.95	3.12	2.87	2.74	3.43	3.39	3.32
7.68e-03	1.89	1.97	2.00	2.87	2.50	2.30	4.58	4.14	3.81
3.84e-03	1.90	1.86	1.89	2.90	2.80	2.73	3.32	3.01	2.82
1.92e-03	2.00	2.00	1.99	2.90	2.83	2.79	3.80	3.64	3.52
9.60e-04	1.99	1.98	1.99	3.05	3.01	2.94	4.03	4.00	3.93

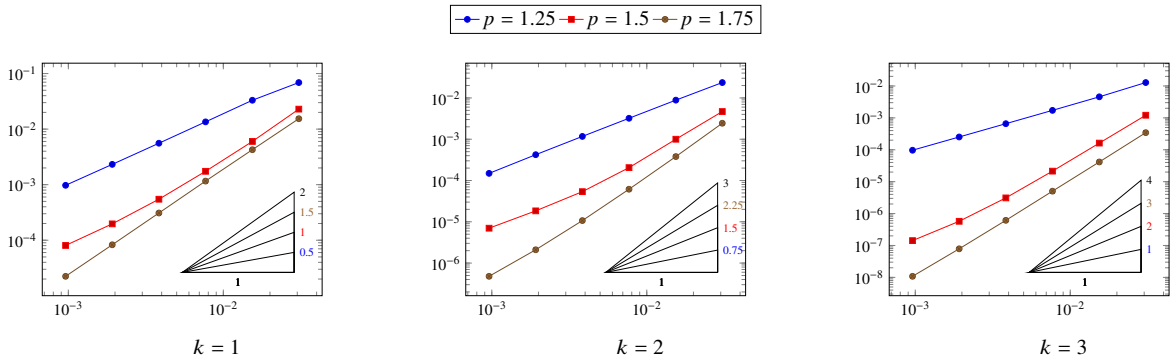

Figure 4: Numerical results for the test case of Subsection 6.4. The steeper slope (in black) indicates the $k+1$ convergence rate. The other slopes indicate the $(k+1)(p-1)$ convergence rate according to p.

Table 4: Convergence rates for the test case of Subsection 6.4. The bold numbers in each column correspond to the $(k+1)(p-1) \sim (k+1)$ convergence rates.

k	1			2			3		
h \ p	1.25	1.5	1.75	1.25	1.5	1.75	1.25	1.5	1.75
3.07e-02	**0.5 ~ 2**	**1 ~ 2**	**1.5 ~ 2**	**0.75 ~ 3**	**1.5 ~ 3**	**2.25 ~ 3**	**1 ~ 4**	**2 ~ 4**	**3 ~ 4**
1.54e-02	1.06	1.94	1.86	1.42	2.24	2.70	1.49	2.91	3.98
7.68e-03	1.29	1.78	1.87	1.45	2.28	2.62	1.40	2.91	3.96
3.84e-03	1.28	1.67	1.91	1.46	1.93	2.52	1.39	2.79	3.98
1.92e-03	1.26	1.47	1.90	1.47	1.55	2.35	1.38	2.44	3.98
9.60e-04	1.25	1.29	1.89	1.50	1.39	2.14	1.38	2.01	3.98

6.4 Degenerate problem

We consider $\delta = 0$ (the p-Laplacian case), and the potential u is given such that for all $(x, y) \in \Omega$,

$$u(x,y) = \frac{1}{10} \exp\left(-10\left(|x-0.5|^{p+\frac{k+2}{4}} + |y-0.5|^{p+\frac{k+2}{4}}\right)\right) \quad \forall (x,y) \in \Omega.$$

The particular choice of u, which changes with p and k, is driven by the need to ensure that the function and its flux have the required regularity for the error estimate in Theorem 11 to be valid (and thus to potentially avoid some of the issues observed in Section 6.1), all the while not being too smooth or with a simple structure, which might artificially generate a better convergence of the scheme. Since ∇u vanishes on an entire region of the domain and $\delta = 0$, we have, $\tilde{\eta}_h = +\infty$ and we should observe $(k+1)(p-1)$ asymptotic convergence rates. The results for this test are presented in Figure 4 and Table 4. In most cases, they do confirm that the asymptotic rate of convergence is closer to $(k+1)(p-1)$ than $k+1$, with major exceptions for $(p,k) = (1.25, 1)$ and $(p,k) = (1.75, 3)$, for which the observed rate is closer to $k+1$ – probably because, for the considered mesh sizes and parameter values, the first term in (34) might be dominant due to a larger multiplicative constant. Another explanation could be that the asymptotic regime is not reached yet for these tests, or, in view of the recent results in [**carstensen:20**], that the error estimate in Theorem 11, which is valid in a more general setting, is actually sub-optimal for these particular test cases. In any case, these tests are outliers in the results presented here, which otherwise support rather well the theoretical error estimate.

7 Conclusion

We have presented and analysed a Hybrid High-Order scheme of arbitrary order k, for a non-linear model that generalises the p-Laplace equation (with $p \in (1, 2]$) through the addition of an offset in the flux, that potentially remove its singularity at 0. Our error estimate highlights various convergence regimes for the scheme, depending on its degeneracy or lack thereof (the latter occurring in presence of a non-zero offset, or when the gradient of the continuous solution does not vanish); for a degenerate model we recover the known rates of convergence in $(k + 1)(p - 1)$, while an optimal rate of $(k + 1)$, identical to the rate for linear models, is obtained when the model is not degenerate. These regimes are locally driven by a dimensionless number, and intermediate regimes are also identified.

Several numerical tests have been provided, and show a good agreement with the theoretical error estimate, except in a few cases where the convergence appears to be faster than expected (which could be due to the asymptotic regime not yet being attained in that case, or to the specifics of the particular degenerate test case considered here).

References

[1] P. F. Antonietti, N. Bigoni, and M. Verani. "Mimetic finite difference approximation of quasilinear elliptic problems". In: *Calcolo* 52 (1 2014), pp. 45–67. DOI: 10.1007/s10092-014-0107-y.

[2] J. W. Barrett and W. B. Liu. "Finite element approximation of the p-Laplacian". In: *Math. Comp.* 61.204 (1993), pp. 523–537. DOI: 10.2307/2153239.

[3] L. Beirão da Veiga, F. Brezzi, A. Cangiani, G. Manzini, L. D. Marini, and A. Russo. "Basic principles of virtual element methods". In: *Math. Models Methods Appl. Sci. (M3AS)* 199.23 (2013), pp. 199–214. DOI: 10.1142/S0218202512500492.

[4] L. Belenki, L. Diening, and C. Kreuzer. "Optimality of an adaptive finite element method for the p-Laplacian equation". In: *IMA Journal of Numerical Analysis* 32.2 (2012), pp. 484–510.

[5] L. Botti, D. A. Di Pietro, and J. Droniou. "A Hybrid High-Order discretisation of the Brinkman problem robust in the Darcy and Stokes limits". In: *Comput. Methods Appl. Mech. Engrg.* 341 (2018), pp. 278–310. DOI: 10.1016/j.cma.2018.07.004.

[6] M. Botti, D. Castanon Quiroz, D. A. Di Pietro, and A. Harnist. *A Hybrid High-Order method for creeping flows of non-Newtonian fluids*. Submitted. 2020. URL: https://hal.archives-ouvertes.fr/hal-02519233.

[7] D. A. Di Pietro and J. Droniou. "$W^{s,p}$-approximation properties of elliptic projectors on polynomial spaces, with application to the error analysis of a hybrid high-order discretisation of Leray-Lions problems". In: *Math. Models Methods Appl. Sci.* 27.5 (2017), pp. 879–908. DOI: 10.1142/S0218202517500191.

[8] D. A. Di Pietro and J. Droniou. "A Hybrid High-Order method for Leray–Lions elliptic equations on general meshes". In: *Math. Comp.* 86.307 (2017), pp. 2159–2191. DOI: 10.1090/mcom/3180.

[9] D. A. Di Pietro and J. Droniou. *The Hybrid High-Order method for polytopal meshes. Design, analysis, and applications*. Modeling, Simulation and Application 19. Springer International Publishing, 2020. ISBN: 978-3-030-37202-6 (Hardcover) 978-3-030-37203-3 (eBook). DOI: 10.1007/978-3-030-37203-3.

[10] D. A. Di Pietro, J. Droniou, and G. Manzini. "Discontinuous Skeletal Gradient Discretisation methods on polytopal meshes". In: *J. Comput. Phys.* 355 (2018), pp. 397–425. DOI: 10.1016/j.jcp.2017.11.018.

[11] 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.

[12] R. Glowinski and A. Marrocco. “Sur l’approximation, par éléments finis d’ordre un, et la résolution, par pénalisation-dualité, d’une classe de problèmes de Dirichlet non linéaires”. In: *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge Anal. Numér.* 9.R-2 (1975), pp. 41–76.

[13] R. Glowinski and J. Rappaz. “Approximation of a nonlinear elliptic problem arising in a non-Newtonian fluid flow model in glaciology”. In: *M2AN Math. Model. Numer. Anal.* 37.1 (2003), pp. 175–186. DOI: 10.1051/m2an:2003012.

[14] A. Hirn. “Approximation of the p-Stokes equations with equal-order finite elements”. In: *J. Math. Fluid Mech.* 15.1 (2013), pp. 65–88. DOI: 10.1007/s00021-012-0095-0.

[15] J.-L. Lions and E. Magenes. *Non-homogeneous boundary value problems and applications. Vol. I*. Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 181. Springer-Verlag, New York-Heidelberg, 1972, pp. xvi+357.

[16] W. Liu and N. Yan. “Quasi-norm a priori and a posteriori error estimates for the nonconforming approximation of p-Laplacian”. In: *Numer. Math.* 89 (2001), pp. 341–378.