

# UNIVERSIDAD DE CONCEPCIÓN



## CENTRO DE INVESTIGACIÓN EN INGENIERÍA MATEMÁTICA (CI<sup>2</sup>MA)



An a priori error analysis for a class of nonlinear elliptic  
problems applying the hybrid high-order method

ROMMEL BUSTINZA, JONATHAN MUNGUÍA

PREPRINT 2020-08

### SERIE DE PRE-PUBLICACIONES



# An a priori error analysis for a class of nonlinear elliptic problems applying the hybrid high-order method\*

ROMMEL BUSTINZA<sup>†</sup> and JONATHAN MUNGUÍA-LA-COTERA<sup>‡</sup>

## Abstract

In this paper, we analyse a nonlinear elliptic problem in a bounded domain, applying the already known hybrid high-order (HHO) method. Our analysis follows known approaches to deal with diffusion linear problems, and take into account the nonlinear works in elasticity. This approach allows us to obtain high-order approximation of unknowns, by assembling a non-conforming discrete space based on degrees of freedom over the interior volume of each element and on the faces of its boundary, through  $L^2$ -projections. The principal feature is the use of a gradient reconstruction operator, whose codomain in the current context, is bigger than the considered in the linear case. On the other hand, the stabilization term is the same as for the linear case. In addition, the construction does not depend on the spatial dimension, and it offers the possibility to use general meshes with polytopal cells and non-matching interfaces. The discrete unknowns are chosen such that they are supported over each element and its boundary. It is important to emphasize that the cell-based unknowns can be eliminated locally by a Schur complement technique (also known as static condensation). This allows us to obtain a more compact system, reducing significantly its size, which is solved by Newton's method. We establish the well-posedness of the nonlinear discrete scheme, the stability of the method, and confirm that it is optimally convergent in the energy norm and in the  $L^2$ -norm, for the potential and its gradient, for smooth enough solutions. In addition, we also establish the convergence of the  $L^2$ -projection of the potential error and the super convergence of the reconstructive potential error, under the same additional regularity assumption of the exact solution. The latter results, up to the author's knowledge, have not been proven before. Several numerical experiments confirm our theoretical results. We point out that this approach can be also applied/extended to deal with other boundary conditions (such as nonhomogeneous Dirichlet, mixed or pure Neumann).

---

\*This research has been partially supported by CONICYT / PIA / Concurso Apoyo a Centros Científicos y Tecnológicos de Excelencia con Financiamiento Basal AFB170001, by Vicerrectoría de Investigación y Desarrollo de la Universidad de Concepción (Chile) through project VRID-Enlace No. 218.013.044-1.0, and by Centro de Investigación en Ingeniería Matemática (CI<sup>2</sup>MA), Universidad de Concepción (Chile), by CONCYTEC-Perú through FONDECYT project "Programas de Doctorado en Universidades Peruanas" CG-176-2015, Instituto de Matemática y Ciencias Afines (IMCA), and to Universidad Nacional de Ingeniería (Lima-Perú)

<sup>†</sup>Centro de Investigación en Ingeniería Matemática (CI<sup>2</sup>MA) and Departamento de Ingeniería Matemática, Universidad de Concepción, Concepción, Chile, e-mail: rbustinz@ing-mat.udec.cl

<sup>‡</sup>Instituto de Matemática y Ciencias Afines (IMCA) and Universidad Nacional de Ingeniería (UNI), Lima, Perú, e-mail: jmunguia1@uni.edu.pe

**Keywords:** nonlinear elliptic problems, Polytopal Meshes, Hybrid High-Order, Gradient Reconstruction.

# 1 Introduction

The numerical resolution of nonlinear elliptic partial differential equations (PDEs) is nowadays of great interest in the engineering practice. These arise in many areas, including differential geometry (the Monge-Ampère equation) [28], mass transportation (the Monge-Kantorovich problem) [1, 43], dynamic programming (Bellman’s equation), fluid dynamics (the geostrophic equations) [40], phase change problems governed by Stefan’s model [29], the modelling of phenomena such as: glacier motion [35], incompressible turbulent flows in porous media [22], and airfoil design [34]. We consider a class of nonlinear elliptic problems arising in heat conduction and fluid mechanics, that reads as: *Find  $u$  such that*

$$-\operatorname{div}(\mathbf{a}(\cdot, \nabla u)) = f \quad \text{in } \Omega, \quad (1a)$$

$$u = 0 \quad \text{on } \partial\Omega, \quad (1b)$$

where  $\Omega \subset \mathbb{R}^d$ ,  $d \in \{2, 3\}$ , is a polytopal bounded connected domain, with Lipschitz boundary  $\partial\Omega$ ,  $\mathbf{a} : \Omega \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  is a nonlinear function, and  $f$  represents the source term. The design of convergent numerical schemes for the solution of such equations, is thus an important and very active research topic. In this work, we focus in particular, in the well-known HHO method, under certain hypothesis on nonlinear function  $\mathbf{a}$  that will be described in the Section 2. The homogeneous Dirichlet boundary condition (1b) is considered only for the sake of simplicity. We can modify appropriately the mathematical analysis, to deal with more general boundary conditions such as mixed boundary conditions, non-homogeneous Dirichlet boundary condition, and also Neumann boundary condition.

This approach has several advantageous features, initially described in [19, 18]: (i) it is independent of the spatial dimension; (ii) it supports arbitrary polynomial orders  $k \geq 0$  on fairly general meshes including, e.g., polyhedral elements and non-matching interfaces; (iii) it is devised from local reconstruction operators and a linear local stabilization term; (iv) it is efficiently parallelable (the local stencil only connects a mesh element with its faces); (v) it has moderate computational costs, thanks to the possibility of locally eliminating, by static condensation, the volumetric unknowns for linearized versions of the problem (encountered, e.g., when solving the corresponding system of nonlinear algebraic equations by the Newton method or Kačanov’s one).

It is important to point out that a kind of nonlinear elasticity problem, under weak hypotheses, has been analysed with HHO approach in [6]. For a more general context, we refer to [13, 14], where the authors analyse the  $p$ -Laplacian problem in Banach spaces, with a stabilization term depending also of the nonlinearity. In the case of  $p = 2$ , we obtain the linear diffusion problem described in [21], but with a slightly different definition for the reconstruction operator. The case of  $p$ -Laplacian and Leray–Lions equations with pure Neumann boundary conditions is analysed in [20, Chapter 7]. Albeit, in [13, 6] we can find part of the HHO analysis for (1), in this

paper we propose another mathematical analysis (taking into account that we are dealing only with the case  $p = 2$ ). Moreover, we are able to deduce an a priori error estimate for  $L^2$ -error of reconstructive potential, which, up to author's knowledge, has not been established yet.

Conforming finite element methods on standard meshes have been considered in [31, 33], in the context of elasticity for mixed formulation, while in [3], the authors deal with the  $p$ -laplacian applying a continuous piecewise linear finite element approximation. Discontinuous Galerkin (DG) methods on standard meshes have been considered in [42], while in [8], the authors apply local discontinuous Galerkin (LDG) method for the same class of nonlinear elliptic problem we study. Polytopal meshes have been considered in DG methods in [4], and also with in Virtual Element (VE) method [47], based on a low-order approximation of the displacement field.

The lowest-order version of the HHO method ( $k = 0$ ), can be linked with Hybrid Finite Volume (HFV) methods in [19], when face unknowns are not eliminated by interpolation. Also, HHO approach is related to Hybrid Mixed Mimetic (HMM) methods in [15]. Here, HMM methods can be seen as the unified formulation for the Hybrid Finite Volume (HFV, [26]) methods, the Mixed Finite Volume (MFV, [24]) methods, and the Mimetic Finite Difference (MFD, [46]) methods proposed in [44, 25]. It is important to mention that in [14], the authors extend the analysis proposed in [23] for Leray–Lions equations. Moreover, in the case of fully nonlinear models, one strategy to recover the high-order rate of convergence, seems to be defining the gradient reconstruction operator in the full polynomial space  $\mathbb{P}^k(T)^d$ . For more details, we refer to Section 4 in [16].

There is also a connection of HHO technique with other approaches, such as Hybridizable Discontinuous Galerkin (HDG) methods in [12, Section 4], with the help of an equivalent mixed formulation is necessary. Also, in [12, Section 2.4], we find a link with High-order mimetic (HOM) approach, introduced in [39], and with the non-conforming version of the VE method introduced in [2]. In addition, we find a unified analysis between HHO and VE methods in [38], but differs from the standard VE method described in [45].

Finally, in [16], the authors show that the HHO method [19, 13], the non-conforming versions of MFD [39] and VE [45, 10] methods, can be included in the generic framework of Gradient scheme (GS), which was originally developed for encompassing low numerical methods ( $k = 0$ ). We point out that our analysis is not included, since the stabilization term is not embedded into the discrete symmetric gradient operator. We refer to Remark 24 in [16], for further details.

The rest of the paper is organized as follows. In Section 2, we introduce the continuous model problem and discuss the well-posedness of a weak formulation of this one. In Section 3, we introduce the discrete approximation spaces, where the unknowns of the HHO scheme will be looked for. Moreover, we also introduce the potential and gradient reconstruction operators, emphasizing their key properties. In Section 4, we introduce the discrete problem and study its stability. In Section 5, we establish the well-posedness of the discrete problem, while in Section 6 we perform the a priori error analysis. First, in the energy-norm, and then in the  $L^2$ -norm, under additional elliptic regularity assumptions. Finally, in Section 8, we present some examples, which are in agreement with our theoretical results.

## 2 Continuous setting

In this section we present the continuous problem, and give appropriate assumptions on the function  $\mathbf{a}$  that will let us to prove the well-posedness of the corresponding weak formulation. Then, given  $f \in L^2(\Omega)$ , we look for  $u : \Omega \rightarrow \mathbb{R}$ , solution of the following class of nonlinear elliptic problem:

$$-\operatorname{div}(\mathbf{a}(\cdot, \nabla u)) = f \quad \text{in } \Omega, \quad (2a)$$

$$u = 0 \quad \text{on } \partial\Omega, \quad (2b)$$

where  $\Omega \subset \mathbb{R}^d$ ,  $d \in \{2, 3\}$ , is a polytopal bounded connected domain, with Lipschitz boundary  $\partial\Omega$ , while  $\mathbf{a} : \Omega \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  is a nonlinear function.

**Assumption 2.1** *The vectorial nonlinear function  $\mathbf{a}$  satisfies the following properties/hypotheses (see [8]):*

- (H.1) *Carathéodory condition: The function  $a_j(x, \cdot)$ ,  $j = 1, \dots, d$ , is continuous in  $\mathbb{R}^d$  for almost all  $x \in \Omega$ , and  $a_j(\cdot, \zeta)$ ,  $j = 1, \dots, d$ , is measurable in  $\Omega$  for all  $\zeta \in \mathbb{R}^d$ .*
- (H.2) *Growth condition: There exist  $\phi_j \in L^2(\Omega)$ ,  $j = 1, \dots, d$ , and  $C > 0$  such that  $|a_j(x, \zeta)| \leq C(1 + |\zeta|) + |\phi_j(x)|$ , for almost all  $x \in \Omega$  and for all  $\zeta \in \mathbb{R}^d$ .*
- (H.3) *Ellipticity condition: The function  $\mathbf{a} = (a_1, \dots, a_d)$  has continuous first order partial derivatives in  $\mathbb{R}^d$ , for almost all  $x \in \Omega$ . In addition, there exists  $C > 0$  such that*

$$\sum_{i,j=1}^d \frac{\partial}{\partial \zeta_j} a_i(x, \zeta) \tilde{\zeta}_i \tilde{\zeta}_j \geq C |\tilde{\zeta}|^2,$$

*for all  $\zeta := (\zeta_i, \zeta_j)^\top$ ,  $\tilde{\zeta} := (\tilde{\zeta}_i, \tilde{\zeta}_j)^\top \in \mathbb{R}^2$  and for almost all  $x \in \Omega$ .*

- (H.4) *Regularity condition: The function  $a_i(x, \cdot)$ ,  $i = 1, \dots, d$ , has continuous first order partial derivatives in  $\mathbb{R}^d$  for almost all  $x \in \Omega$ . In addition, there exists  $C > 0$  such that for each  $i, j \in \{1, \dots, d\}$ ,  $\frac{\partial}{\partial \zeta_j} a_i(x, \zeta)$  satisfies the Carathéodory condition (H.1), and  $\left| \frac{\partial}{\partial \zeta_j} a_i(x, \zeta) \right| \leq C$ , for all  $\zeta \in \mathbb{R}^d$ , and for almost all  $x \in \Omega$ .*

The weak formulation of problem (2), that will serve as a starting point for the development and analysis of the HHO method, reads: Find  $u \in U := H_0^1(\Omega)$  such that

$$\int_{\Omega} \mathbf{a}(x, \nabla u(x)) \cdot \nabla v(x) dx = \int_{\Omega} f(x) v(x) dx \quad \forall v \in U. \quad (3)$$

We remark that (H.1) - (H.2) guarantee that  $\mathbf{a}$  is well defined, while (H.3) - (H.4) ensure that the nonlinear operator in the left-hand side of (3) is strongly monotone and Lipschitz-continuous, thanks to Theorem 32.6 in [48, page 240]. Finally, the well-posedness of the problem (3) follows after invoking Theorem 34.1 in [48, page 245].

### 3 Discrete settings

In this section, we recall the notion of admissible mesh sequences, and reasonable assumptions on such meshes, which are required to define the HHO scheme. First, we let  $(\mathcal{T}_h)_{h \in \mathcal{H}}$  be an  $h$ -refined regular mesh sequence of  $\bar{\Omega}$ , consisting of polytopal meshes (cf. Chapter 1 in [20]), with  $\mathcal{H}$  being a countable set of positive numbers having 0 as its only accumulation point. Next, for any  $h \in \mathcal{H}$ ,  $\mathcal{T}_h$  is a finite collection of nonempty, disjoint, open, polytope elements  $T$  with diameter  $h_T$ , such that  $\cup_{T \in \mathcal{T}_h} \bar{T} = \bar{\Omega}$ , and there is a matching simplicial submesh  $\mathfrak{J}_h$  such that any simplex and any face in  $\mathfrak{J}_h$  belongs to just one element and face of  $\mathcal{T}_h$ . Besides, there exists  $\rho > 0$  (called the mesh regularity parameter), that is independent of  $h$ , such that for every  $h \in \mathcal{H}$ : (i) for any simplex  $S \in \mathfrak{J}_h$  of diameter  $h_S$  and inradius  $r_S$ , there holds  $\rho h_S \leq r_S$ ; and (ii)  $\forall T \in \mathcal{T}_h, \forall S \in \mathfrak{J}_h : (S \subseteq T \Rightarrow \rho h_T \leq h_S)$ . As usual, we introduce  $h := \max_{T \in \mathcal{T}_h} h_T$ . The regularity condition of the mesh sequence, ensures important geometric bounds on  $\mathcal{T}_h$  and  $\mathcal{F}_h$ , for any  $h \in \mathcal{H}$ . We refer to Lemma 1.2 in [20] for more details in this context.

Given any  $\mathcal{T}_h$ , with  $h \in \mathcal{H}$ , we call a face any hyperplanar closed connected subset  $F$  of  $\bar{\Omega}$  with positive  $(d-1)$ -dimensional measure, and such that: (1) either there exists  $T \in \mathcal{T}_h$  such that  $F \subset \partial T \cap \partial \Omega$  (in this case,  $F$  is called a boundary face), or (2) there exist  $T_1, T_2 \in \mathcal{T}_h$  such that  $F \subset \partial T_1 \cap \partial T_2$  (and then,  $F$  is called an interior face). Boundary faces are collected in  $\mathcal{F}_h^b$ , interior faces are collected in the set  $\mathcal{F}_h^{\text{int}}$ , and we set  $\mathcal{F}_h := \mathcal{F}_h^{\text{int}} \cup \mathcal{F}_h^b$ . Now, by  $h_F$  we denote the diameter of a face  $F \in \mathcal{F}_h$ . For each  $T \in \mathcal{T}_h$ , we define  $\mathcal{F}_T := \{F \in \mathcal{F}_h : F \subset \partial T\}$ , and then, given  $F \in \mathcal{F}_T$ , by  $\mathbf{n}_{TF}$  we denote the unit normal to  $F$ , outward to  $T$ .

Throughout this paper, given two different non negative numbers  $x, y$ , by  $x \lesssim y$ , we denote the inequality  $x \leq C y$ , where  $C$  is a positive constant that is independent of the nonhomogeneous, but may depend on the mesh regularity parameter  $\rho$ . In addition, we also introduce the notation  $x \approx y$ , which means that  $y \lesssim x$  and  $x \lesssim y$ .

It is important to emphasize that in this context, the very well-known inequalities and embeddings, are also valid. We remind here the discrete Poincaré-Wirtinger inequality [20, Remark 1.46], i.e. given any  $T \in \mathcal{T}_h$ :

$$\|w\|_T \lesssim h_T \|\nabla w\|_T \quad \forall w \in H^1(T) \cap L_0^2(T), \quad (4)$$

which will be invoked in this paper. Other known results are, for example, the discrete Sobolev embedding [13, Proposition 5.4], the discrete inverse inequality [17, Lemma 1.44], the discrete and continuous trace inequalities [17, Lemmas 1.46 and 1.49].

Now, we recall the definition and properties of the  $L^2$ -orthogonal projector. Given any  $T \in \mathcal{T}_h$  and any integer  $l \geq 0$ , we define the  $L^2$ -orthogonal projector  $\pi_T^l : L^1(T) \rightarrow \mathbb{P}_d^l(T)$  such that: For all  $v \in L^1(T)$ ,  $\pi_T^l v$  is the unique polynomial in  $\mathbb{P}_d^l(T)$  that satisfies

$$(\pi_T^l v - v, \nabla w)_T = 0 \quad \forall w \in \mathbb{P}_d^l(T).$$

We also introduce the so called elliptic projector  $\pi_T^{1,l} : W^{1,1}(T) \rightarrow \mathbb{P}_d^l(T)$  such that: For all  $v \in W^{1,1}(T)$ ,  $\pi_T^{1,l} v$  is a polynomial in  $\mathbb{P}_d^l(T)$  that satisfies (cf. Definition 1.39 in [20])

$$\begin{aligned} (\nabla(\pi_T^{1,l} v - v), \nabla w)_T &= 0 \quad \forall w \in \mathbb{P}_d^l(T), \\ (\pi_T^{1,l} v - v, 1)_T &= 0. \end{aligned}$$

For the next result, we set  $\pi_T^{0,l} := \pi_T^l$ .

**Lemma 3.1 (Approximation properties).** *Given an integer  $l \geq 0$ ,  $T \in \mathcal{T}_h$  and  $\kappa \in \{0, 1\}$ . Let  $\pi_T^{\kappa, l}$  be the  $L^2$ -orthogonal projection or elliptic projector onto  $\mathbb{P}_d^k(T)$ . Then, for all  $s \in \{\kappa, \dots, l+1\}$  and all  $v \in H^s(T)$ , there holds*

$$|v - \pi_T^{\kappa, l} v|_{H^m(T)} \leq C'_{app} h_T^{s-m} |v|_{H^s(T)} \quad \forall m \in \{0, \dots, s\}. \quad (5)$$

Moreover, assuming additionally that  $s \geq 1$ , there holds for all  $F \in \mathcal{F}_T$ ,

$$|v - \pi_T^{\kappa, l} v|_{H^m(F)} \leq C''_{app} h_T^{s-m-1/2} |v|_{H^s(T)} \quad \forall m \in \{0, \dots, s-1\}. \quad (6)$$

Here  $C'_{app} > 0$  and  $C''_{app} > 0$  depend only on  $d$ , the regularity of the mesh,  $l$ , and  $s$ .

*Proof.* We refer to the proofs of Theorems 1.45 and 1.48 in [20].  $\square$

### 3.1 Degrees of freedom (DOFs)

Let  $k \geq 0$  be fixed. For all  $T \in \mathcal{T}_h$ , we define the local space of DOFs as  $\underline{\mathbf{U}}_T^k := \mathbb{P}_d^k(T) \times (\prod_{F \in \mathcal{F}_T} \mathbb{P}_{d-1}^k(F))$ , where  $\mathbb{P}_d^k(T)$  (resp.,  $\mathbb{P}_{d-1}^k(F)$ ) is spanned by the restrictions to  $T$  (resp., to  $F$ ) of  $d$ -variate (resp.,  $(d-1)$ -variate) polynomials of total degree  $\leq k$ . The global space of DOFs on the domain  $\Omega$ , is

$$\underline{\mathbf{U}}_h^k := \left( \prod_{T \in \mathcal{T}_h} \mathbb{P}_d^k(T) \right) \times \left( \prod_{F \in \mathcal{F}_h} \mathbb{P}_{d-1}^k(F) \right).$$

The space  $\underline{\mathbf{U}}_h^k$  is equipped with the following discrete semi-norm,  $\|\cdot\|_{\epsilon, h} : \underline{\mathbf{U}}_h^k \rightarrow \mathbb{R}$

$$\|\underline{\mathbf{v}}_h\|_{\epsilon, h}^2 := \sum_{T \in \mathcal{T}_h} \|\underline{\mathbf{v}}_T\|_{\epsilon, T}^2, \quad \|\underline{\mathbf{v}}_T\|_{\epsilon, T}^2 := \|\nabla v_T\|_T^2 + \sum_{F \in \mathcal{F}_T} h_F^{-1} \|v_F - v_T\|_F^2. \quad (7)$$

For each  $T \in \mathcal{T}_h$ , we define the local reduction operator  $\underline{\mathbf{I}}_T^k : H^1(T) \rightarrow \underline{\mathbf{U}}_T^k$  such that, for all  $v \in H^1(T)$ ,

$$\underline{\mathbf{I}}_T^k v := (\pi_T^k v, (\pi_F^k v)_{F \in \mathcal{F}_T}) \in \underline{\mathbf{U}}_T^k, \quad (8)$$

where  $\pi_T^k$  and  $\pi_F^k$  denote the usual  $L^2$ -orthogonal projectors onto  $\mathbb{P}_d^k(T)$  and  $\mathbb{P}_{d-1}^k(F)$ , respectively. The corresponding global interpolation operator  $\underline{\mathbf{I}}_h^k : H^1(\Omega) \rightarrow \underline{\mathbf{U}}_h^k$  is then defined, such that for all  $v \in H^1(\Omega)$ ,

$$\underline{\mathbf{I}}_h^k v := ((\pi_T^k v)_{T \in \mathcal{T}_h}, (\pi_F^k v)_{F \in \mathcal{F}_h}) \in \underline{\mathbf{U}}_h^k. \quad (9)$$

### 3.2 Gradient and Potential reconstruction operators

Based on the local DOFs, we introduce reconstructions of the gradient and of the potential that will be helpful in the formulation of the method. In what follows,  $(\cdot, \cdot)_T$  and  $(\cdot, \cdot)_F$  will denote the usual  $L^2$ -inner products on  $T \in \mathcal{T}_h$  and  $F \in \mathcal{F}_h$ , respectively. We define the **local discrete gradient operator**  $G_T^k : \underline{\mathbf{U}}_T^k \rightarrow [\mathbb{P}_d^k(T)]^d$  such that, given  $\underline{\mathbf{v}}_T \in \underline{\mathbf{U}}_T^k$ , then,  $G_T^k \underline{\mathbf{v}}_T \in [\mathbb{P}_d^k(T)]^d$  satisfies

$$(G_T^k \underline{\mathbf{v}}_T, \phi)_T = (\nabla v_T, \phi)_T + \sum_{F \in \mathcal{F}_T} (v_F - v_T, \phi \cdot \mathbf{n}_{TF})_F \quad \forall \phi \in [\mathbb{P}_d^k(T)]^d, \quad (10)$$

where, we recall that  $\mathbf{n}_{TF}$  is the unit normal to  $F$  pointing out of  $T$ . The following proposition shows us the relation between the exact gradient and the discrete gradient operators.



**Proposition 3.1 (Commuting property).** *For all  $v \in H^1(T)$ , there holds*

$$G_T^k \mathbf{I}_T^k v = \pi_T^k(\nabla v), \quad (11)$$

where  $\pi_T^k$  acts componentwise.

*Proof.* For the sake of completeness, we rewrite the proof of Proposition 10 in [14]. First, we fix  $T \in \mathcal{T}_h$  and  $\phi \in [\mathbb{P}_d^k(T)]^d$ . Then, plugging the definition (8) of  $\mathbf{I}_T^k$  into (10), and applying integration by parts, taking into account the definition of orthogonal projection, together with the fact that  $\operatorname{div} \phi \in \mathbb{P}_d^{k-1}(T) \subset \mathbb{P}_d^k(T)$  and that  $\phi|_F \cdot \mathbf{n}_{TF} \in \mathbb{P}_{d-1}^k(F)$  for all  $F \in \mathcal{F}_T$ , we have

$$\begin{aligned} (G_T^k \mathbf{I}_T^k v, \phi)_T &= (\nabla \pi_T^k v, \phi)_T + \sum_{F \in \mathcal{F}_T} (\pi_F^k v - \pi_T^k v, \phi \cdot \mathbf{n}_{TF})_F \\ &= -(\pi_T^k v, \operatorname{div} \phi)_T + \sum_{F \in \mathcal{F}_T} (\pi_F^k v, \phi \cdot \mathbf{n}_{TF})_F \\ &= -(v, \operatorname{div} \phi)_T + \sum_{F \in \mathcal{F}_T} (v, \phi \cdot \mathbf{n}_{TF})_F \\ &= (\nabla v, \phi)_T, \end{aligned}$$

where, we have applied integration by parts again. Then, the conclusion is derived and we end the proof.  $\square$

**Remark 3.1** *In the Appendix of this paper, we show the relevance of introducing (10), with a numerical example.*

We also define the **local potential reconstruction operator**  $p_T^{k+1} : \underline{U}_T^k \rightarrow \mathbb{P}_d^{k+1}(T)$  such that, for all  $\underline{v}_T := (v_T, (v_F)_{F \in \mathcal{F}_T}) \in \underline{U}_T^k$ ,  $p_T^{k+1} \underline{v}_T \in \mathbb{P}_d^{k+1}(T)$  is the unique solution of

$$(\nabla p_T^{k+1} \underline{v}_T - G_T^k \underline{v}_T, \nabla w)_T = 0, \quad \forall w \in \mathbb{P}_d^{k+1}(T), \quad (12a)$$

$$(p_T^{k+1} \underline{v}_T - v_T, 1)_T = 0. \quad (12b)$$

This operator satisfies the following orthogonal property (cf. Section 4 in [13]).

**Corollary 3.1 (Euler equation).** *For all  $v \in H^1(T)$ , there holds*

$$(\nabla(p_T^{k+1} \mathbf{I}_T^k v - v), \nabla w)_T = 0 \quad \forall w \in \mathbb{P}_d^{k+1}(T). \quad (13)$$

*Proof.* We have the following identity

$$(\nabla(p_T^{k+1} \mathbf{I}_T^k v - v), \nabla w)_T = (\nabla p_T^{k+1} \mathbf{I}_T^k v - G_T^k \mathbf{I}_T^k v, \nabla w)_T + (G_T^k \mathbf{I}_T^k v - \nabla v, \nabla w)_T \quad \forall w \in \mathbb{P}_d^{k+1}(T).$$

Then, the conclusion follows from (12a) and the commuting property (11).  $\square$

**Remark 3.2** *We observe from (12b) and the definition of  $\pi_T^k$ , that  $(p_T^{k+1} \mathbf{I}_T^k v, 1)_T = (\pi_T^k v, 1)_T = (v, 1)_T$ . Then, there holds*

$$(p_T^{k+1} \mathbf{I}_T^k v - v, 1)_T = 0 \quad \forall T \in \mathcal{T}_h. \quad (14)$$

The following lemma is similar to the one established for the original Gradient reconstruction operator (defined for linear elliptic problems), and will be applied later in this paper (cf. proof of Theorem 6.3).

**Lemma 3.2 (Approximation properties for  $p_T^{k+1}\underline{\mathbf{I}}_T^k$ ).** *There exists a real number  $C > 0$ , depending on  $\rho$  (mesh regularity parameter) and  $d$ , but independent of the polynomial degree, and the meshsize. So that for any  $v \in H^{s+2}(T)$ ,  $s \in \{0, \dots, k\}$ , there holds:*

$$\begin{aligned} & \|v - p_T^{k+1}\underline{\mathbf{I}}_T^k v\|_T + h_T^{1/2} \|v - p_T^{k+1}\underline{\mathbf{I}}_T^k v\|_{\partial T} + h_T \|\nabla(v - p_T^{k+1}\underline{\mathbf{I}}_T^k v)\|_T \\ & + h_T^{3/2} \|\nabla(v - p_T^{k+1}\underline{\mathbf{I}}_T^k v)\|_{\partial T} \leq Ch_T^{s+2} \|v\|_{H^{s+2}(T)}. \end{aligned} \quad (15)$$

*Proof.* The proof follows immediately from Theorems 1.1 and 1.2 in [14], and observing that  $p_T^{k+1}\underline{\mathbf{I}}_T^k$  is an elliptic projector, thanks to the Euler equation (13), fixing its constant by (14).  $\square$

## 4 Formulation

Here, we introduce the following subspace of  $\underline{\mathbf{U}}_h^k$ , which strongly incorporates the homogeneous Dirichlet boundary condition (2b):

$$\underline{\mathbf{U}}_{h,0}^k := \left\{ \underline{\mathbf{v}}_h \in \underline{\mathbf{U}}_h^k : v_F = 0 \quad \forall F \in \mathcal{F}_h^b \right\}. \quad (16)$$

**Proposition 4.1** *The map  $\|\cdot\|_{\epsilon,h} : \underline{\mathbf{U}}_h^k \rightarrow \mathbb{R}$ , given in (7), defines a norm on  $\underline{\mathbf{U}}_{h,0}^k$ .*

*Proof.* The proof follows the same ideas from the proof of Corollary 2.16 in [20]. For the sake of completeness, we proceed to establish this result. First, we notice that it is enough to prove

$$\forall \underline{\mathbf{v}}_h := ((v_T)_{T \in \mathcal{T}_h}, (v_F)_{F \in \mathcal{F}_h}) \in \underline{\mathbf{U}}_{h,0}^k : \left( \|\underline{\mathbf{v}}_h\|_{\epsilon,h} = 0 \quad \Rightarrow \quad \underline{\mathbf{v}}_h = 0 \in \underline{\mathbf{U}}_{h,0}^k \right).$$

We let  $\underline{\mathbf{v}}_h := ((v_T)_{T \in \mathcal{T}_h}, (v_F)_{F \in \mathcal{F}_h}) \in \underline{\mathbf{U}}_{h,0}^k$  be such that  $\|\underline{\mathbf{v}}_h\|_{\epsilon,h} = 0$ . This implies, from (7), that there exists  $C_0 \in \mathbb{R}$ , such that

$$v_F = v_T = C_0 \quad \forall F \in \mathcal{F}_T, \quad \forall T \in \mathcal{T}_h.$$

Since  $v_F$  is single value on each  $F \in \mathcal{F}_h$  and  $v_F = 0$  on  $\mathcal{F}_h^b$ , we conclude that  $\underline{\mathbf{v}}_h = 0 \in \underline{\mathbf{U}}_{h,0}^k$ , and we end the proof.  $\square$

Hereafter, we recall that given  $\underline{\mathbf{v}}_h := ((v_T)_{T \in \mathcal{T}_h}, (v_F)_{F \in \mathcal{F}_h}) \in \underline{\mathbf{U}}_h^k$ , we set  $v_h \in \mathbb{P}_d^k(\mathcal{T}_h)$  such that  $v_h|_T := v_T$ , for all  $T \in \mathcal{T}_h$ .

Now, we consider the following discrete variational formulation of (3): Find  $\underline{\mathbf{u}}_h \in \underline{\mathbf{U}}_{h,0}^k$  such that, for any  $\underline{\mathbf{v}}_h \in \underline{\mathbf{U}}_{h,0}^k$ ,

$$A_h(\underline{\mathbf{u}}_h, \underline{\mathbf{v}}_h) := N_h(\underline{\mathbf{u}}_h, \underline{\mathbf{v}}_h) + s_h(\underline{\mathbf{u}}_h, \underline{\mathbf{v}}_h) = (f, v_h)_{0,\Omega} =: b_h(\underline{\mathbf{v}}_h), \quad (17)$$

where the nonlinear form  $N_h : \underline{\mathbf{U}}_h^k \times \underline{\mathbf{U}}_h^k \rightarrow \mathbb{R}$  and the bilinear form  $s_h : \underline{\mathbf{U}}_h^k \times \underline{\mathbf{U}}_h^k \rightarrow \mathbb{R}$ , are assembled element-wise as

$$N_h(\underline{\mathbf{u}}_h, \underline{\mathbf{v}}_h) := \sum_{T \in \mathcal{T}_h} N_T(\underline{\mathbf{u}}_T, \underline{\mathbf{v}}_T), \quad \text{and} \quad s_h(\underline{\mathbf{u}}_h, \underline{\mathbf{v}}_h) := \sum_{T \in \mathcal{T}_h} s_T(\underline{\mathbf{u}}_T, \underline{\mathbf{v}}_T), \quad (18)$$

from the local contributions  $N_T : \underline{\mathbf{U}}_T^k \times \underline{\mathbf{U}}_T^k \rightarrow \mathbb{R}$ , and  $s_T : \underline{\mathbf{U}}_T^k \times \underline{\mathbf{U}}_T^k \rightarrow \mathbb{R}$ ,  $T \in \mathcal{T}_h$ , given by

$$N_T(\underline{\mathbf{u}}_T, \underline{\mathbf{v}}_T) := \int_T \mathbf{a}(x, G_T^k \underline{\mathbf{u}}_T(x)) \cdot G_T^k \underline{\mathbf{v}}_T(x) dx, \quad (19a)$$

$$s_T(\underline{\mathbf{u}}_T, \underline{\mathbf{v}}_T) := \sum_{F \in \mathcal{F}_T} h_F^{-1} \left( \pi_F^k(u_F - P_T^k \underline{\mathbf{u}}_T), \pi_F^k(v_F - P_T^k \underline{\mathbf{v}}_T) \right)_F. \quad (19b)$$

We recall that the local potential reconstruction  $P_T^k : \underline{\mathbf{U}}_T^k \rightarrow \mathbb{P}_d^{k+1}(T)$  is defined such that, for all  $\underline{\mathbf{v}}_T \in \underline{\mathbf{U}}_T^k$ ,

$$P_T^k \underline{\mathbf{v}}_T := v_T + (p_T^{k+1} \underline{\mathbf{v}}_T - \pi_T^k p_T^{k+1} \underline{\mathbf{v}}_T). \quad (20)$$

Next, we introduce the global discrete gradient operator  $G_h^k : \underline{\mathbf{U}}_h^k \rightarrow [\mathbb{P}_d^k(\mathcal{T}_h)]^d$  such that, for all  $\underline{\mathbf{v}}_h \in \underline{\mathbf{U}}_h^k$ ,

$$G_h^k \underline{\mathbf{v}}_h|_T := G_T^k \underline{\mathbf{v}}_T \quad \forall T \in \mathcal{T}_h. \quad (21)$$

**Lemma 4.1** *Let  $(\mathcal{T}_h)_{h \in \mathcal{H}}$  be an admissible mesh sequence and  $k \geq 0$ . For any  $T \in \mathcal{T}_h$ , there holds the following equivalence of local semi norms on  $\underline{\mathbf{U}}_T^k$*

$$\|\underline{\mathbf{v}}_T\|_{\epsilon, T}^2 \approx \|G_T^k \underline{\mathbf{v}}_T\|_T^2 + s_T(\underline{\mathbf{v}}_T, \underline{\mathbf{v}}_T) \quad \forall \underline{\mathbf{v}}_T \in \underline{\mathbf{U}}_T^k \quad (22)$$

$$\approx \|p_T^{k+1} \underline{\mathbf{v}}_T\|_T^2 + s_T(\underline{\mathbf{v}}_T, \underline{\mathbf{v}}_T) \quad \forall \underline{\mathbf{v}}_T \in \underline{\mathbf{U}}_T^k. \quad (23)$$

Besides, the hidden constant is independent of the meshsize and  $\underline{\mathbf{v}}_T$ . Consequently, for all  $\underline{\mathbf{v}}_h \in \underline{\mathbf{U}}_h^k$ , there holds

$$\|\underline{\mathbf{v}}_h\|_{\epsilon, h}^2 \approx \|G_h^k \underline{\mathbf{v}}_h\|_{0, \Omega}^2 + s_h(\underline{\mathbf{v}}_h, \underline{\mathbf{v}}_h). \quad (24)$$

*Proof.* We refer to the proofs of Lemma 4 in [19] and Lemma 5.2 in [13].  $\square$

**Remark 4.1**  $\|G_h^k \underline{\mathbf{v}}_h\|_{0, \Omega}^2 + s_h(\underline{\mathbf{v}}_h, \underline{\mathbf{v}}_h)$  is usually called energy norm.

**Lemma 4.2** *The nonlinear form  $N_h$ , defined in (18), is strongly monotone on  $\underline{\mathbf{U}}_h^k$ .*

*Proof.* The strong monotonicity of  $N_h$  follows straightforwardly from (22) and (H.3).  $\square$

**Lemma 4.3 (Consistency of  $s_T$ ).** *Given  $T \in \mathcal{T}_h$ , we let  $s_T$  be the stabilization bilinear form defined in (18). Then, for each  $v \in H^{r+2}(T)$ , with  $r \in \{0, \dots, k\}$ , there holds*

$$s_T(\underline{\mathbf{I}}_T^k v, \underline{\mathbf{I}}_T^k v)^{1/2} \lesssim h_T^{r+1} \|v\|_{H^{r+2}(T)}. \quad (25)$$

*Proof.* We refer to the proof of Lemma 2.14 in [20]. We omit further details.  $\square$

## 5 Well-posedness of (17)

Our purpose here, is to establish the unique solvability of (17). To this aim, we notice that the assumptions **(H.3)** and **(H.4)** on the coefficients  $a_i$ , imply that the nonlinear operator induced by  $\mathbf{a}$  is strongly monotone and Lipschitz continuous on  $[L^2(\Omega)]^d$  (see, e.g. [30, 32]). This ensures the existence of  $C_1, C_2 > 0$ , such that for all  $\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}} \in [L^2(\Omega)]^d$ , there hold

$$\int_{\Omega} \left( \mathbf{a}(\cdot, \boldsymbol{\theta}) - \mathbf{a}(\cdot, \tilde{\boldsymbol{\theta}}) \right) \cdot (\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}}) \geq C_1 \|\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}}\|_{0,\Omega}^2, \quad (26)$$

and

$$\|\mathbf{a}(\cdot, \boldsymbol{\theta}) - \mathbf{a}(\cdot, \tilde{\boldsymbol{\theta}})\|_{0,\Omega} \leq C_2 \|\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}}\|_{0,\Omega}. \quad (27)$$

Next result shows that the nonlinear operator  $A_h$  is also Lipschitz continuous.

**Lemma 5.1** *There exists  $C_{\text{LC}} > 0$ , independent of the meshsize, such that*

$$\|A_h(\underline{\mathbf{v}}_h, \cdot) - A_h(\underline{\mathbf{w}}_h, \cdot)\|_{\underline{\mathbf{U}}_{h,0}^{k,*}} \leq C_{\text{LC}} \|\underline{\mathbf{v}}_h - \underline{\mathbf{w}}_h\|_{\epsilon,h} \quad \forall \underline{\mathbf{v}}_h, \underline{\mathbf{w}}_h \in \underline{\mathbf{U}}_{h,0}^k, \quad (28)$$

where  $\underline{\mathbf{U}}_{h,0}^{k,*}$  represents the dual of  $\underline{\mathbf{U}}_{h,0}^k$ .

*Proof.* Given  $\underline{\mathbf{v}}_h, \underline{\mathbf{w}}_h, \underline{\mathbf{z}}_h \in \underline{\mathbf{U}}_{h,0}^k$ , we notice that

$$A_h(\underline{\mathbf{v}}_h, \underline{\mathbf{z}}_h) - A_h(\underline{\mathbf{w}}_h, \underline{\mathbf{z}}_h) = \sum_{T \in \mathcal{T}_h} \left[ \left( \mathbf{a}(\cdot, G_T^k \underline{\mathbf{v}}_T) - \mathbf{a}(\cdot, G_T^k \underline{\mathbf{w}}_T), G_T^k \underline{\mathbf{z}}_T \right)_T + s_T(\underline{\mathbf{v}}_T - \underline{\mathbf{w}}_T, \underline{\mathbf{z}}_T) \right].$$

Now, noticing that **(H.4)** also implies that  $\mathbf{a}$  is Lipschitz continuous on  $[L^2(T)]^d$ , for each  $T \in \mathcal{T}_h$ , we have, after applying Cauchy-Schwarz inequality, that

$$|A_h(\underline{\mathbf{v}}_h, \underline{\mathbf{z}}_h) - A_h(\underline{\mathbf{w}}_h, \underline{\mathbf{z}}_h)| \lesssim \sum_{T \in \mathcal{T}_h} \|G_T^k(\underline{\mathbf{v}}_T - \underline{\mathbf{w}}_T)\|_T \|G_T^k \underline{\mathbf{z}}_T\|_T + s_T(\underline{\mathbf{v}}_T - \underline{\mathbf{w}}_T, \underline{\mathbf{z}}_T).$$

Next, we take into account Minkowski inequality, to derive

$$|A_h(\underline{\mathbf{v}}_h, \underline{\mathbf{z}}_h) - A_h(\underline{\mathbf{w}}_h, \underline{\mathbf{z}}_h)| \lesssim \left( \sum_{T \in \mathcal{T}_h} \|G_T^k(\underline{\mathbf{v}}_T - \underline{\mathbf{w}}_T)\|_T^2 + s_T(\underline{\mathbf{v}}_T - \underline{\mathbf{w}}_T, \underline{\mathbf{v}}_T - \underline{\mathbf{w}}_T) \right)^{1/2} \left( \sum_{T \in \mathcal{T}_h} \|G_T^k \underline{\mathbf{z}}_T\|_T^2 + s_T(\underline{\mathbf{z}}_T, \underline{\mathbf{z}}_T) \right)^{1/2}.$$

Then, (28) follows after invoking Lemma 4.1.  $\square$

**Lemma 5.2** *There exists  $C_{\text{SM}} > 0$ , independent of the meshsize, such that*

$$A_h(\underline{\mathbf{v}}_h, \underline{\mathbf{v}}_h - \underline{\mathbf{w}}_h) - A_h(\underline{\mathbf{w}}_h, \underline{\mathbf{v}}_h - \underline{\mathbf{w}}_h) \geq C_{\text{SM}} \|\underline{\mathbf{v}}_h - \underline{\mathbf{w}}_h\|_{\epsilon,h}^2 \quad \forall \underline{\mathbf{v}}_h, \underline{\mathbf{w}}_h \in \underline{\mathbf{U}}_h^k. \quad (29)$$

*Proof.* Let  $\underline{\mathbf{v}}_h, \underline{\mathbf{w}}_h \in \underline{\mathbf{U}}_h^k$ . Thanks to the definition of  $A_h$ , we derive

$$\begin{aligned} & A_h(\underline{\mathbf{v}}_h, \underline{\mathbf{v}}_h - \underline{\mathbf{w}}_h) - A_h(\underline{\mathbf{w}}_h, \underline{\mathbf{v}}_h - \underline{\mathbf{w}}_h) \\ &= \sum_{T \in \mathcal{T}_h} \left( \mathbf{a}(\cdot, G_T^k \underline{\mathbf{v}}_T) - \mathbf{a}(\cdot, G_T^k \underline{\mathbf{w}}_T), G_T^k \underline{\mathbf{v}}_T - G_T^k \underline{\mathbf{w}}_T \right)_T + s_h(\underline{\mathbf{v}}_h - \underline{\mathbf{w}}_h, \underline{\mathbf{v}}_h - \underline{\mathbf{w}}_h). \end{aligned}$$

Next, applying (26), we obtain

$$A_h(\underline{\mathbf{v}}_h, \underline{\mathbf{v}}_h - \underline{\mathbf{w}}_h) - A_h(\underline{\mathbf{w}}_h, \underline{\mathbf{v}}_h - \underline{\mathbf{w}}_h) \geq C_1 \sum_{T \in \mathcal{T}_h} \|G_T^k(\underline{\mathbf{v}}_T - \underline{\mathbf{w}}_T)\|_T^2 + s_h(\underline{\mathbf{v}}_h - \underline{\mathbf{w}}_h, \underline{\mathbf{v}}_h - \underline{\mathbf{w}}_h).$$

Finally, applying Lemma 4.1 again, (29) is concluded.  $\square$

In order to exhibit that  $b_h \in \underline{\mathbf{U}}_{h,0}^{k,*}$ , we require the following result.

**Lemma 5.3** *There exists  $C_P > 0$ , independent of the meshsize, such that*

$$\|v_h\|_{0,\Omega} \leq C_P \|\underline{\mathbf{v}}_h\|_{\epsilon,h} \quad \forall \underline{\mathbf{v}}_h \in \underline{\mathbf{U}}_{h,0}^k. \quad (30)$$

*Proof.* We refer to the proof of Lemma 2.15 in [20]  $\square$

An immediate consequence of Lemma 5.3, is the boundedness of linear functional  $b_h$ . This allows us to ensure the unique solvability of the nonlinear HHO formulation (17).

**Theorem 5.1** *There is one and only one solution  $\underline{\mathbf{u}}_h \in \underline{\mathbf{U}}_{h,0}^k$  of (17), which satisfies*

$$\|\underline{\mathbf{u}}_h\|_{\epsilon,h} \leq \frac{1}{C_{SM}} \left( C_P \|f\|_{0,\Omega} + \|A_h(0, \cdot)\|_{\underline{\mathbf{U}}_{h,0}^{k,*}} \right). \quad (31)$$

*Proof.* Thanks to Lemmas 5.1 and 5.2, the existence and uniqueness of the solution of (17) is consequence of a well known result in nonlinear functional analysis (see, e.g. Theorem 3.3.23 in Chapter III of [41, page 50], or Theorem 35.4 in [48, page 251]). On the other hand, (31) is obtained taking into account the strong monotonicity of  $A_h$ , (17), and Lemma 5.3. We omit further details.  $\square$

## 6 Error analysis

In this Section, we obtain a priori error estimates of the method, in the energy-norm and also in  $L^2$ -norm, under additional regularity assumption on exact solution. We emphasize that the analysis presented here is quite different to the one shown/described in [13], since we are dealing just with a kind of nonlinear version of the 2-Laplacian problem.

The following result will help us to conclude the error estimates of the method. From now on, we let  $u \in U$  be the exact solution of (3), and  $\underline{\mathbf{u}}_h \in \underline{\mathbf{U}}_{h,0}^k$  be the unique solution of (17) on the mesh  $\mathcal{T}_h$ . We define the *consistency error* as the linear functional  $\mathcal{E}_h(u; \cdot) : \underline{\mathbf{U}}_{h,0}^k \rightarrow \mathbb{R}$  such that  $\mathcal{E}_h(u; \cdot) := A_h(\underline{\mathbf{I}}_h^k u, \cdot) - b_h(\cdot)$ .

For the sake of completeness, we introduce the following notations for the flux and potential fields, and their approximations. First, we recall that  $\boldsymbol{\sigma} := \mathbf{a}(\cdot, \nabla u)$  and  $\widehat{\underline{\mathbf{u}}}_h := \underline{\mathbf{I}}_h^k u$ . Next, for each  $T \in \mathcal{T}_h$ , we introduce

$$\boldsymbol{\sigma}_T := \mathbf{a}(\cdot, G_T^k \underline{\mathbf{u}}_T), \quad \widehat{\underline{\mathbf{u}}}_T := \underline{\mathbf{I}}_T^k u \quad \text{and} \quad \widehat{\boldsymbol{\sigma}}_T := \mathbf{a}(\cdot, G_T^k \widehat{\underline{\mathbf{u}}}_T).$$

**Lemma 6.1 (Consistency error estimate).** *Let  $u \in H_0^1(\Omega)$  be the unique solution of (3). Let  $(\mathcal{T}_h)_{h \in \mathcal{H}}$  be a regular mesh family. Let a polynomial degree  $k \geq 0$  be fixed, and for all  $h \in \mathcal{H}$ , we take the additional regularity  $u \in H^{r+2}(\mathcal{T}_h)$  and  $\mathbf{a}(\cdot, \nabla u) \in H^{r+1}(\mathcal{T}_h)$ , for some  $r \in \{0, \dots, k\}$ . Then, there holds*

$$\sup_{\mathbf{v}_h \in \underline{\mathbf{U}}_{h,0}^k, \|\mathbf{v}_h\|_{\epsilon,h}=1} \mathcal{E}_h(u; \mathbf{v}_h) \lesssim h^{r+1} \left( \|u\|_{H^{r+2}(\mathcal{T}_h)} + \|\mathbf{a}(\cdot, \nabla u)\|_{H^{r+1}(\mathcal{T}_h)^d} \right). \quad (32)$$

*Proof.* We adapt the proof of Theorem 16 in [6]. Using the definition (17), fixing  $\mathbf{v}_h \in \underline{\mathbf{U}}_{h,0}^k$ , we obtain

$$\begin{aligned} \mathcal{E}_h(u; \mathbf{v}_h) &= \sum_{T \in \mathcal{T}_h} (\widehat{\boldsymbol{\sigma}}_T, G_T^k \mathbf{v}_T)_T + s_h(\widehat{\mathbf{u}}_h, \mathbf{v}_h) - (f, v_h)_{0,\Omega} \\ &= \sum_{T \in \mathcal{T}_h} (\widehat{\boldsymbol{\sigma}}_T - \boldsymbol{\sigma}, G_T^k \mathbf{v}_T)_T + s_h(\widehat{\mathbf{u}}_h, \mathbf{v}_h) \\ &\quad + \sum_{T \in \mathcal{T}_h} (\boldsymbol{\sigma}, G_T^k \mathbf{v}_T)_T - (f, v_h)_{0,\Omega}. \end{aligned} \quad (33)$$

Considering, for any  $T \in \mathcal{T}_h$ , the definition (10) of  $G_T^k$  with  $\boldsymbol{\phi} := \pi_T^k \boldsymbol{\sigma} \in [\mathbb{P}_d^k(T)]^d$ , we have that

$$\sum_{T \in \mathcal{T}_h} (G_T^k \mathbf{v}_T, \boldsymbol{\sigma})_T = \sum_{T \in \mathcal{T}_h} \left\{ (\boldsymbol{\sigma}, \nabla v_T)_T + \sum_{F \in \mathcal{F}_T} (\pi_T^k \boldsymbol{\sigma} \cdot \mathbf{n}_{TF}, v_F - v_T)_F \right\}, \quad (34)$$

where we have used the fact that  $\nabla v_T \in [\mathbb{P}_d^{k-1}(T)]^d$ , together with the definition of the  $L^2$ -orthogonal projector  $\pi_T^k$  componentwise.

On the other hand, knowing that  $v_h|_T = v_T$  for all  $T \in \mathcal{T}_h$ , and  $f = -\nabla \cdot \boldsymbol{\sigma}$  a.e. in  $\Omega$ , and after integrating by parts element by element, we get that

$$(f, v_h)_{0,\Omega} = \sum_{T \in \mathcal{T}_h} \left\{ (\boldsymbol{\sigma}, \nabla v_T)_T + \sum_{F \in \mathcal{F}_T} (\boldsymbol{\sigma} \cdot \mathbf{n}_{TF}, v_F - v_T)_F \right\}, \quad (35)$$

where we have additionally taken into account that  $\boldsymbol{\sigma} \in H(\text{div}; \Omega)$ ,  $(v_F)_{F \in \mathcal{F}_h}$  is single valued, and  $v_F = 0 \quad \forall F \in \mathcal{F}_h^b$ .

Now, replacing (34) and (35) in (33), we deduce

$$\begin{aligned} \mathcal{E}_h(u; \mathbf{v}_h) &= \sum_{T \in \mathcal{T}_h} (\mathbf{a}(\cdot, G_T^k \widehat{\mathbf{u}}_T) - \mathbf{a}(\cdot, \nabla u), G_T^k \mathbf{v}_T)_T + s_h(\widehat{\mathbf{u}}_h, \mathbf{v}_h) \\ &\quad + \sum_{T \in \mathcal{T}_h} \sum_{F \in \mathcal{F}_T} \left( \mathbf{a}(\cdot, \nabla u) - \pi_T^k \mathbf{a}(\cdot, \nabla u) \cdot \mathbf{n}_{TF}, v_F - v_T \right)_F. \end{aligned} \quad (36)$$

Then, after applying Cauchy-Schwarz and Minkowski inequalities, we obtain

$$|\mathcal{E}_h(u; \mathbf{v}_h)| \lesssim \mathfrak{T}^{1/2} \|\mathbf{v}_h\|_{\epsilon,h}, \quad (37)$$

where

$$\mathfrak{T} := \underbrace{\sum_{T \in \mathcal{T}_h} \|\mathbf{a}(\cdot, G_T^k \hat{\mathbf{u}}_T) - \mathbf{a}(\cdot, \nabla u)\|_T^2}_{\mathfrak{T}_1} + \underbrace{\sum_{T \in \mathcal{T}_h} \sum_{F \in \mathcal{F}_T} h_F \|\mathbf{a}(\cdot, \nabla u) - \pi_T^k \mathbf{a}(\cdot, \nabla u)\|_F^2}_{\mathfrak{T}_2} + \underbrace{s_h(\hat{\mathbf{u}}_h, \hat{\mathbf{u}}_h)}_{\mathfrak{T}_3}. \quad (38)$$

To bound  $\mathfrak{T}_1$ , we apply the Lipschitz continuity (27) with  $\boldsymbol{\theta} := \nabla u$ ,  $\tilde{\boldsymbol{\theta}} := G_h^k \hat{\mathbf{u}}_h$ , and invoke the optimal approximation properties of  $G_T^k \mathbf{I}_T^k$  (cf. (11)) together with (5), with  $m = 1$  and  $s = r + 2$ , to deduce

$$\mathfrak{T}_1 \lesssim \sum_{T \in \mathcal{T}_h} \|\nabla u - G_T^k \hat{\mathbf{u}}_T\|_T^2 \lesssim \sum_{T \in \mathcal{T}_h} h_T^{2(r+1)} \|u\|_{H^{r+2}(T)}^2. \quad (39)$$

Now, thanks to (11) again, we derive

$$\mathfrak{T}_2 \lesssim \sum_{T \in \mathcal{T}_h} h_T^{2(r+1)} \|\mathbf{a}(\cdot, \nabla u)\|_{[H^{r+1}(T)]^d}^2. \quad (40)$$

Finally, invoking Lemma 4.3, we bound  $\mathfrak{T}_3$ , and (32) is straightforwardly implied. We end the proof.  $\square$

**Theorem 6.1 (Energy error estimate).** *Under the same assumptions and notations given in Lemma 6.1, there holds, for some  $r \in \{0, \dots, k\}$ :*

$$\|\underline{\mathbf{u}}_h - \mathbf{I}_h^k u\|_{\epsilon, h} \lesssim h^{r+1} \left( \|u\|_{H^{r+2}(\mathcal{T}_h)} + \|\mathbf{a}(\cdot, \nabla u)\|_{[H^{r+1}(\mathcal{T}_h)]^d} \right). \quad (41)$$

Moreover, applying Lemma 4.1, there holds

$$\|\nabla u - G_h^k \underline{\mathbf{u}}_h\|_{0, \Omega} + s_h(\underline{\mathbf{u}}_h, \underline{\mathbf{u}}_h)^{1/2} \lesssim h^{r+1} \left( \|u\|_{H^{r+2}(\mathcal{T}_h)} + \|\mathbf{a}(\cdot, \nabla u)\|_{[H^{r+1}(\mathcal{T}_h)]^d} \right). \quad (42)$$

*Proof.* First, we take into account the strong monotonicity of nonlinear form  $A_h$ , and obtain

$$\begin{aligned} \|\underline{\mathbf{u}}_h - \mathbf{I}_h^k u\|_{\epsilon, h}^2 &\lesssim A_h(\underline{\mathbf{u}}_h, \underline{\mathbf{u}}_h - \mathbf{I}_h^k u) - A_h(\mathbf{I}_h^k u, \underline{\mathbf{u}}_h - \mathbf{I}_h^k u) \\ &= b_h(\underline{\mathbf{u}}_h - \mathbf{I}_h^k u) - A_h(\mathbf{I}_h^k u, \underline{\mathbf{u}}_h - \mathbf{I}_h^k u) \\ &= \mathcal{E}_h(u; \mathbf{I}_h^k u - \underline{\mathbf{u}}_h) \\ &\lesssim \sup_{\underline{\mathbf{v}}_h \in \mathbf{U}_{h,0}^k, \|\underline{\mathbf{v}}_h\|_{\epsilon, h}=1} \mathcal{E}_h(u; \underline{\mathbf{v}}_h) \|\underline{\mathbf{u}}_h - \mathbf{I}_h^k u\|_{\epsilon, h}, \end{aligned}$$

and then, we deduce

$$\|\underline{\mathbf{u}}_h - \mathbf{I}_h^k u\|_{\epsilon, h} \lesssim \sup_{\underline{\mathbf{v}}_h \in \mathbf{U}_{h,0}^k, \|\underline{\mathbf{v}}_h\|_{\epsilon, h}=1} \mathcal{E}_h(u; \underline{\mathbf{v}}_h). \quad (43)$$

Thus, (41) follows once we apply Lemma 6.1. The proof of (42) relies on Lemma 4.1. We omit further details.  $\square$

Let us turn our attention to the  $L^2$ -norm for the volumetric part of the error  $\underline{\mathbf{u}}_h - \underline{\mathbf{I}}_h^k u$ , that is  $e_h := \pi_h^k u - u_h$ , where  $\pi_h^k u|_T := \pi_T^k u$  and  $u_h|_T := u_T$ . To this aim, we rewrite problem (17) as: Find  $\underline{\mathbf{u}}_h \in V_h := \underline{\mathbf{U}}_{h,0}^k$  such that, for any  $\underline{\mathbf{v}}_h \in V_h$ ,

$$[\mathbb{A}_h(\underline{\mathbf{u}}_h), \underline{\mathbf{v}}_h] := [\mathbb{N}_h(\underline{\mathbf{u}}_h), \underline{\mathbf{v}}_h] + s_h(\underline{\mathbf{u}}_h, \underline{\mathbf{v}}_h) = (f, v_h)_{0,\Omega} =: [F_h, \underline{\mathbf{v}}_h], \quad (44)$$

where  $[\cdot, \cdot]$  represents the duality product,  $V(h) := \underline{\mathbf{U}}_h^k$ ,  $\mathbb{A}_h : V(h) \rightarrow V(h)'$  and  $\mathbb{N}_h : V(h) \rightarrow V(h)'$  are the nonlinear operators induced by  $A_h$  and  $N_h$ , respectively.

Now, we can ensure the existence of the Gâteaux derivative of  $\mathbb{N}_h$  at each  $\underline{\mathbf{z}}_h \in V(h)$ , thanks to the assumption **(H.I)**, as the bounded bilinear form  $D\mathbb{N}_h(\underline{\mathbf{z}}_h) : V(h) \times V(h) \rightarrow \mathbb{R}$ , given by

$$\begin{aligned} D\mathbb{N}_h(\underline{\mathbf{z}}_h)(\underline{\mathbf{v}}_h, \underline{\mathbf{w}}_h) &= \lim_{t \rightarrow 0} \frac{[\mathbb{N}_h(\underline{\mathbf{z}}_h + t\underline{\mathbf{v}}_h), \underline{\mathbf{w}}_h] - [\mathbb{N}_h(\underline{\mathbf{z}}_h), \underline{\mathbf{w}}_h]}{t} \\ &= \lim_{t \rightarrow 0} \frac{\sum_{T \in \mathcal{T}_h} \left\{ (\mathbf{a}(\cdot, G_T^k \underline{\mathbf{z}}_T + tG_T^k \underline{\mathbf{v}}_T), G_T^k \underline{\mathbf{w}}_T)_T - (\mathbf{a}(\cdot, G_T^k \underline{\mathbf{z}}_T), G_T^k \underline{\mathbf{w}}_T)_T \right\}}{t} \\ &= \sum_{T \in \mathcal{T}_h} \left( \lim_{t \rightarrow 0} \frac{\mathbf{a}(\cdot, G_T^k \underline{\mathbf{z}}_T + tG_T^k \underline{\mathbf{v}}_T) - \mathbf{a}(\cdot, G_T^k \underline{\mathbf{z}}_T)}{t}, G_T^k \underline{\mathbf{w}}_T \right)_T \\ &= \sum_{T \in \mathcal{T}_h} \int_T D\mathbf{a}(\cdot, G_T^k \underline{\mathbf{z}}_T) G_T^k \underline{\mathbf{v}}_T \cdot G_T^k \underline{\mathbf{w}}_T \quad \forall \underline{\mathbf{v}}_h, \underline{\mathbf{w}}_h \in V(h). \end{aligned} \quad (45)$$

As consequence, we define  $D\mathbb{A}_h(\underline{\mathbf{z}}_h) : V(h) \times V(h) \rightarrow \mathbb{R}$  as the bilinear form

$$D\mathbb{A}_h(\underline{\mathbf{z}}_h)(\underline{\mathbf{v}}_h, \underline{\mathbf{w}}_h) := D\mathbb{N}_h(\underline{\mathbf{z}}_h)(\underline{\mathbf{v}}_h, \underline{\mathbf{w}}_h) + s_h(\underline{\mathbf{v}}_h, \underline{\mathbf{w}}_h) \quad \forall \underline{\mathbf{v}}_h, \underline{\mathbf{w}}_h \in V(h). \quad (46)$$

In what follows, we assume that the jacobian tensor  $D\mathbf{a}(\cdot, \boldsymbol{\theta})$  is symmetric for all  $\boldsymbol{\theta} \in [L^2(\Omega)]^d$ , and that  $D\mathbb{N}_h$  is *hemi-continuous*, that is for any  $\underline{\mathbf{v}}_h, \underline{\mathbf{w}}_h \in V(h)$ , the mapping  $\mathbb{R} \ni \mu \rightarrow D\mathbb{N}_h(\underline{\mathbf{v}}_h + \mu \underline{\mathbf{w}}_h)(\underline{\mathbf{w}}_h, \cdot) \in V(h)'$  is continuous. Thus, applying the mean value theorem, there exists  $\tilde{\underline{\mathbf{u}}}_h \in V(h)$ , a convex combination of  $\underline{\mathbf{I}}_h^k u$  and  $\underline{\mathbf{u}}_h$ , such that

$$D\mathbb{N}_h(\tilde{\underline{\mathbf{u}}}_h)(\underline{\mathbf{I}}_h^k u - \underline{\mathbf{u}}_h, \underline{\mathbf{v}}_h) = [\mathbb{N}_h(\underline{\mathbf{I}}_h^k u) - \mathbb{N}_h(\underline{\mathbf{u}}_h), \underline{\mathbf{v}}_h] \quad \forall \underline{\mathbf{v}}_h \in V(h). \quad (47)$$

Further, it follows from (46) and (47), that

$$D\mathbb{A}_h(\tilde{\underline{\mathbf{u}}}_h)(\underline{\mathbf{I}}_h^k u - \underline{\mathbf{u}}_h, \underline{\mathbf{v}}_h) = [\mathbb{A}_h(\underline{\mathbf{I}}_h^k u) - \mathbb{A}_h(\underline{\mathbf{u}}_h), \underline{\mathbf{v}}_h] \quad \forall \underline{\mathbf{v}}_h \in V(h). \quad (48)$$

Next, we establish the following linear auxiliary problem with additional elliptic regularity assumption: Given  $g \in L^2(\Omega)$ , we let  $z \in U := H_0^1(\Omega)$  be the unique solution of

$$-\operatorname{div}(\widetilde{\mathbf{K}} \nabla z) = g \quad \text{in } \Omega, \quad (49a)$$

$$z = 0 \quad \text{on } \partial\Omega, \quad (49b)$$

with  $\widetilde{\mathbf{K}} \in [L^\infty(\mathcal{T}_h)]^{d \times d}$ , such that  $\widetilde{\mathbf{K}}|_T := D\mathbf{a}(\cdot, G_T^k \underline{\mathbf{u}}_T)$ , for each  $T \in \mathcal{T}_h$ . Assuming further regularity on  $\widetilde{\mathbf{K}}$  so that  $z \in U \cap H^2(\Omega)$  and  $\widetilde{\mathbf{K}} \nabla z \in [H^1(\Omega)]^d$ , there exists  $C_{\text{ell}} > 0$  only depending on  $\Omega$ , such that:

$$\|z\|_{H^2(\Omega)} \leq C_{\text{ell}} \|g\|_{0,\Omega} \quad \text{and} \quad \|\widetilde{\mathbf{K}} \nabla z\|_{H^1(\Omega)^d} \leq C_{\text{ell}} \|g\|_{0,\Omega}. \quad (50)$$



This holds, for example, when  $\Omega$  is convex, and the tensor  $\widetilde{\mathbf{K}}$  is symmetric and Lipschitz continuous (cf. [36]). We also consider the HHO formulation of problem (49): *Find  $\underline{\mathbf{z}}_h \in V_h$  such that*

$$D\mathbb{A}_h(\underline{\mathbf{u}}_h)(\underline{\mathbf{z}}_h, \underline{\mathbf{v}}_h) = (g, v_h)_{0,\Omega} =: [\tilde{F}_h, \underline{\mathbf{v}}_h] \quad \forall \underline{\mathbf{v}}_h \in V_h. \quad (51)$$

We emphasize that invoking Theorem 4.16 in [20] with  $r = 0$ , and applying (50), we derive

$$\|\mathbb{I}_h^k \underline{\mathbf{z}} - \underline{\mathbf{z}}_h\|_{\epsilon,h} \lesssim h \|g\|_{0,\Omega}. \quad (52)$$

Further, we introduce another nonlinear operator  $\widehat{\mathbb{N}}_h : [L^2(\Omega)]^d \rightarrow [L^2(\Omega)]^d$  as

$$[\widehat{\mathbb{N}}_h(\boldsymbol{\theta}), \boldsymbol{\zeta}] := \sum_{T \in \mathcal{T}_h} (\mathbf{a}(\cdot, \boldsymbol{\theta}|_T), \boldsymbol{\zeta})_T \quad \forall \boldsymbol{\theta}, \boldsymbol{\zeta} \in [L^2(\Omega)]^d. \quad (53)$$

Then, the corresponding Gâteaux derivative of  $\widehat{\mathbb{N}}_h$  at each  $\boldsymbol{\theta} \in [L^2(\Omega)]^d$  is given by

$$D\widehat{\mathbb{N}}_h(\boldsymbol{\theta})(\boldsymbol{\eta}, \boldsymbol{\zeta}) := \sum_{T \in \mathcal{T}_h} (D\mathbf{a}(\cdot, \boldsymbol{\theta}|_T) \boldsymbol{\eta}, \boldsymbol{\zeta})_T, \quad \forall \boldsymbol{\eta}, \boldsymbol{\zeta} \in [L^2(\Omega)]^d. \quad (54)$$

Now, assuming that  $D\widehat{\mathbb{N}}_h$  is hemi-continuous, we ensure that there exists  $\widehat{\boldsymbol{\theta}} \in [L^2(\Omega)]^d$ , a convex combination of  $\nabla u$  and  $\pi_h^k \nabla u$ , such that

$$D\widehat{\mathbb{N}}_h(\widehat{\boldsymbol{\theta}})(\nabla u - \pi_h^k \nabla u, \boldsymbol{\zeta}) = [\widehat{\mathbb{N}}_h(\nabla u) - \widehat{\mathbb{N}}_h(\pi_h^k \nabla u), \boldsymbol{\zeta}] \quad \forall \boldsymbol{\zeta} \in [L^2(\Omega)]^d. \quad (55)$$

Then, introducing  $\widehat{\boldsymbol{\theta}}_T := \widehat{\boldsymbol{\theta}}|_T$ , we set  $\widehat{\mathbf{K}} \in [L^\infty(\mathcal{T}_h)]^{d \times d}$  such that  $\widehat{\mathbf{K}}|_T := D\mathbf{a}(\cdot, \widehat{\boldsymbol{\theta}}_T)$ , and we can derive the following representation of  $D\widehat{\mathbb{N}}_h(\widehat{\boldsymbol{\theta}})$  as the bilinear form

$$D\widehat{\mathbb{N}}_h(\widehat{\boldsymbol{\theta}})(\boldsymbol{\eta}, \boldsymbol{\zeta}) := \sum_{T \in \mathcal{T}_h} (\widehat{\mathbf{K}} \boldsymbol{\eta}, \boldsymbol{\zeta})_T, \quad \forall \boldsymbol{\eta}, \boldsymbol{\zeta} \in [L^2(\Omega)]^d. \quad (56)$$

**Theorem 6.2 (Error estimate of  $L^2$ -projection of the potential)** *Under the same hypothesis given in Theorem 6.1, and assuming that  $\widehat{\mathbf{K}} \in [W^{1,\infty}(\mathcal{T}_h)]^{d \times d}$ . Let a integer  $k \geq 0$  be fixed, and for all  $h \in \mathcal{H}$ , we take the additional regularity  $u \in H^{q+2}(\mathcal{T}_h)$  and  $\mathbf{a}(\cdot, \nabla u) \in H^{q+1}(\mathcal{T}_h)$ , for some  $q \in \{0, \dots, k\}$ . Then, for  $k \geq 1$ , there holds*

$$\|\pi_h^k u - u_h\|_{0,\Omega} \lesssim h^{q+2} \left( (1 + |\widehat{\mathbf{K}}|_{[W^{1,\infty}(\mathcal{T}_h)]^{d \times d}}) \|u\|_{H^{q+2}(\mathcal{T}_h)} + \|\mathbf{a}(\cdot, \nabla u)\|_{[H^{q+1}(\mathcal{T}_h)]^d} \right). \quad (57)$$

Moreover, when  $k = 0$ , assuming that  $f \in H^1(\mathcal{T}_h)$ , there holds

$$\|\pi_h^0 u - u_h\|_{0,\Omega} \lesssim h^2 \left( (1 + |\widehat{\mathbf{K}}|_{[W^{1,\infty}(\mathcal{T}_h)]^{d \times d}}) \|u\|_{H^2(\mathcal{T}_h)} + \|\mathbf{a}(\cdot, \nabla u)\|_{[H^1(\mathcal{T}_h)]^d} + \|f\|_{H^1(\mathcal{T}_h)} \right). \quad (58)$$

*Proof.* From problem (51), with  $g := \pi_h^k u - u_h \in L^2(\Omega)$  and  $\underline{\mathbf{v}}_h := \mathbb{I}_h^k u - \underline{\mathbf{u}}_h \in V_h$ , we derive

$$\|\pi_h^k u - u_h\|_{0,\Omega}^2 = [\tilde{F}_h, \mathbb{I}_h^k u - \underline{\mathbf{u}}_h] = D\mathbb{A}_h(\underline{\mathbf{u}}_h)(\underline{\mathbf{z}}_h, \mathbb{I}_h^k u - \underline{\mathbf{u}}_h).$$

Now, as  $D\mathbb{A}_h(\tilde{\mathbf{u}}_h)$  is symmetric, we apply (48) to deduce

$$\begin{aligned} \|\pi_h^k u - u_h\|_{0,\Omega}^2 &= [\mathbb{A}_h(\mathbf{I}_h^k u) - \mathbb{A}_h(\mathbf{u}_h), \mathbf{z}_h] \\ &= \underbrace{\left\{ [\mathbb{A}_h(\mathbf{I}_h^k u), \mathbf{z}_h - \mathbf{I}_h^k z] - [F_h, \mathbf{z}_h - \mathbf{I}_h^k z] \right\}}_{\mathfrak{T}_1} + \underbrace{\left\{ [\mathbb{A}_h(\mathbf{I}_h^k u), \mathbf{I}_h^k z] - [F_h, \mathbf{I}_h^k z] \right\}}_{\mathfrak{T}_2}. \end{aligned} \quad (59)$$

Since  $\mathfrak{T}_1$  in (59) corresponds to the consistency error of (44),  $\mathcal{E}_h(u; \mathbf{z}_h - \mathbf{I}_h^k z)$ , we invoke (32) with  $r = q$ , and after taking into account (52), we obtain

$$\begin{aligned} |\mathfrak{T}_1| &\lesssim h^{q+1} \left( \|u\|_{H^{q+2}(\mathcal{T}_h)} + \|\mathbf{a}(\cdot, \nabla u)\|_{[H^{q+1}(\mathcal{T}_h)]^d} \right) \|\mathbf{I}_h^k z - \mathbf{z}_h\|_{\epsilon, h} \\ &\lesssim h^{q+1} \left( \|u\|_{H^{q+2}(\mathcal{T}_h)} + \|\mathbf{a}(\cdot, \nabla u)\|_{[H^{q+1}(\mathcal{T}_h)]^d} \right) h \|\pi_h^k u - u_h\|_{0,\Omega}. \end{aligned} \quad (60)$$

On the other hand, we notice that  $\mathfrak{T}_2$  in (59) is the consistence term  $\mathcal{E}_h(u; \mathbf{I}_h^k z)$ . Here, we consider two cases.

(i.A)  $k \geq 1$ . We proceed as in the derivation of (36), to have

$$\begin{aligned} \mathcal{E}_h(u; \mathbf{I}_h^k z) &= \underbrace{\sum_{T \in \mathcal{T}_h} (\mathbf{a}(\cdot, \pi_T^k \nabla u) - \mathbf{a}(\cdot, \nabla u), \pi_T^k \nabla z)_T}_{\mathfrak{J}_1} \\ &\quad + \underbrace{\sum_{T \in \mathcal{T}_h} \sum_{F \in \mathcal{F}_T} ([\mathbf{a}(\cdot, \nabla u) - \pi_T^k \mathbf{a}(\cdot, \nabla u)] \cdot \mathbf{n}_{TF}, \pi_F^k z - \pi_T^k z)_F}_{\mathfrak{J}_2} \\ &\quad - \underbrace{s_h(\mathbf{I}_h^k u, \mathbf{I}_h^k z)}_{\mathfrak{J}_3}. \end{aligned} \quad (61)$$

We bound  $\mathfrak{J}_2$  applying Cauchy-Schwarz, resulting

$$|\mathfrak{J}_2| \leq \left( \sum_{T \in \mathcal{T}_h} \sum_{F \in \mathcal{F}_T} \|\mathbf{a}(\cdot, \nabla u) - \pi_T^k \mathbf{a}(\cdot, \nabla u)\|_F^2 \right)^{1/2} \left( \sum_{T \in \mathcal{T}_h} \sum_{F \in \mathcal{F}_T} \|\pi_F^k z - \pi_T^k z\|_F^2 \right)^{1/2},$$

By approximation property (6), with  $m = 0$  and  $s := q + 1$ , we obtain

$$\|\mathbf{a}(\cdot, \nabla u) - \pi_T^k \mathbf{a}(\cdot, \nabla u)\|_F \lesssim h_T^{q+1/2} \|\mathbf{a}(\cdot, \nabla u)\|_{H^{q+1}(T)^d}. \quad (62)$$

Similarly, applying approximation property (6), we infer

$$\|\pi_F^k z - \pi_T^k z\|_F \leq \|z - \pi_T^k z\|_F \lesssim h_T^{3/2} \|z\|_{H^2(T)}. \quad (63)$$

Then, from (62), (63), and (50), with  $g := \pi_h^k u - u_h$ , we derive

$$|\mathfrak{J}_2| \lesssim h^{q+2} \|\mathbf{a}(\cdot, \nabla u)\|_{H^{q+1}(\mathcal{T}_h)} \|\pi_h^k u - u_h\|_{0,\Omega}. \quad (64)$$

Now, we bound  $\mathfrak{J}_3$ , using (25) with  $r := q$  for  $u$ , and  $r := 0$  for  $z$ , and the first elliptic regularity estimate in (50), yielding to

$$\begin{aligned} |\mathfrak{J}_3| &\leq s_h(\mathbf{I}_h^k u, \mathbf{I}_h^k u)^{1/2} s_h(\mathbf{I}_h^k z, \mathbf{I}_h^k z)^{1/2} \\ &\lesssim h^{q+1} \|u\|_{H^{q+2}(\mathcal{T}_h)} h \|z\|_{H^2(\Omega)} \\ &\lesssim h^{q+2} \|u\|_{H^{q+2}(\mathcal{T}_h)} \|\pi_h^k u - u_h\|_{0,\Omega}. \end{aligned} \quad (65)$$

Finally, to bound  $\mathfrak{J}_1$ , we invoke the bilinear form (56), obtaining

$$\begin{aligned} \mathfrak{J}_1 &= D\hat{\mathbf{N}}_h(\hat{\boldsymbol{\theta}})(\nabla u - \pi_h^k \nabla u, \pi_h^k \nabla z) \\ &= \sum_{T \in \mathcal{T}_h} (\widehat{\mathbf{K}}(\nabla u - \pi_T^k \nabla u), \pi_T^k \nabla z)_T \\ &= \underbrace{\sum_{T \in \mathcal{T}_h} (\widehat{\mathbf{K}}(\nabla u - \pi_T^k \nabla u), \pi_T^k \nabla z - \pi_T^0 \nabla z)_T}_{\mathfrak{J}_{1,1}} \\ &\quad + \underbrace{\sum_{T \in \mathcal{T}_h} (\nabla u - \pi_T^k \nabla u, \widehat{\mathbf{K}} \pi_T^0 \nabla z - \pi_T^0 (\widehat{\mathbf{K}} \pi_T^0 \nabla z))_T}_{\mathfrak{J}_{1,2}}. \end{aligned}$$

Invoking **(H.4)**, the fact that  $\|\pi_T^k \nabla z - \pi_T^0 \nabla z\|_T \leq \|\nabla z - \pi_T^0 \nabla z\|_T$ , and that  $\widehat{\mathbf{K}}$  is uniformly bounded, we apply approximation property (5) and the first elliptic regularity estimate in (50), and deduce

$$|\mathfrak{J}_{1,1}| \lesssim h^{q+2} \|u\|_{H^{q+2}(\mathcal{T}_h)} \|\pi_h^k u - u_h\|_{0,\Omega}. \quad (66)$$

In order to bound  $\mathfrak{J}_{1,2}$ , we notice that  $\pi_T^0 (\widehat{\mathbf{K}} \pi_T^0 \nabla z) = (\pi_T^0 \widehat{\mathbf{K}})(\pi_T^0 \nabla z)$ , where  $\pi_T^0 \widehat{\mathbf{K}}$  is defined componentwise. Then, since  $\widehat{\mathbf{K}} \in [W^{1,\infty}(\mathcal{T}_h)]^{d \times d}$ , we invoke the *tensorial version of* Theorem 1.45 in [20], and establish that  $\|\widehat{\mathbf{K}} - \pi_T^0 \widehat{\mathbf{K}}\|_{[W^{0,\infty}(T)]^{d \times d}} \lesssim h_T |\widehat{\mathbf{K}}|_{[W^{1,\infty}(\mathcal{T}_h)]^{d \times d}}$ . Now, after applying again approximation properties (5) as in (66), it yields

$$|\mathfrak{J}_{1,2}| \lesssim |\widehat{\mathbf{K}}|_{[W^{1,\infty}(\mathcal{T}_h)]^{d \times d}} h^{q+2} \|u\|_{H^{q+2}(\mathcal{T}_h)} \|\pi_h^k u - u_h\|_{0,\Omega}. \quad (67)$$

Then, from (64), (65), (66), and (67), we deduce the estimate for  $\mathcal{E}_h(u; \mathbf{I}_h^k z)$  in (61), and (57) is concluded.

(i.B)  $\mathbf{k} = \mathbf{0}$ . We can rewrite

$$\begin{aligned} \mathcal{E}_h(u; \mathbf{I}_h^0 z) &= [\mathbb{A}(\mathbf{I}_h^0 u) - \mathbb{A}(\mathbf{u}_h), \mathbf{I}_h^0 z] \\ &= \sum_{T \in \mathcal{T}_h} (\mathbf{a}(\cdot, \pi_T^0 \nabla u), \pi_T^0 \nabla z)_T + s_h(\mathbf{I}_h^0 u, \mathbf{I}_h^0 z) - \sum_{T \in \mathcal{T}_h} (f, \pi_T^0 z)_T. \end{aligned} \quad (68)$$

Playing with the definition of  $\pi_T^0$ , we obtain

$$(f, \pi_T^0 z)_T = (\pi_T^0 f, z)_T = (\pi_T^0 f - f, z - \pi_T^0 z) + (f, z)_T. \quad (69)$$

Then, from (68) and (69), and the fact  $(\pi_T^0 \nabla z - \nabla z, \pi_T^0 \mathbf{a}(\cdot, \nabla u))_T = 0$  for all  $T \in \mathcal{T}_T$ , we deduce

$$\begin{aligned}
\mathcal{E}_h(u; \mathbf{I}_h^0 z) &= \sum_{T \in \mathcal{T}_h} (\mathbf{a}(\cdot, \pi_T^0 \nabla u), \pi_T^0 \nabla z)_T - (f, z)_\Omega \\
&\quad + s_h(\mathbf{I}_h^0 u, \mathbf{I}_h^0 z) + \sum_{T \in \mathcal{T}_h} (\pi_T^0 f - f, \pi_T^0 z - z)_T \\
&= \underbrace{\sum_{T \in \mathcal{T}_h} (\mathbf{a}(\cdot, \pi_T^0 \nabla u) - \mathbf{a}(\cdot, \nabla u), \pi_T^0 \nabla z)_T}_{\mathfrak{T}_1} \\
&\quad + \underbrace{\sum_{T \in \mathcal{T}_h} (\mathbf{a}(\cdot, \nabla u) - \pi_T^0 \mathbf{a}(\cdot, \nabla u), \pi_T^0 \nabla z - \nabla z)_T}_{\mathfrak{T}_2} \\
&\quad + \underbrace{s_h(\mathbf{I}_h^0 u, \mathbf{I}_h^0 z)}_{\mathfrak{T}_3} + \underbrace{\sum_{T \in \mathcal{T}_h} (\pi_T^0 f - f, \pi_T^0 z - z)_T}_{\mathfrak{T}_4}. \tag{70}
\end{aligned}$$

$\mathfrak{T}_2$ ,  $\mathfrak{T}_3$  and  $\mathfrak{T}_4$  are bounded by applying Cauchy-Schwarz inequality and very well-known approximation property (5). Then, after take into account the first elliptic regularity estimate in (50), we find

$$|\mathfrak{T}_2| \lesssim h^2 \|\mathbf{a}(\cdot, \nabla u)\|_{[H^1(\mathcal{T}_h)]^d} \|\pi_h^0 u - u_h\|_{0,\Omega} \tag{71}$$

$$|\mathfrak{T}_3| \lesssim h^2 \|u\|_{H^2(\mathcal{T}_h)} \|\pi_h^0 u - u_h\|_{0,\Omega} \tag{72}$$

$$|\mathfrak{T}_4| \lesssim h^2 \|f\|_{H^1(\mathcal{T}_h)} \|\pi_h^0 u - u_h\|_{0,\Omega}. \tag{73}$$

Now, we aim to bound the remaining term  $\mathfrak{T}_1$ . First, we notice that

$$\begin{aligned}
\mathfrak{T}_1 &= D\widehat{\mathbf{N}}_h(\widehat{\boldsymbol{\theta}})(\nabla u - \pi_h^0 \nabla u, \pi_h^0 \nabla z) \\
&= \sum_{T \in \mathcal{T}_h} (\widehat{\mathbf{K}}(\nabla u - \pi_T^0 \nabla u), \pi_T^0 \nabla z)_T \\
&= \sum_{T \in \mathcal{T}_h} (\nabla u - \pi_T^0 \nabla u, \widehat{\mathbf{K}} \pi_T^0 \nabla z - \pi_T^0 (\widehat{\mathbf{K}} \pi_T^0 \nabla z))_T.
\end{aligned}$$

Then, proceeding in analogous way for bounding  $\mathfrak{J}_{1,2}$ , we derive

$$|\mathfrak{T}_1| \lesssim |\widehat{\mathbf{K}}|_{[W^{1,\infty}(\mathcal{T}_h)]^{d \times d}} h^2 \|u\|_{H^2(\mathcal{T}_h)} \|\pi_h^k u - u_h\|_{0,\Omega}. \tag{74}$$

As a result, we establish

$$|\mathcal{E}_h(u; \mathbf{I}_h^0 z)| \lesssim h^2 \left( (1 + |\widehat{\mathbf{K}}|_{[W^{1,\infty}(\mathcal{T}_h)]^{d \times d}}) \|u\|_{H^2(\mathcal{T}_h)} + \|\mathbf{a}(\cdot, \nabla u)\|_{[H^1(\mathcal{T}_h)]^d} + \|f\|_{H^1(\mathcal{T}_h)} \right), \tag{75}$$

and the estimate (58) is implied.  $\square$

Now, we establish another important result.

**Theorem 6.3** ( $L^2$ -error estimate of reconstructive potential). *Under the same assumptions of Theorem 6.2, elliptic regularity (50), we have, for  $k \geq 1$ :*

$$\|p_h^{k+1}\underline{\mathbf{u}}_h - u\|_{0,\Omega} \lesssim h^{q+2} \left( (1 + |\widehat{\mathbf{K}}|_{[W^{1,\infty}(\mathcal{T}_h)]^{d \times d}}) \|u\|_{H^{q+2}(\mathcal{T}_h)} + \|\mathbf{a}(\cdot, \nabla u)\|_{H^{q+1}(\mathcal{T}_h)^d} \right), \quad (76)$$

for some  $q \in \{0, \dots, k\}$ . Moreover, when  $k = 0$ , assuming that  $f \in H^1(\mathcal{T}_h)$ , there holds

$$\|p_h^1 u - u_h\|_{0,\Omega} \lesssim h^2 \left( (1 + |\widehat{\mathbf{K}}|_{[W^{1,\infty}(\mathcal{T}_h)]^{d \times d}}) \|u\|_{H^2(\mathcal{T}_h)} + \|\mathbf{a}(\cdot, \nabla u)\|_{H^1(\mathcal{T}_h)^d} + \|f\|_{H^1(\mathcal{T}_h)} \right). \quad (77)$$

*Proof.* We proceed as in the proof of Theorem 2.32 from [20]. First, we take  $\tilde{u}_h := p_h^{k+1}\underline{\mathbf{I}}_h^k(u)$ , and applying triangle inequality, we have

$$\|p_h^{k+1}\underline{\mathbf{u}}_h - u\|_{0,\Omega} \leq \underbrace{\|u - \tilde{u}_h\|_{0,\Omega}}_{\mathfrak{T}_1} + \underbrace{\|p_h^{k+1}(\widehat{\underline{\mathbf{u}}}_h - \underline{\mathbf{u}}_h)\|_{0,\Omega}}_{\mathfrak{T}_2}. \quad (78)$$

Next, after invoking (15) with  $s := q$ , we derive

$$\mathfrak{T}_1 \lesssim h^{q+2} \|u\|_{H^{q+2}(\mathcal{T}_h)}. \quad (79)$$

Now, using the local Poincaré-Wirtinger (4) and the fact

$$\left( p_T^{k+1}(\widehat{\underline{\mathbf{u}}}_T - \underline{\mathbf{u}}_T) - \pi_T^0(\widehat{u}_T - u_T), 1 \right)_T = 0 \quad \forall T \in \mathcal{T}_h,$$

we obtain, after applying triangle inequality

$$\begin{aligned} \mathfrak{T}_2^2 &\lesssim \sum_{T \in \mathcal{T}_h} \left\{ \|p_h^{k+1}(\widehat{\underline{\mathbf{u}}}_h - \underline{\mathbf{u}}_h) - \pi_T^0(\widehat{u}_T - u_T)\|_T^2 + \|\pi_T^0(\widehat{u}_T - u_T)\|_T^2 \right\} \\ &\lesssim \underbrace{\sum_{T \in \mathcal{T}_h} h_T^2 \|\nabla p_T^{k+1}(\widehat{\underline{\mathbf{u}}}_T - \underline{\mathbf{u}}_T)\|_T^2}_{\mathfrak{E}_1} + \underbrace{\sum_{T \in \mathcal{T}_h} \|\pi_T^0(\widehat{u}_T - u_T)\|_T^2}_{\mathfrak{E}_2}. \end{aligned} \quad (80)$$

We observe that,  $\mathfrak{E}_1$  is bounded thanks to Lemma 4.1 and Theorem 6.1, while the boundedness of  $\mathfrak{E}_2$  is deduced by invoking Theorem 6.2. Finally, (79) and (80) help us to bound (78), and we conclude the proof.  $\square$

## 7 Other boundary conditions

Here, we briefly discuss how the HHO scheme is written for non-homogeneous Dirichlet, mixed, and nonhomogeneous Neumann boundary conditions, and hint at the modifications required for establishing the corresponding a priori error estimates.

## 7.1 Nonhomogeneous Dirichlet boundary conditions

We consider the nonlinear problem

$$-\operatorname{div}(\mathbf{a}(\cdot, \nabla u)) = f \quad \text{in } \Omega, \quad u = g \quad \text{on } \partial\Omega, \quad (81)$$

with  $g \in H^{1/2}(\partial\Omega)$ . Denoting by  $\gamma : H^1(\Omega) \rightarrow H^{1/2}(\partial\Omega)$  the trace operator, the continuous weak formulation is expressed as: *Find*  $u \in H_g^1(\Omega) := \{w \in H^1(\Omega) : \gamma(w) = g\}$  *such that*

$$(\mathbf{a}(\cdot, \nabla u), \nabla v)_\Omega = (f, v)_\Omega \quad \forall v \in H_0^1(\Omega). \quad (82)$$

Now, if we decompose  $u = u_0 + u_g$ , such that  $u_0 \in H_0^1(\Omega)$  and  $u_g \in H_g^1(\Omega)$ , then we can rewrite (82) as: *Find*  $u_0 \in H_0^1(\Omega)$  *such that*

$$(\mathbf{a}(\cdot, \nabla u_0 + \nabla u_g), \nabla v)_\Omega = (f, v)_\Omega \quad \forall v \in H_0^1(\Omega). \quad (83)$$

The problem (83) is well-posed, since its associated nonlinear form results to be also Lipschitz continuous and strongly monotone in  $H_0^1(\Omega)$ . Now, to establish the HHO discrete scheme corresponding to (83), we follow the ideas given in Subsection 7.1 in [13], and introduce

$\underline{\mathbf{w}}_{g,h} := \left( (w_{g,T})_{T \in \mathcal{T}_h}, (w_{g,F})_{F \in \mathcal{F}_h} \right) \in \underline{\mathbf{U}}_h^k$  such that

$$w_{g,T} = 0 \quad \forall T \in \mathcal{T}_h, \quad w_{g,F} = 0 \quad \forall F \in \mathcal{F}_h^{\text{int}}, \quad w_{g,F} = \pi_F^k g \quad \forall F \in \mathcal{F}_h^{\text{b}}.$$

In order to show a procedure to compute  $\underline{\mathbf{u}}_h$ , the HHO-approximation of the solution of (83), we decompose it as  $\underline{\mathbf{u}}_h := \underline{\mathbf{u}}_{0,h} + \underline{\mathbf{w}}_{g,h}$ , with  $\underline{\mathbf{u}}_{0,h} \in \underline{\mathbf{U}}_{h,0}^k$ . As a result, the proposed HHO formulation reads as: *Find*  $\underline{\mathbf{u}}_{0,h} \in \underline{\mathbf{U}}_{h,0}^k$  *such that*

$$A_h(\underline{\mathbf{u}}_{0,h} + \underline{\mathbf{w}}_{g,h}, \underline{\mathbf{v}}_h) = b_h(\underline{\mathbf{v}}_h) \quad \forall \underline{\mathbf{v}}_h \in \underline{\mathbf{U}}_{h,0}^k, \quad (84)$$

with  $A_h$  and  $b_h$  defined as in (17). The well-posedness of (84) also relies on Lipschitz continuity and strong monotonicity properties. It is not difficult to establish similar results to Theorems 6.1, 6.2 and 6.3 in this situation.

## 7.2 Mixed boundary conditions

We consider a nonlinear problem (2) with mixed boundary conditions, for which we assume that there exists a partition  $\{\Gamma_D, \Gamma_N\}$  of the boundary  $\Gamma := \partial\Omega$ , such that  $\Gamma = \Gamma_N \cup \Gamma_D$ ,  $\overset{\circ}{\Gamma}_N \cap \overset{\circ}{\Gamma}_D = \emptyset$  and  $|\Gamma_D| > 0$ . Then, the nonlinear problem with mixed boundary conditions reads as: *Find*  $u : \Omega \rightarrow \mathbb{R}$  *such that*

$$-\operatorname{div}(\mathbf{a}(\cdot, \nabla u)) = f \quad \text{in } \Omega, \quad (85a)$$

$$\mathbf{a}(\cdot, \nabla u) \cdot \mathbf{n} = g_N \quad \text{on } \Gamma_N, \quad (85b)$$

$$u = g_D \quad \text{on } \Gamma_D, \quad (85c)$$

where  $f \in L^2(\Omega)$ ,  $g_N \in L^2(\Gamma_N)$ ,  $g_D \in H^{1/2}(\Gamma_D)$ , and  $\mathbf{n}$  denotes the exterior unit normal to  $\Gamma$ . We notice that (85) does not degenerate into the pure Neumann case. The continuous weak formulation of (85) reads as: *Find*  $u \in H_*^1(\Omega) := \{w \in H^1(\Omega) : w|_{\Gamma_D} = g_D\}$  *such that*

$$(\mathbf{a}(\cdot, \nabla u), \nabla v)_\Omega = (f, v)_\Omega + (g_N, v)_{\Gamma_N} \quad \forall v \in H_D^1(\Omega) := \{w \in H^1(\Omega) : w|_{\Gamma_D} = 0\}. \quad (86)$$

The well-posedness of (86) follows from [27]. Next, for the HHO discretization, we require the following hypothesis.

**Assumption 7.1**  $\mathcal{T}_h$  fits the partition  $\{\Gamma_D, \Gamma_N\}$ , in the sense that we can define two sets,  $\mathcal{F}_h^D := \{F \in \mathcal{F}_h^b : F \subset \Gamma_D\}$  and  $\mathcal{F}_h^N := \{F \in \mathcal{F}_h^b : F \subset \Gamma_N\}$ , such that  $\mathcal{F}_h^D \cup \mathcal{F}_h^N = \mathcal{F}_h^b$ .

Next, we introduce the discrete spaces

$$\underline{\mathbf{U}}_{h,g_D,D}^k := \{\underline{\mathbf{w}}_h \in \underline{\mathbf{U}}_h^k : w_F = \pi_F^k g_D \ \forall F \in \mathcal{F}_h^D\}. \quad (87)$$

Then, the HHO discretization of (86) reads as: Find  $\underline{\mathbf{u}}_h \in \underline{\mathbf{U}}_{h,g_D,D}^k$  such that

$$A_h(\underline{\mathbf{u}}_h, \underline{\mathbf{v}}_h) = (f, v_h)_\Omega + \sum_{F \in \mathcal{F}_h^N} (g_N, v_F)_F =: b_h(\underline{\mathbf{v}}_h) \quad \forall \underline{\mathbf{v}}_h \in \underline{\mathbf{U}}_{h,0,D}^k. \quad (88)$$

with  $A_h$  defined as in (17). Now, to compute  $\underline{\mathbf{u}}_h$ , we can proceed as described in the previous subsection. This motivates the introduction of  $\underline{\mathbf{w}}_{g_D,h} := \left( (w_{g_D,T})_{T \in \mathcal{T}_h}, (w_{g_D,F})_{F \in \mathcal{F}_h} \right) \in \underline{\mathbf{U}}_{h,g_D,D}^k$ , such that

$$w_{g_D,T} = 0 \quad \forall T \in \mathcal{T}_h, \quad w_{g_D,F} = 0 \quad \forall F \in \mathcal{F}_h^{\text{int}}, \quad w_{g_D,F} = 0 \quad \forall F \in \mathcal{F}_h^N.$$

After that, we compute  $\underline{\mathbf{u}}_h = \underline{\mathbf{u}}_{0,h} + \underline{\mathbf{w}}_{g_D,h}$ , where  $\underline{\mathbf{u}}_{0,h} \in \underline{\mathbf{U}}_{h,0,D}^k$  verifies

$$A_h(\underline{\mathbf{u}}_{0,h} + \underline{\mathbf{w}}_{g_D,h}, \underline{\mathbf{v}}_h) = (f, v_h)_\Omega + \sum_{F \in \mathcal{F}_h^N} (g_N, v_F)_F =: b_h(\underline{\mathbf{v}}_h) \quad \forall \underline{\mathbf{v}}_h \in \underline{\mathbf{U}}_{h,0,D}^k, \quad (89)$$

where the nonlinear form  $A_h$  is defined as in (17).

### 7.3 Nonhomogeneous Neumann boundary condition

Now, we consider a nonlinear problem with pure Neumann condition: Find  $u : \Omega \rightarrow \mathbb{R}$  such that

$$-\text{div}(\mathbf{a}(\cdot, \nabla u)) = f \quad \text{in } \Omega, \quad (90a)$$

$$\mathbf{a}(\cdot, \nabla u) \cdot \mathbf{n} = g \quad \text{on } \Gamma, \quad (90b)$$

$$\int_\Omega u = 0, \quad (90c)$$

where  $f \in L^2(\Omega)$ ,  $g \in L^2(\Gamma)$ . We look for the continuous weak solution in the Hilbert space

$$U := \{v \in H^1(\Omega) : (v, 1)_\Omega = 0\}. \quad (91)$$

The corresponding weak formulation to (90) is given as: Find  $u \in U$  such that

$$(\mathbf{a}(\cdot, \nabla u), \nabla v)_\Omega = (f, v)_\Omega + (g, v)_\Gamma \quad \forall v \in U. \quad (92)$$

Now, applying the HHO approach, we seek the discrete weak solution of (92) in the discrete space

$$\underline{\mathbf{U}}_h^{k,0} := \left\{ \underline{\mathbf{v}}_h \in \underline{\mathbf{U}}_h^k : \sum_{T \in \mathcal{T}_h} (v_T, 1)_T = 0 \right\}. \quad (93)$$

Then, the associated discrete HHO scheme reads as: *Find*  $\underline{\mathbf{u}}_h \in \underline{\mathbf{U}}_h^{k,0}$  *such that*

$$A_h(\underline{\mathbf{u}}_h, \underline{\mathbf{v}}_h) = (f, v_h)_\Omega + \sum_{F \in \mathcal{F}_h} (g, v_F)_F =: b_h(\underline{\mathbf{v}}_h) \quad \forall \underline{\mathbf{v}}_h \in \underline{\mathbf{U}}_h^{k,0}. \quad (94)$$

with the nonlinear form  $A_h$  given as in (17). The problem (94) has been analysed in a more general context in [20, Chapter 6], but the authors have only developed the a priori error analysis for the  $p$ -Laplacian nonlinear operator, which does not satisfy our hypotheses here, unless  $p = 2$  (that is the linear Laplacian operator). Their a priori error analysis does not cover our proposed kind of nonlinear operators.

## 8 Numerical results

In this section we present a comprehensive set of numerical tests with different boundary conditions, each of one with its corresponding explicit exact solution. We start by introducing the families of polytopal meshes we consider in our numerical computations. We call them: Simplicial I (cf. Figure 14), Cartesian (cf. Figure 15), Hexagonal (cf. Figure 18), and Tilted hexagonal (cf. Figure 19), which can be downloaded from <https://github.com/wareHHouse/diskpp>. Other families of meshes are called Graduated (Cartesian with hanging nodes) as in Figure 16 and Fractured (cf. Figure 20), which can be founded in FVCA 5 benchmark [37]. We also consider a family of Trapezoidal (cf. Figure 17) and uniform refined simplicial (Simplicial II) meshes (cf. Figure 21).

We remark that Example (8.1) provides HHO approximations for the nonlinear model problem (2), while Examples (8.2)-(8.5) solve (84), testing the robustness of the scheme considering smooth and non-smooth solutions. Examples (8.6)-(8.7) deal with (85) and (90), respectively. In all examples, we compute the following errors:

- $L^2$ -projection of potential error:  $\|\pi_h^k u - u_h\|_{0,\Omega},$
- Flux error:  $\|\nabla u - G_h^k \underline{\mathbf{u}}_h\|_{0,\Omega},$
- Reconstructive potential error:  $\|u - p_h^{k+1} \underline{\mathbf{u}}_h\|_{0,\Omega}.$

We also remark that we have considered different families of polytopal meshes, and three nonlinear functions  $\mathbf{a}$ , that satisfy the hypotheses **(H.1)**, **(H.2)**, **(H.3)**, and **(H.4)**. The list of examples with their corresponding nonlinear function  $\mathbf{a}$ , the exact solution, domain, polytopal meshes considered (see Section 8.8), and type of boundary conditions (B.C.), are resumed in Table 1.

For each one of the examples presented here, we approximate the exact solution with piecewise polynomials of degree at most  $k$ , with  $k \in \{0, 1, 2, 3, 4\}$ . Besides, the experimental order of convergence (*rate*), is computed as

$$rate = \log(e_{\mathcal{T}}/e_{\tilde{\mathcal{T}}})/\log(h_{\mathcal{T}}/h_{\tilde{\mathcal{T}}}),$$

where  $e_{\mathcal{T}}$  and  $e_{\tilde{\mathcal{T}}}$  represent the errors associated to two consecutive meshes  $\mathcal{T}$  and  $\tilde{\mathcal{T}}$ , respectively.

The nonlinear algebraic system, obtained from (17), is solved by Newton's method with the initial guess given by the solution of the associated Poisson problem. This choice leads to a



considerable reduction of the number of iterations, instead of using the usual null initial guess. We also consider a tolerance of  $10^{-10}$  for the corresponding residual, and fixed the maximum number of Newton's iterations in 10. The number of iterations needed to attain the prescribed tolerance, in all the examples presented here, is  $\leq 5$ .

Ex.	$\mathbf{a}(\cdot, \zeta)$	$u$	$\Omega$	Polytopal Mesh	B.C.
1	$\left(1 + \frac{1}{1+ \zeta ^2}\right) \zeta$	$\sin(\pi x) \sin(\pi y)$	Square	Fractured Hexagonal	Dirichlet
2	$(1 + \exp(- \zeta ^2)) \zeta$	$\exp(x + \pi y)$	Square	Graduated	Dirichlet
3	$\left(2 + \frac{1}{1+ \zeta }\right) \zeta$	$\cos\left(\frac{\pi}{2}y\right) + \chi(x)x^{3.5}$	Square	Simplicial III	Dirichlet
4	$\left(2 + \frac{1}{1+ \zeta }\right) \zeta$	$\cos\left(\frac{\pi}{2}y\right) + \chi(x)x^{3.5}$	$L$ -shaped	Simplicial II	Dirichlet
5	$\left(2 + \frac{1}{1+ \zeta }\right) \zeta$	$r^{2/3} \sin(2\theta/3)$	$L$ -shaped	Simplicial II	Dirichlet
6	$(1 + \exp(- \zeta ^2)) \zeta$	$\sin(\pi x) \sin(\pi y)$	Square	Cartesian Hexagonal	Mixed
7	$\left(1 + \frac{1}{1+ \zeta ^2}\right) \zeta$	$\exp(x + \pi y)$	Square	Trapezoidal Tilted	Neumann

Table 1: Summary of data for the 7 examples

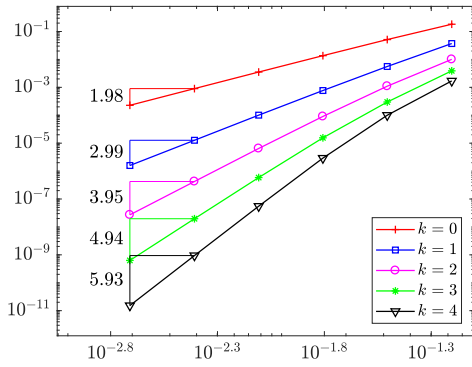
The computational code is based on the one developed by Di Pietro in [19], which is described in [9]. However, we have implemented the computations of the modified versions of the discrete gradient and potential operators (cf. (10) and (12)). The static condensation procedure has been implemented right after the assembling of the local matrices of the linearized problem over the global system. This allows us to recompute, on each iteration, the volumetric terms.

## 8.1 Example 1

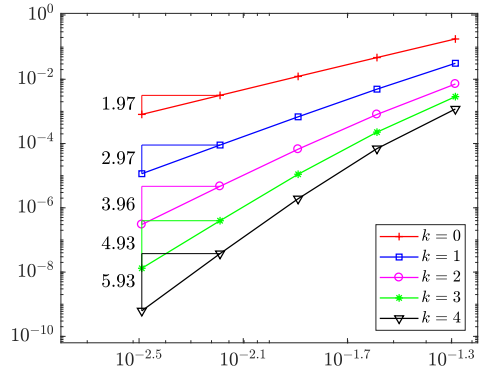
We consider the nonlinear problem (2) with homogeneous Dirichlet boundary condition, defined in the convex domain  $\Omega := (0, 1)^2$ , and with nonlinear coefficient  $\mathbf{a}(\cdot, \zeta) := \left(1 + \frac{1}{1+|\zeta|^2}\right) \zeta$ , for all  $\zeta \in \mathbb{R}^2$ . The datum  $f$  is chosen so that the exact solution is given by the smooth function  $u(x, y) = \sin(\pi x) \sin(\pi y)$ . Here, we consider two families of polytopal meshes: a nonconforming one, that we have called Fractured in Table 1 (cf. Figure 20), and conforming meshes with hexagonal-dominant cells (cf. Figure 18). Tables 2 and 3 show the history of convergence of potential, flux, and reconstructive potential errors, when the exact solution is approximated with piecewise polynomials of degree at most  $k \in \{0, 1, 2, 3, 4\}$ , computed using families of Fractured and Hexagonal meshes, respectively. From both tables, we notice that the scheme converges with the order of convergence of the aforementioned errors behaves as  $\mathcal{O}(h^{k+2})$ ,  $\mathcal{O}(h^{k+1})$  and  $\mathcal{O}(h^{k+2})$ , respectively. These are in agreement with Theorems 6.1 and 6.2. Figures 1 and 3 resume the information provided in Tables 2 and 3, respectively.

Table 2: Histories of convergence of potential, flux, and reconstructive potential errors, considering  $k \in \{0, 1, 2, 3, 4\}$  (Example 1 - Fractured meshes)

Potential										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
6.25e-02	1.82e-01		3.70e-02		1.00e-02		3.89e-03		1.70e-03	
3.12e-02	5.14e-02	1.821	5.58e-03	2.725	1.10e-03	3.185	2.98e-04	3.701	1.00e-04	4.074
1.56e-02	1.36e-02	1.917	7.69e-04	2.860	9.06e-05	3.600	1.55e-05	4.262	2.91e-06	5.103
7.81e-03	3.53e-03	1.952	1.01e-04	2.937	6.45e-06	3.819	5.86e-07	4.737	5.52e-08	5.733
3.91e-03	9.00e-04	1.976	1.27e-05	2.994	4.24e-07	3.936	1.96e-08	4.907	9.41e-10	5.885
1.95e-03	2.27e-04	1.979	1.59e-06	2.990	2.71e-08	3.954	6.31e-10	4.943	1.52e-11	5.927
Flux										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
6.25e-02	1.04e-01		3.08e-02		9.07e-03		3.40e-03		1.36e-03	
3.12e-02	2.91e-02	1.837	7.68e-03	1.999	1.88e-03	2.265	4.20e-04	3.013	1.39e-04	3.284
1.56e-02	7.80e-03	1.898	1.78e-03	2.106	2.60e-04	2.855	3.57e-05	3.556	7.19e-06	4.268
7.81e-03	2.08e-03	1.907	3.53e-04	2.343	2.81e-05	3.215	2.39e-06	3.904	2.20e-07	5.041
3.91e-03	5.39e-04	1.955	6.15e-05	2.523	2.46e-06	3.519	1.27e-07	4.248	5.91e-09	5.227
1.95e-03	1.37e-04	1.972	9.97e-06	2.617	1.88e-07	3.701	6.38e-09	4.295	1.50e-10	5.275
Reconstructive potential										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
6.25e-02	1.94e-01		3.82e-02		1.01e-02		3.90e-03		1.70e-03	
3.12e-02	5.37e-02	1.850	5.71e-03	2.737	1.10e-03	3.189	2.98e-04	3.702	1.00e-04	4.075
1.56e-02	1.41e-02	1.926	7.83e-04	2.868	9.08e-05	3.601	1.55e-05	4.262	2.91e-06	5.103
7.81e-03	3.65e-03	1.956	1.02e-04	2.942	6.46e-06	3.820	5.86e-07	4.737	5.52e-08	5.733
3.91e-03	9.30e-04	1.977	1.29e-05	2.997	4.24e-07	3.937	1.96e-08	4.907	9.41e-10	5.885
1.95e-03	2.35e-04	1.980	1.60e-06	2.992	2.71e-08	3.954	6.31e-10	4.943	1.52e-11	5.927



(a) Fractured



(b) Hexagonal

Figure 1: Rates of convergence of potential error vs.  $h$  (Example 1)

Table 3: Histories of convergence of potential, flux, and reconstructive potential errors, considering  $k \in \{0, 1, 2, 3, 4\}$  (Example 1 - Hexagonal meshes)

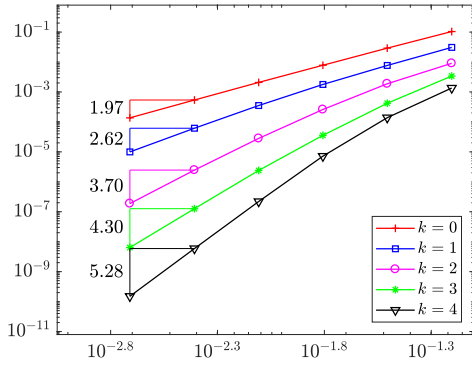
Potential										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
5.18e-02	1.79e-01		3.14e-02		7.13e-03		2.88e-03		1.19e-03	
2.59e-02	4.73e-02	1.920	4.92e-03	2.675	8.04e-04	3.148	2.26e-04	3.669	6.94e-05	4.096
1.29e-02	1.23e-02	1.939	6.78e-04	2.844	6.67e-05	3.572	1.11e-05	4.329	1.92e-06	5.151
6.47e-03	3.15e-03	1.968	8.98e-05	2.930	4.67e-06	3.854	3.97e-07	4.824	3.75e-08	5.699
3.24e-03	8.06e-04	1.973	1.15e-05	2.969	3.02e-07	3.959	1.31e-08	4.927	6.22e-10	5.927

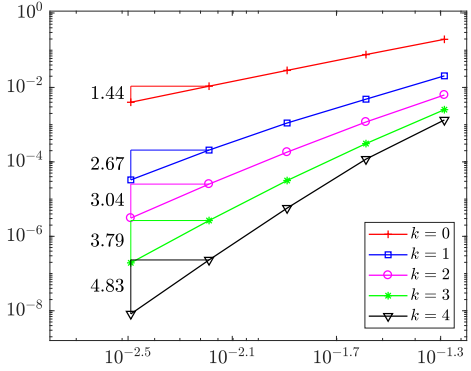
Flux										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
5.18e-02	1.96e-01		2.05e-02		6.27e-03		2.53e-03		1.34e-03	
2.59e-02	7.63e-02	1.361	4.83e-03	2.082	1.18e-03	2.412	3.10e-04	3.027	1.18e-04	3.499
1.29e-02	2.87e-02	1.404	1.11e-03	2.114	1.83e-04	2.676	3.15e-05	3.280	5.72e-06	4.346
6.47e-03	1.08e-02	1.416	2.07e-04	2.428	2.53e-05	2.865	2.66e-06	3.584	2.32e-07	4.646
3.24e-03	3.99e-03	1.438	3.28e-05	2.666	3.09e-06	3.039	1.94e-07	3.785	8.22e-09	4.828

Reconstructive potential										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
5.18e-02	1.96e-01		3.19e-02		7.17e-03		2.88e-03		1.19e-03	
2.59e-02	5.04e-02	1.960	4.97e-03	2.682	8.06e-04	3.153	2.26e-04	3.670	6.95e-05	4.097
1.29e-02	1.28e-02	1.962	6.83e-04	2.847	6.67e-05	3.575	1.11e-05	4.329	1.92e-06	5.151
6.47e-03	3.27e-03	1.981	9.03e-05	2.932	4.67e-06	3.855	3.97e-07	4.824	3.75e-08	5.699
3.24e-03	8.31e-04	1.981	1.16e-05	2.970	3.02e-07	3.959	1.31e-08	4.927	6.22e-10	5.927

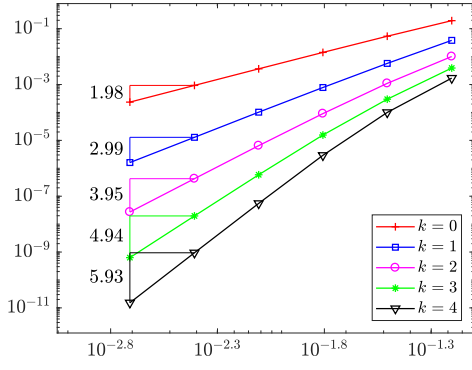


(a) Fractured

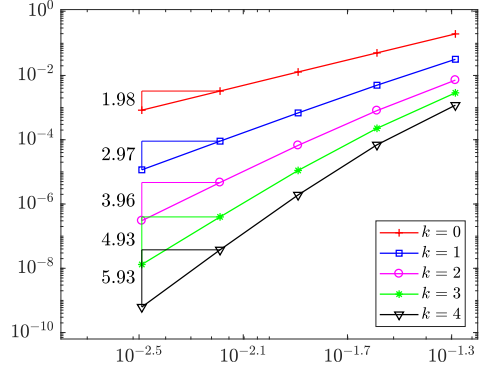


(b) Hexagonal

Figure 2: Rates of convergence of flux error vs.  $h$  (Example 1)

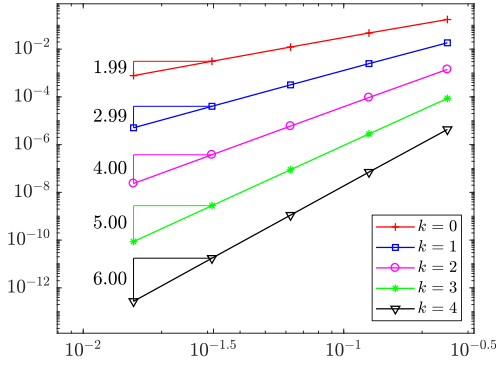


(a) Fracture

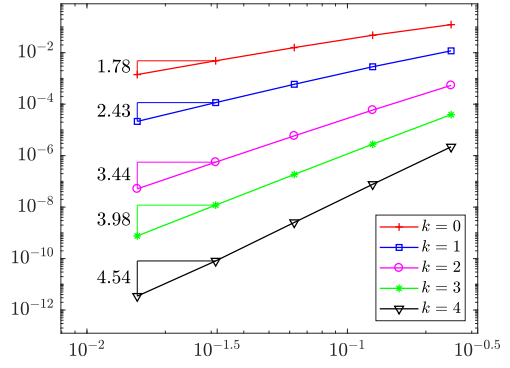


(b) Hexagonal

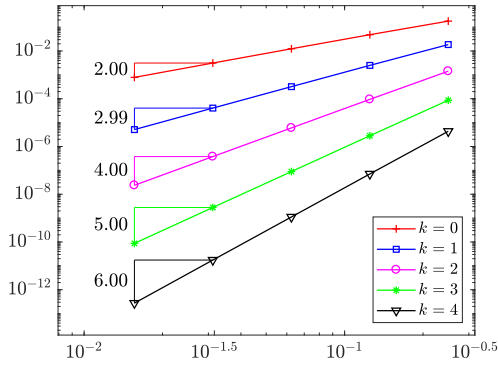
Figure 3: Rates of convergence of reconstructive potential errors (Example 1)



(a) Potential error vs.  $h$



(b) Flux error vs.  $h$



(c) Reconstructive potential error vs.  $h$

Figure 4: Rates of convergence of potential, flux and reconstructive potential errors (Example 2)

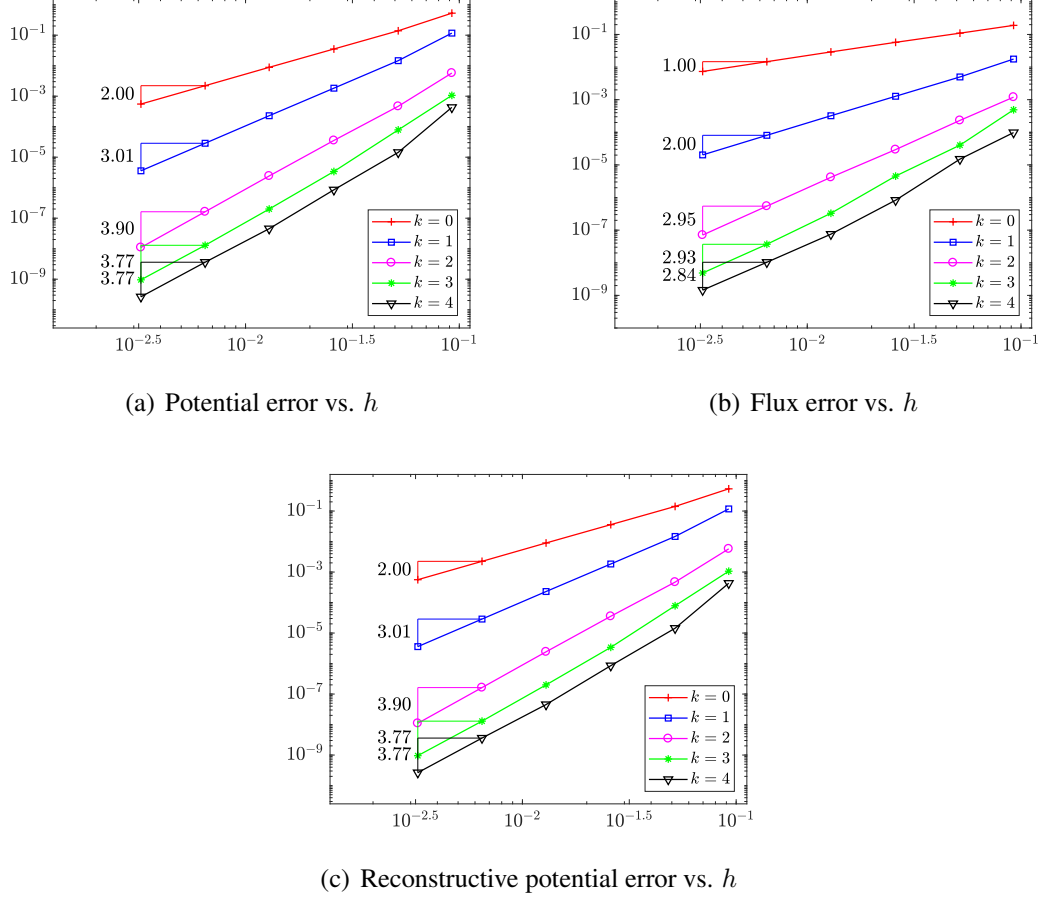


Figure 5: Rates of convergence of potential, flux and reconstructive potential errors (Example 3)

## 8.2 Example 2

Here, we solve the HHO discrete nonlinear scheme (84), corresponding to the nonlinear problem (81). We take  $\Omega := (0, 1)^2$  as the domain, and the nonlinear coefficient  $a(\cdot, \zeta) := (1 + \exp(-|\zeta|^2))\zeta$ , for all  $\zeta \in \mathbb{R}^2$ . We choose the data  $f$  and  $g$  so that the solution  $u$  is the smooth function given by  $u(x, y) := \exp(x + \pi y)$ . The family of meshes considered for this example, corresponds to a (nonconforming) graduated rectangular meshes (cf. Figure 16). Figure 4 shows the well behavior of the potential, flux and reconstructive potential errors, measure in the  $L^2$ -norm, with respect to the nonhomogeneous. Table 4 contains their respective histories of convergence, which are in agreement with Theorems 6.1, 6.2 and 6.3. We point out that for  $k = 4$ , round-off errors are probably affecting the flux error in the last refinement iteration mesh. On the other hand, the apparent super-convergence phenomenon for the flux error (for  $k \in \{0, 1, 2\}$ ) meshes, could be linked to the fact that the mesh quality improves when refining.

Table 4: Histories of convergence of potential, flux, and reconstructive potential errors, considering  $k \in \{0, 1, 2, 3, 4\}$  (Example 2 - Graduated meshes)

Potential										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
2.50e-01	1.74e-01		1.84e-02		1.42e-03		8.61e-05		4.33e-06	
1.25e-01	4.74e-02	1.878	2.48e-03	2.895	9.37e-05	3.921	2.81e-06	4.940	7.05e-08	5.939
6.25e-02	1.22e-02	1.953	3.19e-04	2.959	5.97e-06	3.973	8.86e-08	4.985	1.11e-09	5.984
3.12e-02	3.10e-03	1.979	4.03e-05	2.975	3.76e-07	3.980	2.78e-09	4.984	1.75e-11	5.982
1.56e-02	7.77e-04	1.995	5.07e-06	2.992	2.36e-08	3.995	8.69e-11	4.999	2.72e-13	6.004
Flux										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
2.50e-01	1.23e-01		1.18e-02		5.38e-04		3.92e-05		2.20e-06	
1.25e-01	4.75e-02	1.372	2.84e-03	2.054	5.88e-05	3.192	2.78e-06	3.818	7.69e-08	4.835
6.25e-02	1.58e-02	1.592	5.92e-04	2.263	5.86e-06	3.327	1.85e-07	3.910	2.53e-09	4.927
3.12e-02	4.84e-03	1.700	1.15e-04	2.360	5.57e-07	3.389	1.19e-08	3.945	8.09e-11	4.955
1.56e-02	1.41e-03	1.778	2.13e-05	2.427	5.12e-08	3.443	7.58e-10	3.977	3.47e-12	4.543
Reconstructive Potential										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
2.50e-01	1.79e-01		1.86e-02		1.43e-03		8.64e-05		4.33e-06	
1.25e-01	4.82e-02	1.889	2.50e-03	2.896	9.41e-05	3.922	2.81e-06	4.940	7.07e-08	5.939
6.25e-02	1.24e-02	1.959	3.21e-04	2.960	6.00e-06	3.973	8.89e-08	4.985	1.12e-09	5.984
3.12e-02	3.13e-03	1.982	4.06e-05	2.976	3.78e-07	3.980	2.79e-09	4.984	1.75e-11	5.982
1.56e-02	7.85e-04	1.996	5.10e-06	2.992	2.37e-08	3.995	8.71e-11	4.999	2.74e-13	5.998

### 8.3 Example 3

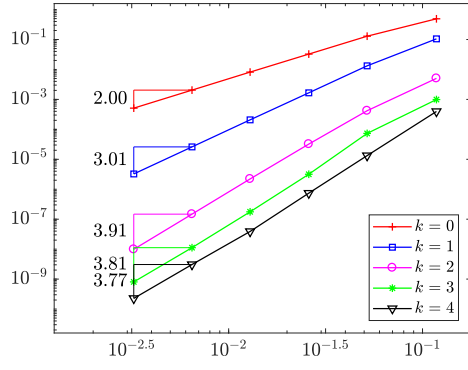
Now, we consider a nonhomogeneous Dirichlet nonlinear problem, defined in the convex domain  $\Omega = (-1, 1)^2$ . The nonlinear coefficient is setting as  $\mathbf{a}(\cdot, \zeta) := (2 + \frac{1}{1+|\zeta|})\zeta$ , for all  $\zeta \in \mathbb{R}^2$ , and the exact solution is given by  $u(x, y) = \cos(\frac{\pi}{2}y) + \chi(x)x^{3.5}$ , where  $\chi(x)$  denotes the characteristic function on  $[0, 1]$  with respect to  $x$ . The HHO approximations are computed considering a family of uniform refined simplicial meshes, that we have called Simplicial-III (cf. Figure 22). We remark that  $u \in H^4(\Omega)$  but does not belong to  $H^{4+\epsilon}(\Omega)$ , for all  $\epsilon > 0$ . We report the histories of convergence of the potential, flux and reconstructive potential errors, with respect to the nonhomogeneous, in Table 5. The results are in agreement with the corresponding versions of Theorems 6.1, 6.2 and 6.3, since the solution  $u$  has a limited regularity. This is the reason why we observe the optimal rates of convergence only for  $k \in \{0, 1, 2\}$ . In Figure 5 we display convergence results for the refined triangulations and polynomial degrees up to 4.

Table 5: Histories of convergence of potential, flux, and reconstructive potential errors, considering  $k \in \{0, 1, 2, 3, 4\}$  (Example 3 - Simplicial-III meshes)

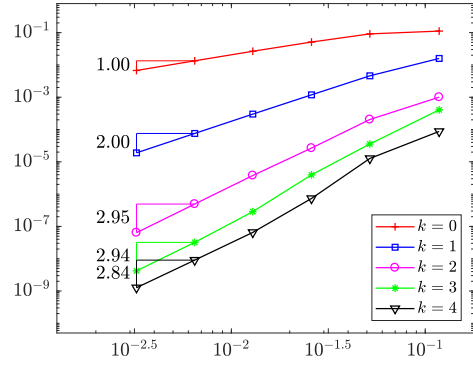
Potential										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
9.23e-02	5.27e-01		1.17e-01		5.74e-03		1.06e-03		4.29e-04	
5.18e-02	1.39e-01	2.302	1.46e-02	3.600	4.66e-04	4.346	7.85e-05	4.512	1.44e-05	5.881
2.59e-02	3.52e-02	1.985	1.83e-03	2.997	3.54e-05	3.720	3.39e-06	4.531	8.51e-07	4.077
1.29e-02	8.82e-03	1.985	2.29e-04	2.982	2.42e-06	3.848	1.99e-07	4.067	4.52e-08	4.211
6.47e-03	2.21e-03	2.008	2.86e-05	3.013	1.62e-07	3.915	1.29e-08	3.965	3.60e-09	3.667
3.24e-03	5.52e-04	2.004	3.58e-06	3.007	1.10e-08	3.897	9.54e-10	3.768	2.65e-10	3.774
Flux										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
9.23e-02	1.90e-01		1.76e-02		1.20e-03		4.90e-04		9.81e-05	
5.18e-02	1.09e-01	0.953	4.98e-03	2.187	2.32e-04	2.845	4.04e-05	4.321	1.50e-05	3.246
2.59e-02	5.73e-02	0.934	1.27e-03	1.967	2.95e-05	2.973	4.55e-06	3.149	8.28e-07	4.183
1.29e-02	2.91e-02	0.971	3.22e-04	1.973	4.17e-06	2.805	3.29e-07	3.768	7.52e-08	3.441
6.47e-03	1.46e-02	0.997	8.09e-05	2.001	5.44e-07	2.952	3.69e-08	3.171	1.04e-08	2.873
3.24e-03	7.33e-03	1.000	2.03e-05	2.000	7.08e-08	2.950	4.85e-09	2.933	1.45e-09	2.844
Reconstructive potential										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
9.23e-02	5.35e-01		1.17e-01		5.77e-03		1.06e-03		4.29e-04	
5.18e-02	1.42e-01	2.301	1.47e-02	3.600	4.68e-04	4.348	7.85e-05	4.513	1.44e-05	5.881
2.59e-02	3.58e-02	1.984	1.84e-03	2.997	3.55e-05	3.721	3.40e-06	4.531	8.51e-07	4.077
1.29e-02	8.97e-03	1.985	2.30e-04	2.982	2.43e-06	3.848	1.99e-07	4.067	4.52e-08	4.211
6.47e-03	2.24e-03	2.008	2.87e-05	3.013	1.63e-07	3.915	1.29e-08	3.965	3.60e-09	3.667
3.24e-03	5.61e-04	2.004	3.59e-06	3.007	1.10e-08	3.897	9.55e-10	3.769	2.65e-10	3.774

## 8.4 Example 4

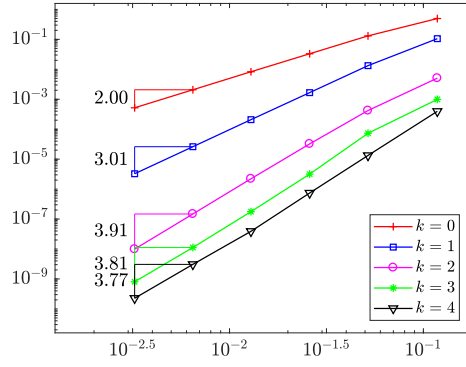
Here, we consider the same problem as in Example 8.3, but defined in the  $L$ -shaped domain  $\Omega = (-1, 1)^2/[0, 1] \times [-1, 0]$ , which is clearly non-convex. We choose a family of uniform refined simplicial meshes (named Simplicial-II) (cf. Figure 21), to establish the HHO formulation. In Figure 6 we display convergence results for the refined triangulations and polynomial degrees up to 4. Despite the non-convexity of the domain, which could not help to ensure the additional regularity require for the  $L^2$ -error estimate of the potential as well as of the reconstructive potential, the obtained results behave similarly to the ones provided in Example (8.3) (in a convex domain), and are still in agreement with Theorems 6.1, 6.2 and 6.3.



(a) Potential error vs.  $h$



(b) Flux error vs.  $h$



(c) Reconstructive potential error vs.  $h$

Figure 6: Rates of convergence of potential, flux and reconstructive potential errors (Example 3)



Table 6: Histories of convergence of potential, flux, and reconstructive potential errors, considering  $k \in \{0, 1, 2, 3, 4\}$  (Example 4 - Simplicial-II meshes)

Potential										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
1.18e-01	4.91e-01		1.05e-01		5.05e-03		9.89e-04		3.96e-04	
5.18e-02	1.29e-01	1.624	1.32e-02	2.520	4.19e-04	3.023	7.32e-05	3.163	1.33e-05	4.123
2.59e-02	3.26e-02	1.985	1.66e-03	2.994	3.22e-05	3.701	3.17e-06	4.530	7.41e-07	4.164
1.29e-02	8.17e-03	1.985	2.08e-04	2.981	2.21e-06	3.846	1.76e-07	4.143	3.89e-08	4.227
6.47e-03	2.04e-03	2.008	2.60e-05	3.013	1.47e-07	3.922	1.12e-08	3.996	3.06e-09	3.683
3.24e-03	5.11e-04	2.004	3.25e-06	3.006	9.87e-09	3.910	8.05e-10	3.806	2.27e-10	3.766
Flux										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
1.18e-01	1.12e-01		1.59e-02		1.01e-03		4.06e-04		8.74e-05	
5.18e-02	9.14e-02	0.244	4.60e-03	1.505	2.05e-04	1.936	3.57e-05	2.955	1.27e-05	2.343
2.59e-02	5.11e-02	0.838	1.19e-03	1.957	2.66e-05	2.942	3.95e-06	3.176	7.34e-07	4.114
1.29e-02	2.65e-02	0.942	3.00e-04	1.969	3.77e-06	2.804	2.87e-07	3.761	6.52e-08	3.472
6.47e-03	1.34e-02	0.987	7.56e-05	1.999	4.93e-07	2.949	3.20e-08	3.176	9.05e-09	2.863
3.24e-03	6.74e-03	0.997	1.90e-05	2.000	6.40e-08	2.951	4.20e-09	2.938	1.27e-09	2.845
Reconstructive potential										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
1.18e-01	4.96e-01		1.06e-01		5.07e-03		9.90e-04		3.96e-04	
5.18e-02	1.31e-01	1.620	1.33e-02	2.520	4.20e-04	3.025	7.32e-05	3.164	1.33e-05	4.123
2.59e-02	3.31e-02	1.984	1.67e-03	2.994	3.23e-05	3.702	3.17e-06	4.530	7.41e-07	4.164
1.29e-02	8.29e-03	1.984	2.08e-04	2.981	2.21e-06	3.846	1.76e-07	4.143	3.89e-08	4.227
6.47e-03	2.08e-03	2.008	2.61e-05	3.013	1.48e-07	3.922	1.12e-08	3.996	3.07e-09	3.683
3.24e-03	5.19e-04	2.004	3.26e-06	3.006	9.89e-09	3.910	8.05e-10	3.806	2.27e-10	3.766

## 8.5 Example 5

We consider again the  $L$ -shaped domain  $\Omega := (-1, -1)^2 \setminus [0, 1] \times [-1, 0]$ , with  $\Gamma := \partial\Omega$ , and another nonlinear coefficient, given by  $\mathbf{a}(\cdot, \zeta) := \left(2 + \frac{1}{1+|\zeta|}\right) \zeta$ , for all  $\zeta \in \mathbb{R}^2$ . We choose the data  $f$  and  $g$ , so that the solution exact  $u$  is the non-smooth function given (in polar coordinates) by  $u(r, \theta) = r^{2/3} \sin(2\theta/3)$ . We point out that  $u \in H^{1+\frac{2}{3}-s}(\Omega)$ , for some  $s > 0$ . We test again over a family of uniform refined simplicial meshes (cf. Figure 21). In Figure 7 we display the rates of convergence corresponding to the potential, flux and reconstructive potential errors, vs.  $h$ , considering polynomials degrees up to 4. The histories of convergence of these errors, are reported in Table 7. We observe that the potential and reconstructive potential errors go to zero at a rate that behaves as  $\mathcal{O}(h^{4/3})$ , while the flux error decreases to zero as  $\mathcal{O}(h^{2/3})$ , for polynomial degrees up to 4.

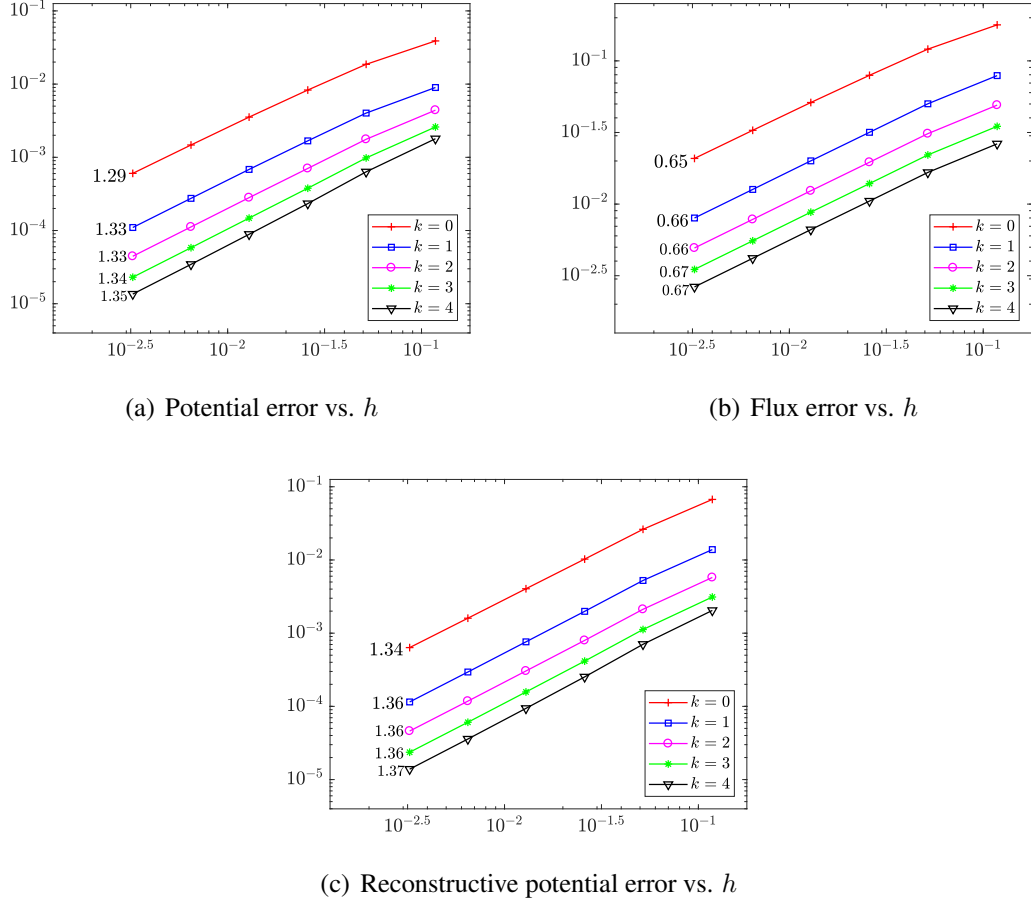


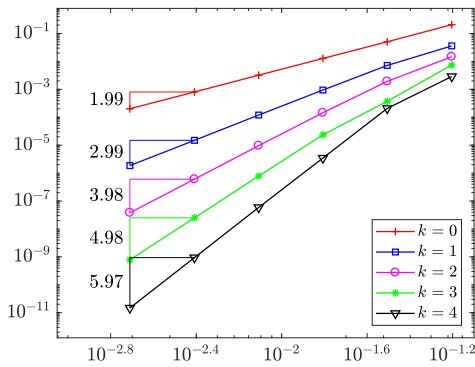
Figure 7: Rates of convergence of potential, flux and reconstructive potential errors (Example 5)

## 8.6 Example 6

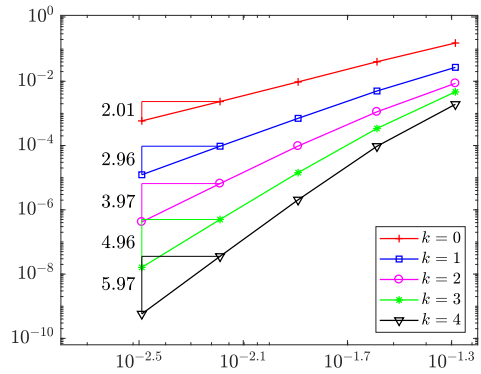
Here, we consider the nonlinear problem with mixed boundary conditions (85), defined in the domain  $\Omega := (0, 1)^2$ , with  $\Gamma_N = \{1\} \times [0, 1]$  and  $\Gamma_D := \Gamma \setminus \Gamma_N$ . The chosen nonlinear coefficient is given by  $\mathbf{a}(\cdot, \zeta) := (1 + \exp(-|\zeta|^2))\zeta$ , for all  $\zeta \in \mathbb{R}^2$ , while the data  $f$ ,  $g_D$  and  $g_N$  are chosen such that the solution  $u$  is given again by  $u(x, y) = \sin(\pi x) \sin(\pi y)$ . We remark that with this choice,  $g_D = 0$ . We solve the HHO nonlinear scheme 88, considering two families of polytopal meshes. The first one is a family of Cartesian meshes (cf. Figure 15), and the second one is a family of Hexagonal-dominant meshes (cf. Figure 18). Tables 8 and 9 report the histories of convergence of the potential, flux and reconstructive potential errors, for the family of Cartesian and the Hexagonal meshes, respectively. We notice certain super-convergence phenomenon for the flux error on the Cartesian (for  $k \in \{0, 1, 2\}$ ) and Hexagonal (for  $k \in \{0, 1\}$ ) meshes, which could be related to the fact that the mesh quality improves when refining. The numerical results provided by this example support the conjecture that the present approach might behave quite well even in a case not fully covered by the theoretical results.

Table 7: Histories of convergence of potential, flux, and reconstructive potential errors, considering  $k \in \{0, 1, 2, 3, 4\}$  (Example 5 - Simplicial II meshes)

Potential										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
1.18e-01	3.89e-02		8.98e-03		4.41e-03		2.60e-03		1.79e-03	
5.18e-02	1.86e-02	0.894	4.02e-03	0.977	1.76e-03	1.112	9.80e-04	1.183	6.33e-04	1.265
2.59e-02	8.30e-03	1.166	1.68e-03	1.261	7.05e-04	1.323	3.78e-04	1.375	2.33e-04	1.441
1.29e-02	3.54e-03	1.222	6.83e-04	1.288	2.81e-04	1.318	1.48e-04	1.347	8.87e-05	1.386
6.47e-03	1.47e-03	1.271	2.75e-04	1.319	1.12e-04	1.334	5.82e-05	1.350	3.45e-05	1.370
3.24e-03	6.03e-04	1.291	1.10e-04	1.326	4.45e-05	1.334	2.30e-05	1.342	1.35e-05	1.352
Flux										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
1.18e-01	1.78e-01		7.89e-02		4.91e-02		3.49e-02		2.63e-02	
5.18e-02	1.21e-01	0.474	5.02e-02	0.551	3.10e-02	0.557	2.20e-02	0.559	1.66e-02	0.559
2.59e-02	7.93e-02	0.607	3.17e-02	0.661	1.96e-02	0.663	1.39e-02	0.664	1.05e-02	0.664
1.29e-02	5.12e-02	0.628	2.00e-02	0.659	1.24e-02	0.659	8.78e-03	0.660	6.63e-03	0.660
6.47e-03	3.27e-02	0.648	1.27e-02	0.665	7.82e-03	0.666	5.54e-03	0.667	4.18e-03	0.667
3.24e-03	2.08e-02	0.654	8.00e-03	0.664	4.94e-03	0.665	3.50e-03	0.666	2.64e-03	0.665
Reconstructive potential										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
1.18e-01	6.71e-02		1.39e-02		5.77e-03		3.12e-03		2.04e-03	
5.18e-02	2.62e-02	1.142	5.24e-03	1.180	2.12e-03	1.216	1.12e-03	1.242	7.04e-04	1.294
2.59e-02	1.03e-02	1.352	1.99e-03	1.401	7.96e-04	1.412	4.15e-04	1.434	2.53e-04	1.479
1.29e-02	4.04e-03	1.337	7.61e-04	1.376	3.05e-04	1.379	1.57e-04	1.392	9.39e-05	1.420
6.47e-03	1.60e-03	1.344	2.95e-04	1.375	1.18e-04	1.375	6.06e-05	1.382	3.58e-05	1.397
3.24e-03	6.35e-04	1.336	1.15e-04	1.362	4.60e-05	1.361	2.36e-05	1.363	1.39e-05	1.371



(a) Cartesian



(b) Hexagonal

Figure 8: Potential error vs.  $h$

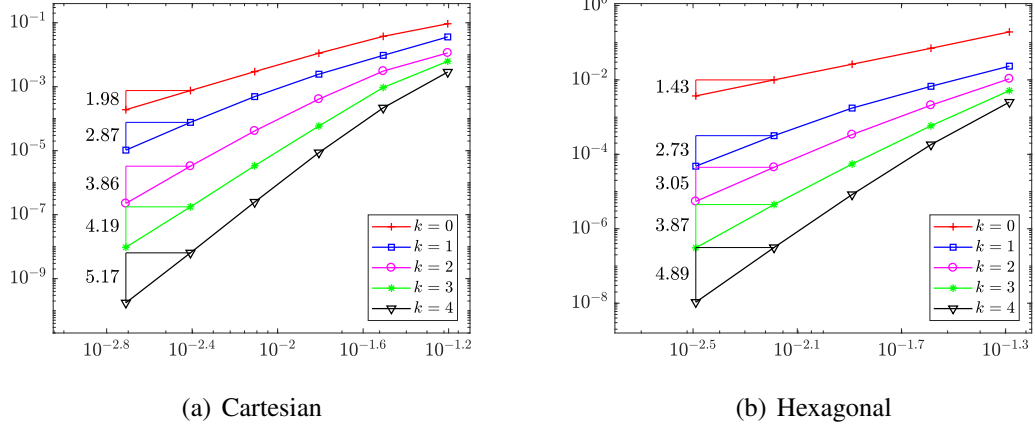


Figure 9: Flux error vs.  $h$

Table 8: Histories of convergence of potential, flux, and reconstructive potential errors, considering  $k \in \{0, 1, 2, 3, 4\}$  (Example 6 - Cartesian meshes)

Potential										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
6.25e-02	2.07e-01		3.60e-02		1.48e-02		7.43e-03		2.88e-03	
3.12e-02	5.06e-02	2.029	7.17e-03	2.321	1.93e-03	2.933	3.74e-04	4.303	2.07e-04	3.786
1.56e-02	1.28e-02	1.985	9.43e-04	2.927	1.45e-04	3.731	2.34e-05	3.996	3.45e-06	5.909
7.81e-03	3.21e-03	1.995	1.20e-04	2.982	9.60e-06	3.924	7.85e-07	4.908	5.87e-08	5.889
3.91e-03	8.04e-04	2.002	1.49e-05	3.014	6.05e-07	3.996	2.49e-08	4.986	9.36e-10	5.982
1.95e-03	2.01e-04	1.992	1.86e-06	2.994	3.78e-08	3.984	7.82e-10	4.976	1.47e-11	5.970
Flux										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
6.25e-02	9.31e-02		3.61e-02		1.14e-02		6.26e-03		2.88e-03	
3.12e-02	3.76e-02	1.306	9.57e-03	1.910	3.08e-03	1.882	9.44e-04	2.723	2.19e-04	3.712
1.56e-02	1.12e-02	1.749	2.47e-03	1.956	4.10e-04	2.911	5.91e-05	3.999	8.69e-06	4.653
7.81e-03	2.97e-03	1.916	4.91e-04	2.332	4.16e-05	3.309	3.39e-06	4.130	2.48e-07	5.143
3.91e-03	7.60e-04	1.970	7.68e-05	2.682	3.28e-06	3.670	1.76e-07	4.279	6.40e-09	5.284
1.95e-03	1.91e-04	1.984	1.04e-05	2.872	2.24e-07	3.857	9.53e-09	4.187	1.75e-10	5.173
Reconstructive potential										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
6.25e-02	2.19e-01		3.76e-02		1.48e-02		7.44e-03		2.88e-03	
3.12e-02	5.32e-02	2.035	7.30e-03	2.360	1.93e-03	2.936	3.75e-04	4.301	2.08e-04	3.787
1.56e-02	1.34e-02	1.993	9.58e-04	2.930	1.45e-04	3.731	2.34e-05	3.999	3.45e-06	5.909
7.81e-03	3.35e-03	1.999	1.21e-04	2.986	9.61e-06	3.925	7.85e-07	4.908	5.87e-08	5.889
3.91e-03	8.38e-04	2.003	1.50e-05	3.017	6.05e-07	3.997	2.49e-08	4.986	9.36e-10	5.982
1.95e-03	2.09e-04	1.993	1.87e-06	2.995	3.78e-08	3.984	7.82e-10	4.976	1.47e-11	5.970

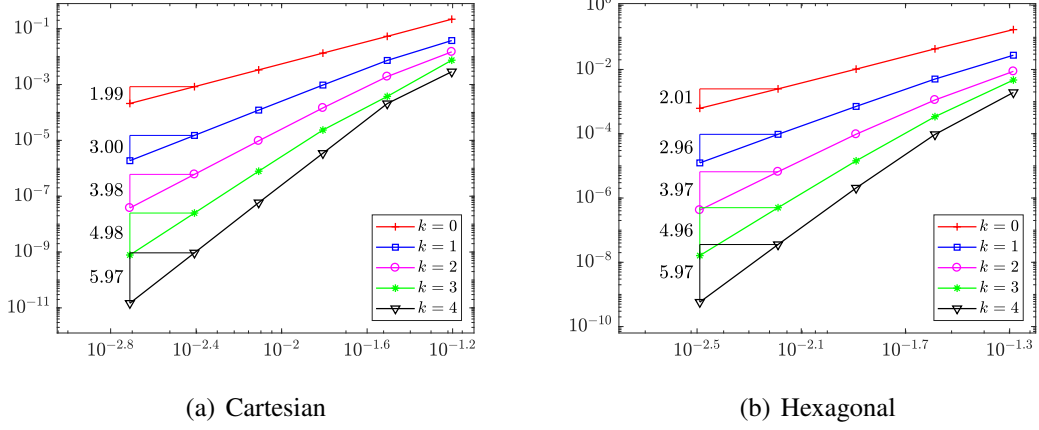


Figure 10: Rates of convergence of potential, flux and reconstructive potential errors (Example 6)

Table 9: Histories of convergence of potential, flux, and reconstructive potential errors, considering  $k \in \{0, 1, 2, 3, 4\}$  (Example 6 - Hexagonal)

Potential										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
5.18e-02	1.55e-01		2.71e-02		8.64e-03		4.68e-03		1.92e-03	
2.59e-02	4.05e-02	1.935	4.97e-03	2.446	1.12e-03	2.949	3.39e-04	3.786	9.48e-05	4.340
1.29e-02	9.55e-03	2.073	7.01e-04	2.811	9.64e-05	3.517	1.43e-05	4.540	2.05e-06	5.502
6.47e-03	2.34e-03	2.041	9.53e-05	2.893	6.55e-06	3.897	5.00e-07	4.863	3.57e-08	5.870
3.24e-03	5.84e-04	2.005	1.23e-05	2.958	4.20e-07	3.971	1.62e-08	4.962	5.76e-10	5.966
Flux										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
5.18e-02	1.94e-01		2.34e-02		1.07e-02		5.15e-03		2.53e-03	
2.59e-02	7.06e-02	1.460	6.72e-03	1.800	2.09e-03	2.360	5.84e-04	3.140	1.82e-04	3.793
1.29e-02	2.63e-02	1.415	1.75e-03	1.931	3.39e-04	2.607	5.49e-05	3.392	8.31e-06	4.430
6.47e-03	9.98e-03	1.406	3.17e-04	2.474	4.48e-05	2.932	4.47e-06	3.635	3.13e-07	4.751
3.24e-03	3.71e-03	1.431	4.80e-05	2.730	5.43e-06	3.053	3.08e-07	3.867	1.06e-08	4.891
Reconstructive potential										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
5.18e-02	1.74e-01		2.77e-02		8.74e-03		4.69e-03		1.92e-03	
2.59e-02	4.38e-02	1.989	5.03e-03	2.462	1.12e-03	2.960	3.40e-04	3.787	9.49e-05	4.341
1.29e-02	1.02e-02	2.085	7.08e-04	2.814	9.66e-05	3.521	1.44e-05	4.540	2.05e-06	5.503
6.47e-03	2.49e-03	2.051	9.59e-05	2.898	6.56e-06	3.898	5.00e-07	4.864	3.57e-08	5.870
3.24e-03	6.18e-04	2.014	1.24e-05	2.960	4.20e-07	3.972	1.62e-08	4.963	5.76e-10	5.966

## 8.7 Example 7

This example, as the previous one, is also not covered by the current analysis. We are interested in solving the nonhomogeneous Neumann nonlinear problem (90), with  $\Omega := (0, 1)^2$ , and  $\mathbf{a}(\cdot, \zeta) := \left(1 + \frac{1}{1+|\zeta|^2}\right) \zeta$ , for all  $\zeta \in \mathbb{R}^2$ . The data  $f$  and  $g$  are chosen so that the solution  $u$  is given by the smooth function  $u(x, y) = \exp(x + \pi y) - (\exp(1 + \pi) - \exp(\pi) - \exp(1) + 1)/\pi$ . Following what we discussed in Subsection 7.3, we proceed to solve the nonlinear HHO scheme (94), considering one family of Trapezoidal meshes (cf. Figure 17), and another of Tilted hexagonal-dominant meshes (cf. Figure 19). Since the HHO discrete space includes the zero mean value condition of its elements, and knowing that it is not simple to find a basis of such space, we circumvent this searching by imposing this restriction through a Lagrange multiplier. We have applied this strategy in [9], where we have solved the Poisson problem with Neumann boundary condition applying the HHO method. The numerical results are displayed in Figure 12, while the histories of convergence of the potential, flux and reconstructive potential errors, with respect to the nonhomogeneous, are reported in Tables 10 (for the family of Trapezoidal meshes), and 11 (for the family of Tilted meshes). They give numerical evidence that the current a priori error analysis could be extended to this kind of boundary conditions.

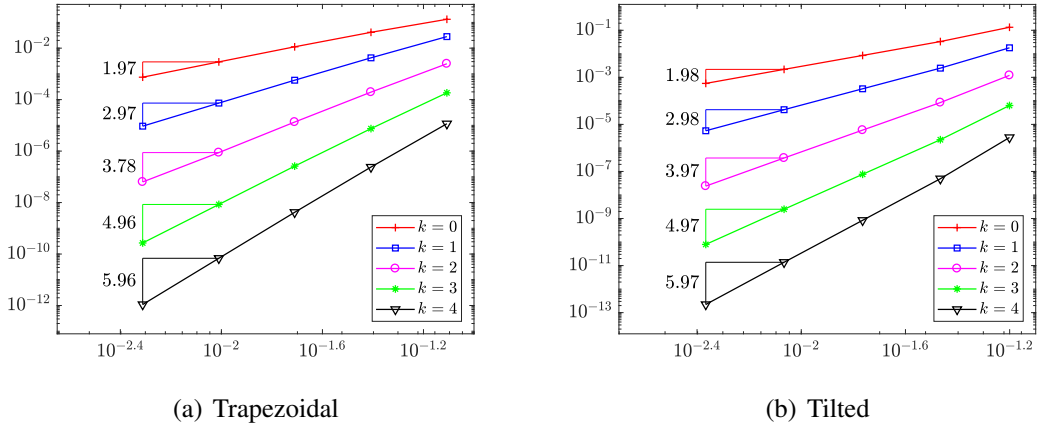
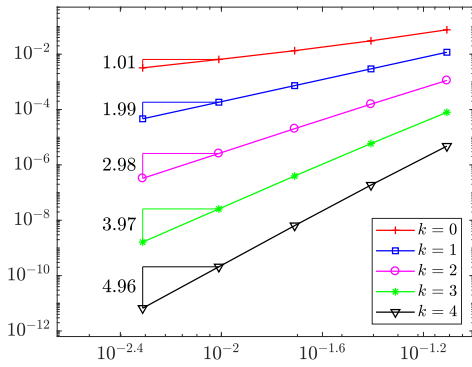


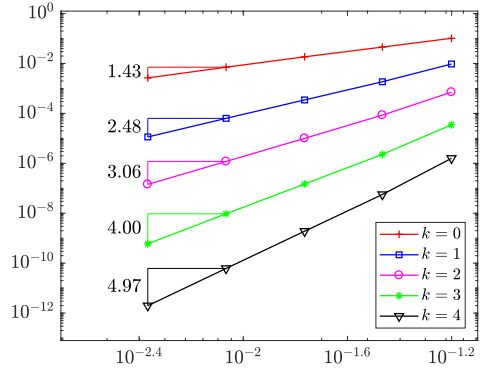
Figure 11: Rates of convergence of potential error vs.  $h$  (Example 7)

Table 10: Histories of convergence of potential, flux, and reconstructive potential errors, considering  $k \in \{0, 1, 2, 3, 4\}$  (Example 7 - Trapezoidal meshes)

Potential										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
6.30e-02	1.32e-01		2.81e-02		2.48e-03		1.84e-04		1.16e-05	
3.42e-02	4.12e-02	1.907	4.20e-03	3.109	1.95e-04	4.156	7.46e-06	5.244	2.39e-07	6.347
1.72e-02	1.13e-02	1.885	5.68e-04	2.914	1.34e-05	3.896	2.60e-07	4.882	4.18e-09	5.890
8.59e-03	2.93e-03	1.943	7.33e-05	2.948	8.74e-07	3.935	8.44e-09	4.938	6.84e-11	5.923
4.30e-03	7.44e-04	1.979	9.30e-06	2.984	6.33e-08	3.795	2.69e-10	4.980	1.09e-12	5.979
Flux										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
6.30e-02	7.61e-02		1.17e-02		1.13e-03		7.94e-05		4.78e-06	
3.42e-02	3.04e-02	1.505	2.94e-03	2.265	1.57e-04	3.229	5.92e-06	4.250	1.87e-07	5.305
1.72e-02	1.34e-02	1.191	7.32e-04	2.025	2.04e-05	2.968	3.98e-07	3.928	6.41e-09	4.907
8.59e-03	6.45e-03	1.051	1.84e-04	1.988	2.59e-06	2.968	2.57e-08	3.946	2.09e-10	4.933
4.30e-03	3.21e-03	1.009	4.63e-05	1.992	3.27e-07	2.991	1.63e-09	3.984	6.66e-12	4.979
Potential reconstructive										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
6.30e-02	1.44e-01		2.84e-02		2.49e-03		1.85e-04		1.16e-05	
3.42e-02	4.36e-02	1.954	4.25e-03	3.108	1.97e-04	4.156	7.50e-06	5.243	2.40e-07	6.347
1.72e-02	1.19e-02	1.895	5.74e-04	2.913	1.35e-05	3.896	2.62e-07	4.882	4.19e-09	5.890
8.59e-03	3.07e-03	1.945	7.42e-05	2.947	8.81e-07	3.934	8.44e-09	4.946	6.87e-11	5.923
4.30e-03	7.81e-04	1.980	9.42e-06	2.983	6.37e-08	3.798	2.71e-10	4.972	1.10e-12	5.980



(a) Trapezoidal



(b) Tilted

Figure 12: Rates of convergence of flux error vs.  $h$  (Example 7)

Table 11: Histories of convergence of potential, flux, and reconstructive potential errors, considering  $k \in \{0, 1, 2, 3, 4\}$  (Example 7 - Tilted meshes)

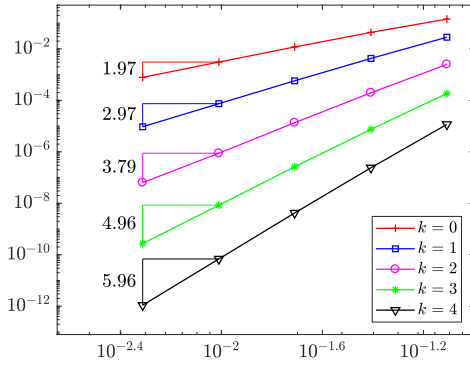
Potential										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
6.30e-02	1.34e-01		1.80e-02		1.21e-03		6.35e-05		1.16e-05	
3.42e-02	3.32e-02	2.287	2.46e-03	3.259	8.48e-05	4.357	2.23e-06	5.480	2.39e-07	6.347
1.72e-02	8.52e-03	1.979	3.26e-04	2.939	5.72e-06	3.924	7.57e-08	4.924	4.18e-09	5.890
8.59e-03	2.18e-03	1.964	4.22e-05	2.947	3.73e-07	3.931	2.48e-09	4.925	6.84e-11	5.923
4.30e-03	5.52e-04	1.984	5.36e-06	2.980	2.39e-08	3.974	7.93e-11	4.973	1.09e-12	5.979

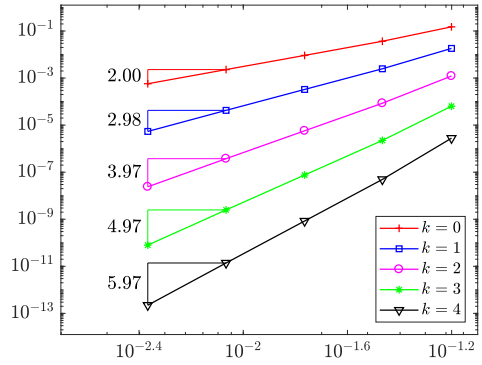
Flux										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
6.30e-02	1.03e-01		9.57e-03		7.12e-04		3.50e-05		4.78e-06	
3.42e-02	4.56e-02	1.327	1.86e-03	2.681	8.55e-05	3.470	2.29e-06	4.460	1.87e-07	5.305
1.72e-02	1.85e-02	1.309	3.48e-04	2.442	1.00e-05	3.115	1.49e-07	3.977	6.41e-09	4.907
8.59e-03	7.13e-03	1.377	6.35e-05	2.449	1.19e-06	3.071	9.49e-09	3.966	2.09e-10	4.933
4.30e-03	2.65e-03	1.428	1.14e-05	2.478	1.44e-07	3.057	5.98e-10	3.995	6.66e-12	4.979

Potential reconstructive										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
6.30e-02	1.50e-01		1.82e-02		1.22e-03		6.37e-05		1.16e-05	
3.42e-02	3.67e-02	2.308	2.48e-03	3.265	8.52e-05	4.358	2.24e-06	5.481	2.40e-07	6.347
1.72e-02	9.16e-03	2.018	3.28e-04	2.941	5.74e-06	3.925	7.59e-08	4.924	4.19e-09	5.890
8.59e-03	2.29e-03	1.995	4.24e-05	2.948	3.75e-07	3.932	2.48e-09	4.925	6.87e-11	5.923
4.30e-03	5.74e-04	2.003	5.39e-06	2.980	2.39e-08	3.974	7.95e-11	4.973	1.10e-12	5.980



(a) Trapezoidal

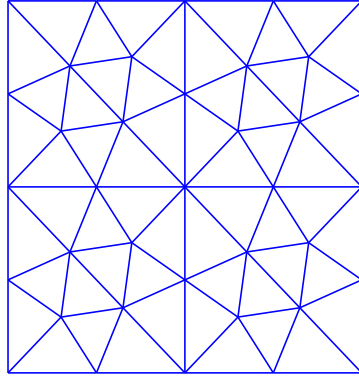


(b) Tilted

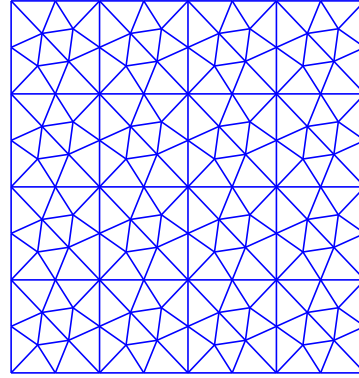
Figure 13: Rates of convergence of reconstructive potential error vs.  $h$  (Example 7)



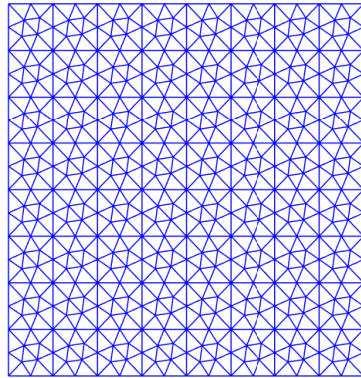
## 8.8 Families of polytopal meshes



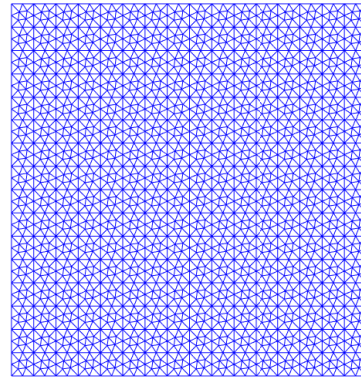
(a) mesh 1



(b) mesh 2

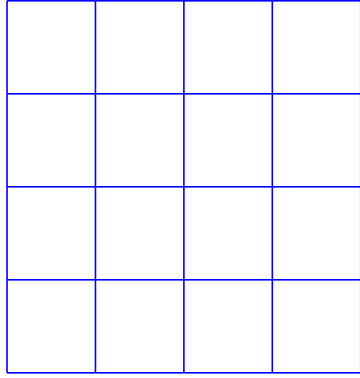


(c) mesh 3

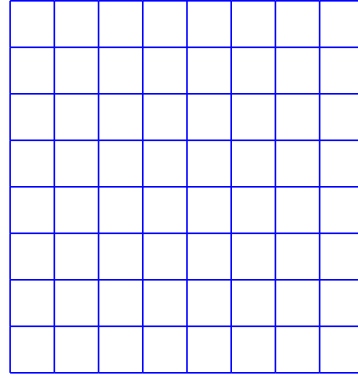


(d) mesh 4

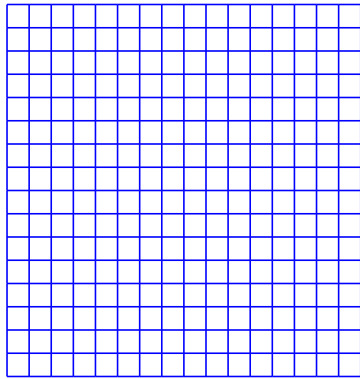
Figure 14: Family of Simplicial-I meshes in a square domain



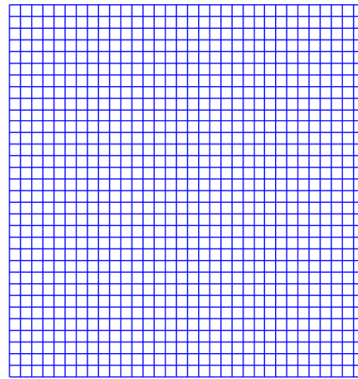
(a) mesh 1



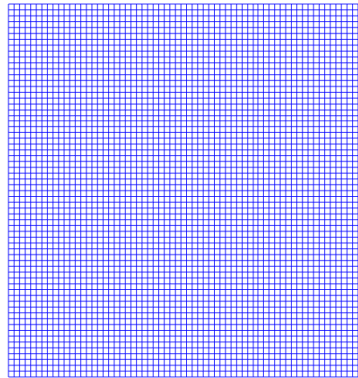
(b) mesh 2



(c) mesh 3

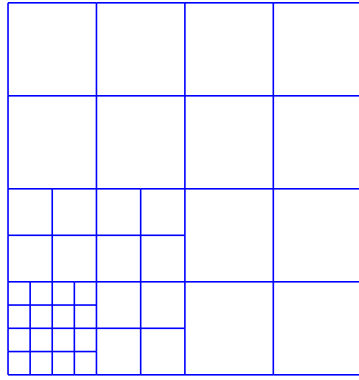


(d) mesh 4

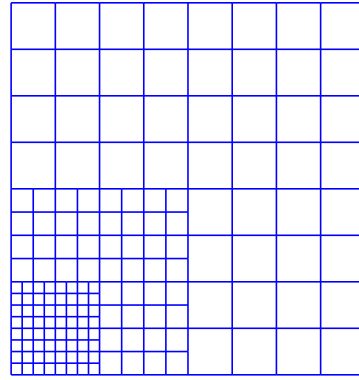


(e) mesh 5

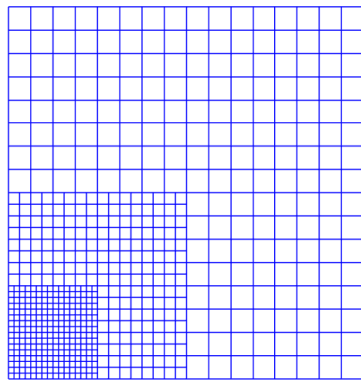
Figure 15: Family of Cartesian meshes in a square domain



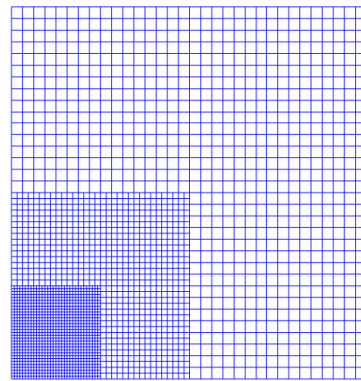
(a) mesh 1



(b) mesh 2

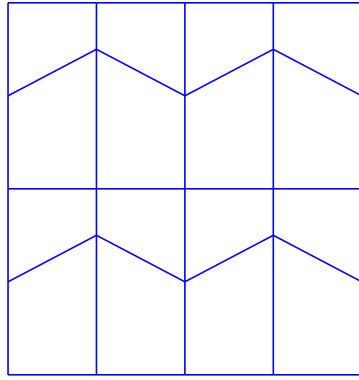


(c) mesh 3

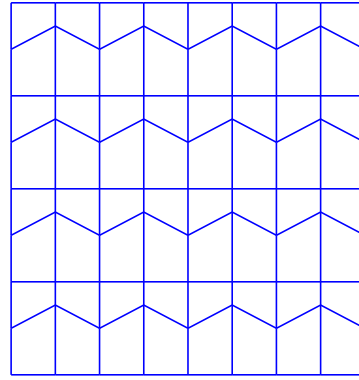


(d) mesh 4

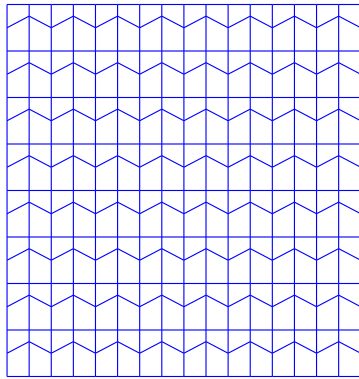
Figure 16: Family of Graduated meshes (with hanging nodes) in a square domain



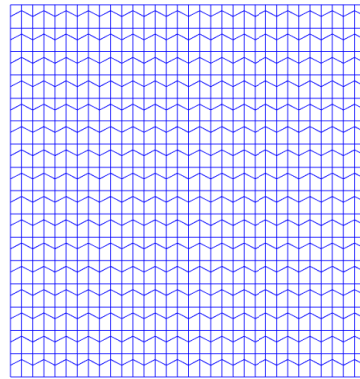
(a) mesh 1



(b) mesh 2

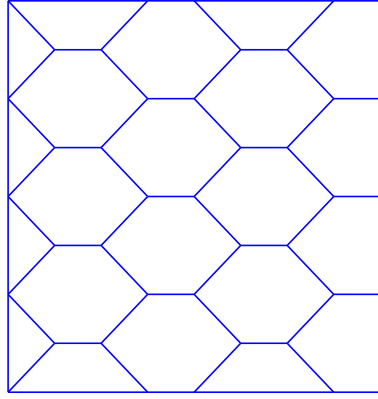


(c) mesh 3

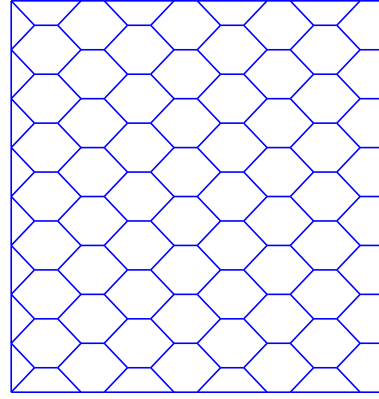


(d) mesh 4

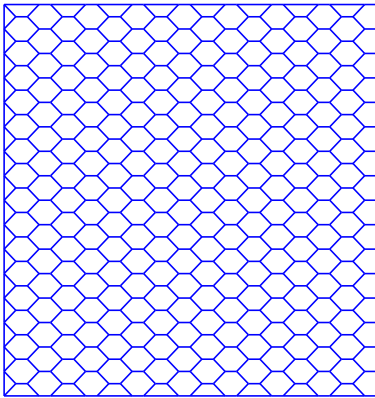
Figure 17: Family of Trapezoidal meshes in a square domain



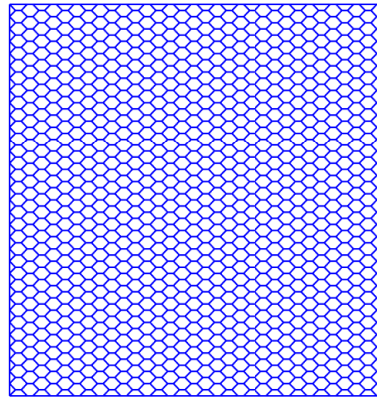
(a) mesh 1



(b) mesh 2

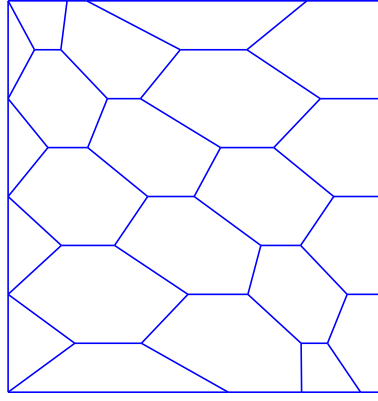


(c) mesh 3

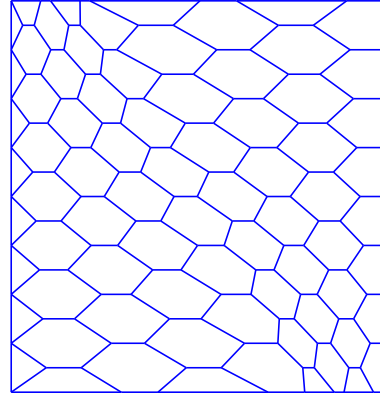


(d) mesh 4

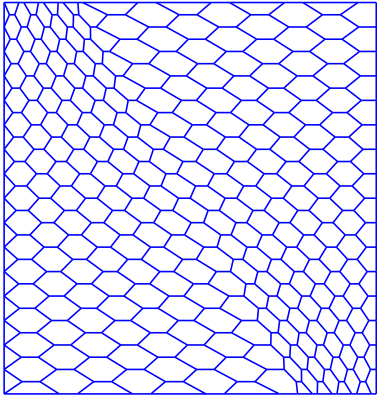
Figure 18: Family of Hexagonal meshes in a square domain



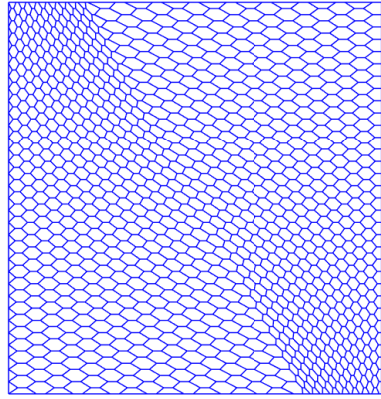
(a) mesh 1



(b) mesh 2

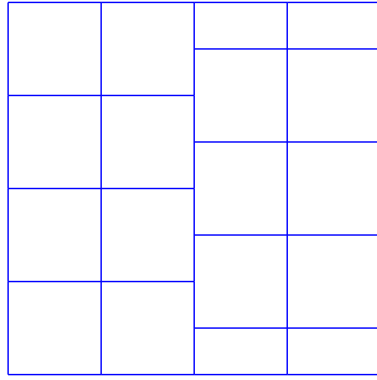


(c) mesh 3

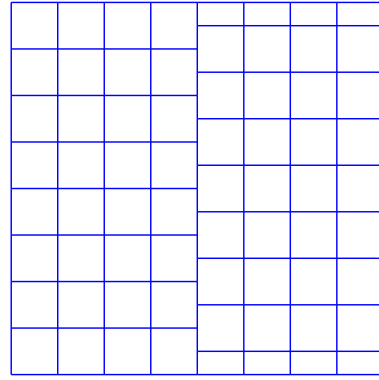


(d) mesh 4

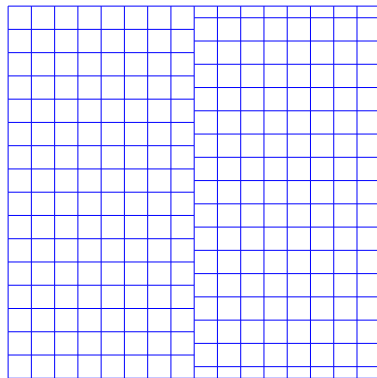
Figure 19: Family of Tilted Hexagonal meshes in a square domain



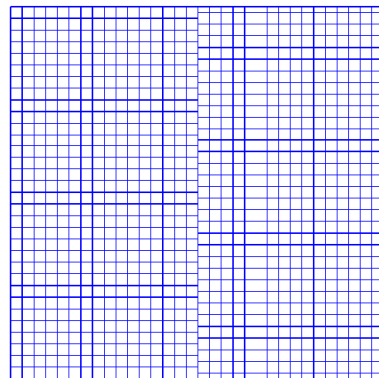
(a) mesh 1



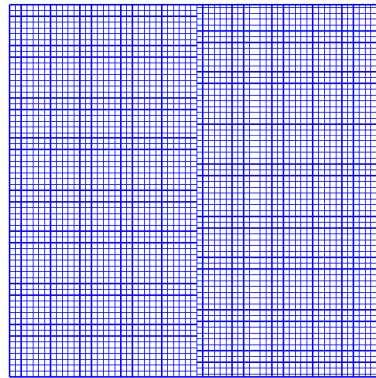
(b) mesh 2



(c) mesh 3

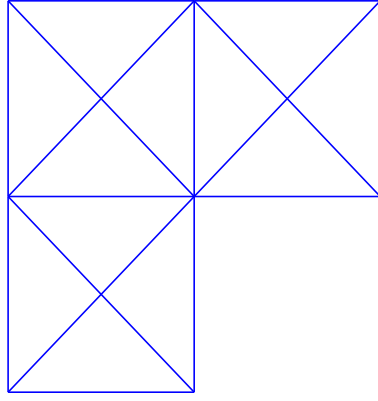


(d) mesh 4

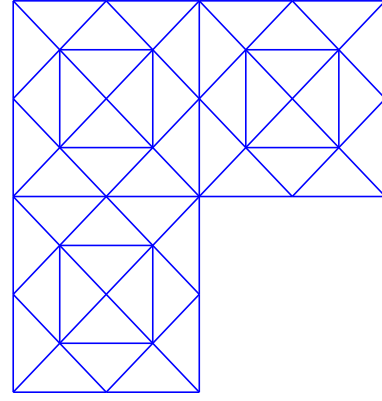


(e) mesh 5

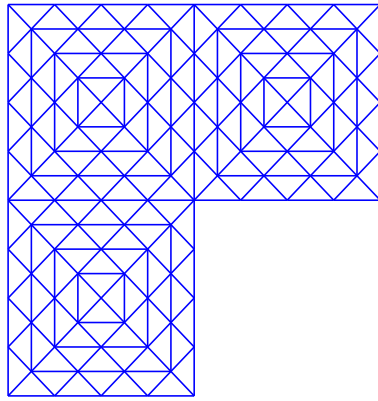
Figure 20: Family of Tilted Fractured meshes in a square domain.



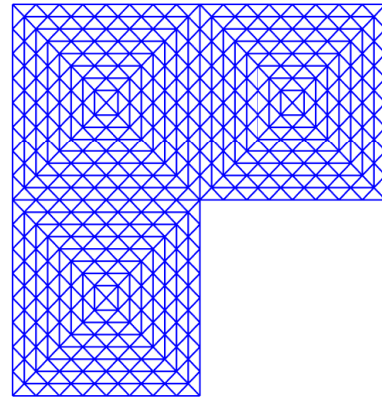
(a) mesh 1



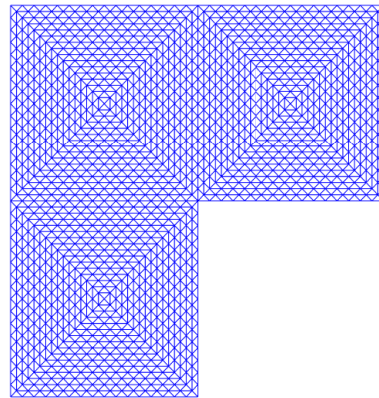
(b) mesh 2



(c) mesh 3



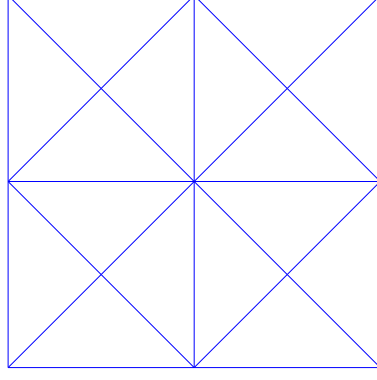
(d) mesh 4



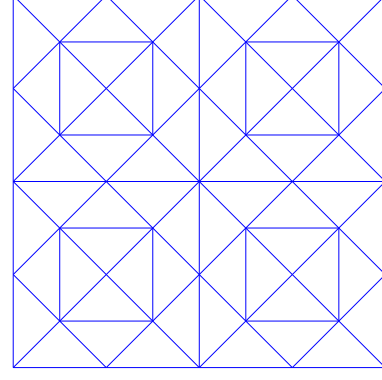
(e) mesh 5

Figure 21: Family of Simplicial-II meshes in  $L$ -shaped domain

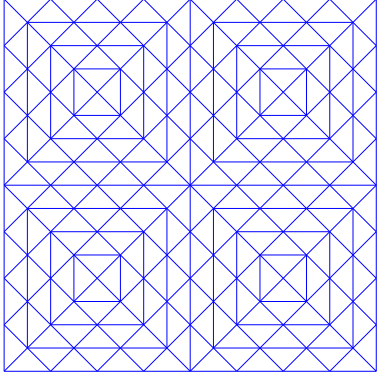




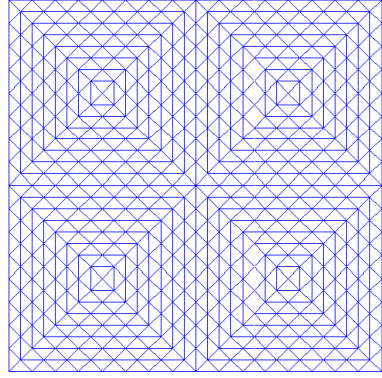
(a) mesh 1



(b) mesh 2



(c) mesh 3



(d) mesh 4

Figure 22: Family of Simplicial-III meshes in Square domain

## Conclusions

In this paper, we have developed a complete a priori error analysis for a class of elliptic non-linear problems, applying the HHO method. It is well known that there are previous works for nonlinear problems in the context of HHO methods, as Leray - Lions equations [13, 14], nonlinear elasticity [6], nonlinear Signorini boundary conditions [11], quasi-Newtonian Stokes problem [49], and nonlinear poroelasticity [5, 7]. However, the contribution of this work relies on Theorems 6.2, and 6.3, which establish the convergence of the  $L^2$ -projection of the potential error ( $||\pi_h^k u - u_h||_{0,\Omega}$ ) and the super-convergence of the reconstructive potential error ( $||p_h^{k+1} \underline{u}_h - u||_{0,\Omega}$ ), respectively. Up to author's knowledge, these errors estimates have not been proven before. It is important to remark that according to what has been discussed in [13, Section 4.1] for nonlinear problems, we need to consider a variant of the gradient reconstruction operator  $G_T^k$ . Otherwise, the rates of convergence will not be the expected. We provide a brief discussion on this in the Appendix 9, which includes a numerical comparison.

We have presented seven computational examples, considering three different nonlinear co-

efficients that verify the Hypotheses **(H.1)**-**(H.4)**, and several admissible families of polytopal meshes, with or without hanging nodes (see Table 1). Numerical examples, provided in this paper, confirm our theoretical results, even for other boundary conditions not covered by us in the present work. In particular, Example 8.5, whose solution is non-smooth, gives us the motivation to develop an a posteriori error analysis for nonlinear elliptic problems. This could be the subject of future work.

## 9 Appendix

Here, we compare the effect of using the standard Gradient reconstruction operator given in [19] and the variant of the Gradient reconstruction operator (with extended codomain, from  $\nabla \mathbb{P}_d^{k+1}(T)$  to  $[\mathbb{P}_d^k(T)]^d$ ) introduced in [13], when solving nonlinear elliptic problems with the HHO method. To clarify the difference, we solve the nonlinear HHO scheme associated to the Example 8.1, with a family of Trapezoidal meshes (cf. Figure 17), considering both kind of Gradient reconstruction operators.

In Figures 23, 24 and 26, we display the rates of convergence of the potential, flux and reconstructive potential errors, with respect to the meshsize. The boxes on the left hand side corresponds to the results obtained when using the Standard gradient reconstruction operator, while the ones on the right hand side are the respective to the Extended gradient reconstruction operator. They are also reported in Tables 12 and 13. As expected, the numerical results associated to the Extended gradient reconstruction operator are in agreement with our theoretical results.

On the other hand, we notice that using the Standard gradient reconstruction operator, the rates of convergence of the potential and reconstructive potential errors, for  $k = 0$  and  $k = 1$ , are the expected ones. For greater values of  $k$ , we observe a notorious loss in the rate of convergence. Respect to the flux error, we only recover the optimal rate of convergence for  $k = 0$  when using the Standard gradient reconstruction operator. Indeed, for  $k = 0$ , both approaches have the same optimal approximation properties, thank to  $\nabla \mathbb{P}^1(T) = [\mathbb{P}^0(T)]^d$ .

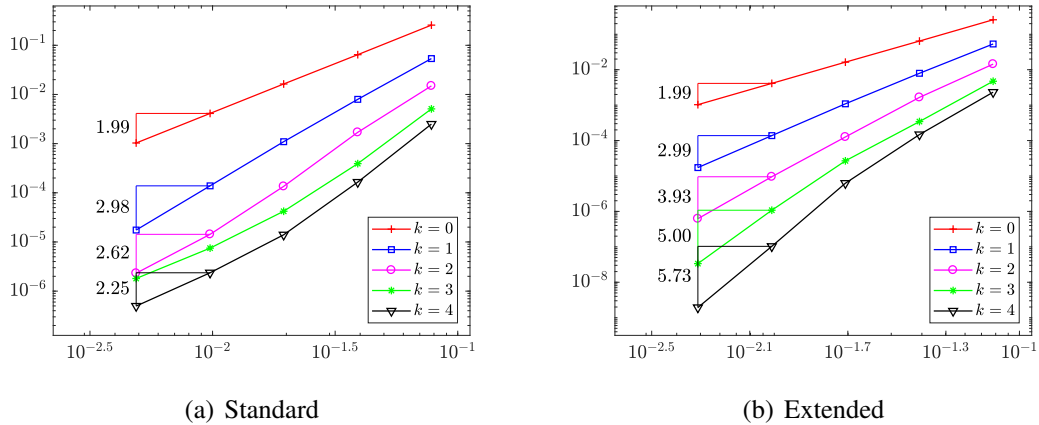


Figure 23: Rates of the Potential error vs.  $h$  (appendix example)

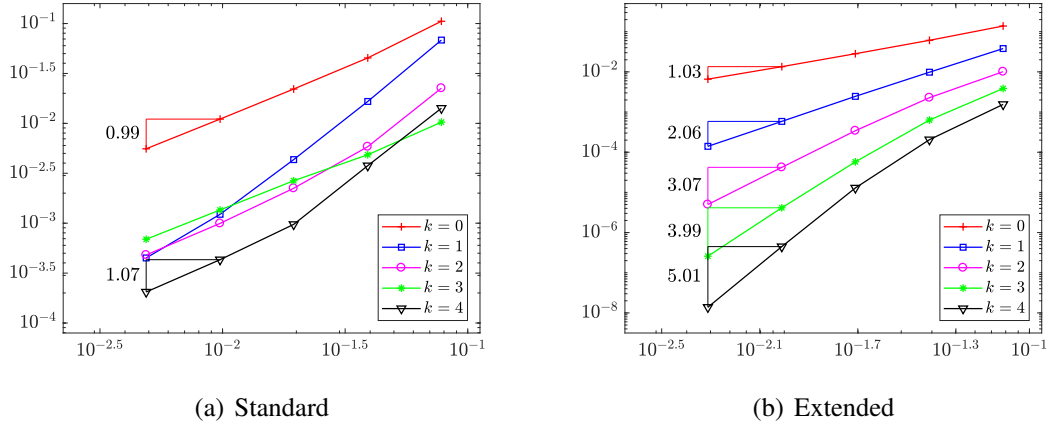


Figure 24: Rates of the Flux error vs.  $h$  (appendix example)

Table 12: Histories of convergence of potential, flux, and reconstructive potential errors, considering  $k \in \{0, 1, 2, 3, 4\}$ , using Standard gradient reconstruction operator (appendix example)

Potential										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
7.81e-02	2.57e-01		5.37e-02		1.49e-02		5.09e-03		2.49e-03	
3.91e-02	6.44e-02	2.001	7.96e-03	2.759	1.70e-03	3.143	3.92e-04	3.706	1.65e-04	3.926
1.95e-02	1.63e-02	1.973	1.09e-03	2.851	1.35e-04	3.638	4.21e-05	3.206	1.40e-05	3.547
9.77e-03	4.11e-03	1.995	1.38e-04	2.992	1.43e-05	3.250	7.48e-06	2.499	2.36e-06	2.574
4.88e-03	1.03e-03	1.994	1.75e-05	2.982	2.31e-06	2.625	1.81e-06	2.048	4.96e-07	2.246
Flux										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
7.81e-02	1.05e-01		6.83e-02		2.25e-02		1.03e-02		1.42e-02	
3.91e-02	4.52e-02	1.223	1.66e-02	2.046	5.84e-03	1.951	4.85e-03	1.087	3.77e-03	1.916
1.95e-02	2.21e-02	1.031	4.33e-03	1.928	2.24e-03	1.378	2.66e-03	0.862	9.73e-04	1.947
9.77e-03	1.10e-02	1.004	1.23e-03	1.826	9.95e-04	1.173	1.36e-03	0.975	4.31e-04	1.178
4.88e-03	5.56e-03	0.986	4.49e-04	1.450	4.81e-04	1.049	6.90e-04	0.973	2.05e-04	1.073
Reconstructive potential										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
7.81e-02	2.74e-01		5.53e-02		1.50e-02		5.10e-03		2.50e-03	
3.91e-02	6.80e-02	2.013	8.12e-03	2.774	1.70e-03	3.146	3.96e-04	3.691	1.68e-04	3.906
1.95e-02	1.71e-02	1.980	1.11e-03	2.856	1.38e-04	3.613	4.49e-05	3.131	1.45e-05	3.523
9.77e-03	4.31e-03	1.998	1.41e-04	2.990	1.56e-05	3.152	8.48e-06	2.412	2.46e-06	2.565
4.88e-03	1.08e-03	1.996	1.79e-05	2.971	2.78e-06	2.485	2.07e-06	2.030	5.21e-07	2.235

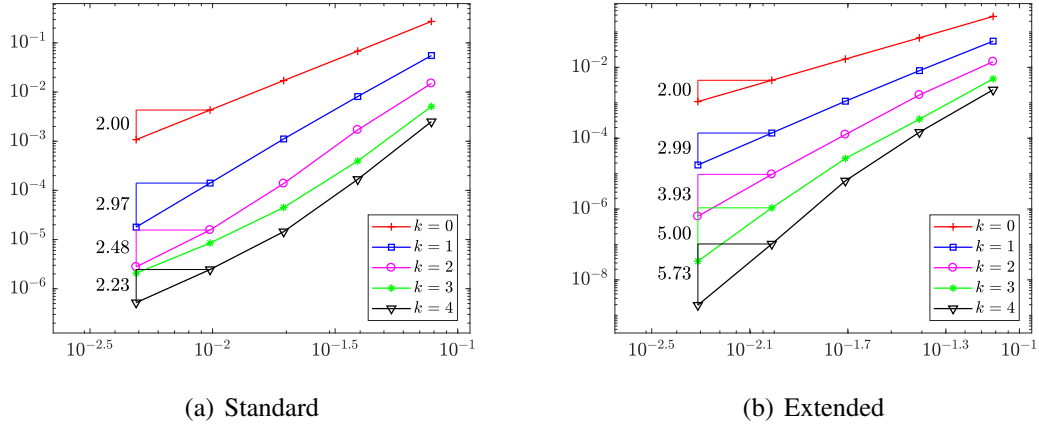


Figure 25: Rates of the Reconstructive potential error vs.  $h$  (appendix example)

Table 13: Histories of convergence of potential, flux, and reconstructive potential errors, considering  $k \in \{0, 1, 2, 3, 4\}$ , using Extended gradient reconstruction operator (appendix example)

Potential										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
7.81e-02	2.57e-01		5.33e-02		1.44e-02		4.73e-03		2.32e-03	
3.91e-02	6.44e-02	2.001	7.93e-03	2.754	1.66e-03	3.126	3.44e-04	3.788	1.48e-04	3.982
1.95e-02	1.63e-02	1.973	1.09e-03	2.852	1.26e-04	3.698	2.68e-05	3.670	6.25e-06	4.546
9.77e-03	4.11e-03	1.995	1.38e-04	2.991	9.50e-06	3.743	1.08e-06	4.645	1.03e-07	5.937
4.88e-03	1.03e-03	1.994	1.73e-05	2.993	6.22e-07	3.928	3.36e-08	5.000	1.93e-09	5.728
Flux										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
7.81e-02	1.38e-01		3.80e-02		1.01e-02		3.85e-03		1.55e-03	
3.91e-02	6.10e-02	1.184	9.81e-03	1.958	2.28e-03	2.156	6.33e-04	2.611	2.05e-04	2.923
1.95e-02	2.83e-02	1.103	2.46e-03	1.990	3.43e-04	2.723	5.76e-05	3.445	1.29e-05	3.976
9.77e-03	1.35e-02	1.073	5.85e-04	2.077	4.20e-05	3.036	4.13e-06	3.812	4.47e-07	4.866
4.88e-03	6.59e-03	1.032	1.40e-04	2.059	4.99e-06	3.069	2.58e-07	3.995	1.39e-08	5.005
Reconstructive potential										
$h$	$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
	error	rate	error	rate	error	rate	error	rate	error	rate
7.81e-02	2.74e-01		5.48e-02		1.45e-02		4.73e-03		2.32e-03	
3.91e-02	6.80e-02	2.013	8.07e-03	2.769	1.66e-03	3.130	3.44e-04	3.788	1.48e-04	3.982
1.95e-02	1.71e-02	1.980	1.11e-03	2.859	1.27e-04	3.698	2.68e-05	3.671	6.25e-06	4.546
9.77e-03	4.31e-03	1.998	1.40e-04	2.993	9.52e-06	3.745	1.08e-06	4.646	1.03e-07	5.937
4.88e-03	1.08e-03	1.996	1.75e-05	2.994	6.23e-07	3.928	3.36e-08	5.000	1.93e-09	5.728

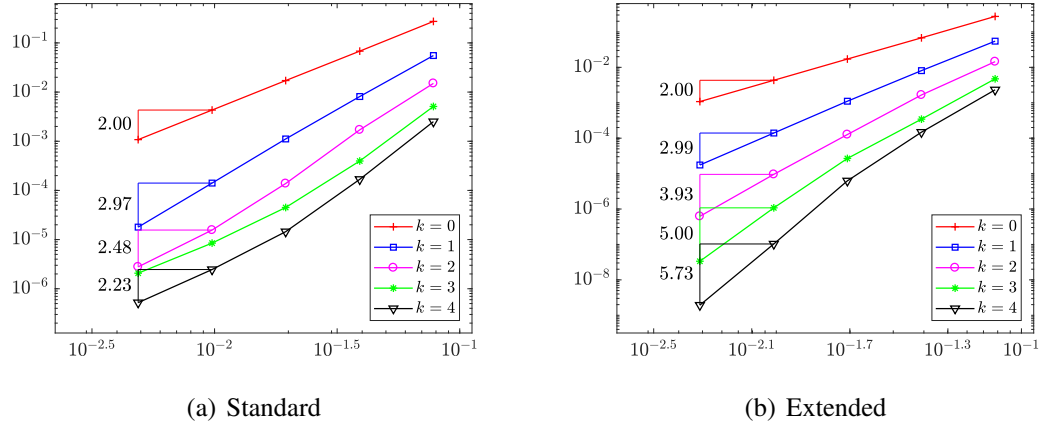


Figure 26: Rates of the reconstructive potential error vs.  $h$  (appendix example)

## References

- [1] AMBROSIO, L. Optimal transport maps in Monge-Kantorovich problem. In *Proceedings of the International Congress of Mathematicians 2002* (2002), T. Li and D. Li, Eds., vol. III, pp. 131–140.
- [2] AYUSO DE DIOS, B., LIPNIKOV, K., AND MANZINI, G. The nonconforming virtual element method. *ESAIM: Mathematical Modelling and Numerical Analysis* 50, 3 (2016), 879–904.
- [3] BARRETT, J., AND LIU, J. G. Finite element approximation of the p-laplacian. *Mathematics of Computation* 61 (1993), 523–537.
- [4] BI, C., AND LIN, Y. Discontinuous galerkin method for monotone nonlinear elliptic problems. *International Journal of Numerical Analysis and Modeling* 9 (2012).
- [5] BOTTI, M., DI PIETRO, D., AND SOCHALA, P. A Hybrid High-Order discretization method for nonlinear poroelasticity. *Computational Methods in Applied Mathematics*. Published online.
- [6] BOTTI, M., DI PIETRO, D., AND SOCHALA, P. A hybrid high-order method for nonlinear elasticity. *SIAM Journal on Numerical Analysis* 55, 6 (2017), 2687–2717.
- [7] BOTTI, M., DI PIETRO, D., AND SOCHALA, P. A nonconforming high-order method for nonlinear poroelasticity. In *Finite Volumes for Complex Applications VIII – Hyperbolic, Elliptic and Parabolic Problems* (2017), C. Cancès and P. Omnes, Eds., pp. 537–545.
- [8] BUSTINZA, R. A unified analysis of the local discontinuous Galerkin method for a class of nonlinear problems. *Applied Numerical Mathematics* 56 (2006), 1293–1306.
- [9] BUSTINZA, R., AND MUNGUÍA-LA-COTERA, J. A hybrid high-order formulation for a Neumann problem on polytopal meshes. *Numerical Methods for Partial Differential Equations* 36, 3 (2020), 524–551.

- [10] CANGIANI, A., MANZINI, G., AND SUTTON, O. Conforming and nonconforming virtual element methods for elliptic problems. *IMA Journal of Numerical Analysis* 37, 3 (2017), 1317–1354.
- [11] CASCAVITA, K., CHOULY, F., AND ERN, A. Hybrid High-Order discretizations combined with Nitsche’s method for Dirichlet and Signorini boundary conditions. *IMA Journal of Numerical Analysis* (2019).
- [12] COCKBURN, B., DI PIETRO, D., AND ERN, A. Bridging the Hybrid High-Order and Hybridizable Discontinuous Galerkin Methods. *ESAIM: Mathematical Modelling and Numerical Analysis* 50 (2015).
- [13] DI PIETRO, D., AND DRONIOU, J. A Hybrid High-Order method for Leray–Lions elliptic equations on general meshes. *Mathematics of Computation* 86, 307 (2017), 2159–2191.
- [14] DI PIETRO, D., AND DRONIOU, J.  $W^{s,p}$ -approximation properties of elliptic projectors on polynomial spaces, with application to the error analysis of a Hybrid High-Order discretisation of Leray-Lions problems. *Mathematical Models and Methods in Applied Sciences* 27, 5 (2017), 879–908.
- [15] DI PIETRO, D., DRONIOU, J., AND ERN, A. A discontinuous-skeletal method for advection-diffusion-reaction on general meshes. *SIAM Journal on Numerical Analysis* 53 (2015), 2135 – 2157.
- [16] DI PIETRO, D., DRONIOU, J., AND MANZINI, G. Discontinuous skeletal gradient discretisation methods on polytopal meshes. *Journal of Computational Physics* 355 (2017).
- [17] DI PIETRO, D., AND ERN, A. *Mathematical Aspects of Discontinuous Galerkin Methods*, vol. 69 of *Mathématiques et Applications*. Springer-Verlag, Berlin, 2012.
- [18] DI PIETRO, D., AND ERN, A. A hybrid high-order locking-free method for linear elasticity on general meshes. *Computer Methods in Applied Mechanics and Engineering* 283 (2014), 1–21.
- [19] DI PIETRO, D., ERN, A., AND LEMAIRE, S. An Arbitrary-Order and Compact-Stencil Discretization of Diffusion on General Meshes Based on Local Reconstruction Operators. *Computational Methods in Applied Mathematics* 14, 4 (2014), 461–472.
- [20] DI PIETRO, D. A., AND DRONIOU, J. *The Hybrid High-Order Method for Polytopal Meshes: Design, Analysis, and Applications.*, vol. 19 of *Modeling, Simulation and Applications series*. Springer International Publishing, 2020. 528 pages.
- [21] DI PIETRO, D. A., AND ERN, A. Arbitrary-order mixed methods for heterogeneous anisotropic diffusion on general meshes. *IMA Journal of Numerical Analysis* 37, 1 (2017), 40–63.
- [22] DIAZ, J. I., AND THELIN, F. On a nonlinear parabolic problem arising in some models related to turbulent flows. *SIAM Journal on Mathematical Analysis* 25 (1994).

- [23] DRONIOU, J. Finite volume schemes for fully non-linear elliptic equations in divergence form. *Esaim-Mathematical Modelling and Numerical Analysis-Modelisation Mathématique et Analyse Numerique* 40 (2006), 1069–1100.
- [24] DRONIOU, J., AND EYMARD, R. A mixed finite volume scheme for anisotropic diffusion problems on any grid. *Numerische Mathematik* 105, 1 (2006), 35–71.
- [25] DRONIOU, J., EYMARD, R., GALLOUET, T., AND HERBIN, R. A unified approach to mimetic finite difference, hybrid finite volume and mixed finite volume methods. *Mathematical Models and Methods in Applied Sciences* 20, 2 (2010), 265 – 295.
- [26] EYMARD, R., GALLOUËT, T., AND HERBIN, R. Discretization of heterogeneous and anisotropic diffusion problems on general nonconforming meshes SUSHI: a scheme using stabilization and hybrid interfaces. *IMA Journal of Numerical Analysis* 30, 4 (2010), 1009–1043.
- [27] FEISTAUER, M., AND ENOK, A. Finite element solution of nonlinear elliptic problems. *Numerische Mathematik* 50 (1986), 451 – 475.
- [28] FIGALLI, A. *The Monge-Ampère Equation and Its Applications*, vol. 22 of *Zurich lectures in advanced mathematics*. European Mathematical Society, 2017.
- [29] FRIEDMAN, A. The stefan problem in several space variables. *Transactions of The American Mathematical Society - TRANS AMER MATH SOC* 133 (1968).
- [30] GATICA, G. N., HEUER, N., AND MEDDAHI, S. On the numerical analysis of nonlinear twofold saddle point problems. *IMA J. Numer. Anal.* 23, 2 (2003), 301–330.
- [31] GATICA, G. N., MARQUEZ, A., AND RUDOLPH, W. A priori and a posteriori error analyses of augmented twofold saddle point formulations for nonlinear elasticity problems. *Computer Methods in Applied Mechanics and Engineering* 264 (2013), 23 – 48.
- [32] GATICA, G. N., AND MEDDAHI, S. A dual-dual mixed formulation for nonlinear exterior transmission problems. *Math. Comp.* 70, 236 (2001), 1461–1480.
- [33] GATICA, G. N., AND STEPHAN, E. A mixed-fem formulation for nonlinear incompressible elasticity in the plane. *Numerical Methods for Partial Differential Equations* 18 (2002), 105 – 128.
- [34] GLOWINSKI, R. *Numerical Methods for Nonlinear Variational Problems*. Springer Series in Computational Physics. Springer-Verlag, New York, 1984.
- [35] GLOWINSKI, R., AND RAPPAZ, J. Approximation of a nonlinear elliptic problem arising in a non-newtonian fluid flow model in glaciology. *ESAIM: Mathematical Modelling and Numerical Analysis* 37 (2003).
- [36] GRISVARD, P. *Elliptic problems in nonsmooth domains*. Classics in Applied Mathematics. SIAM, USA, 2011. Reprint of the 1983 edition.

- [37] HERBIN, R., AND HUBERT, F. Benchmark on discretization schemes for anisotropic diffusion problems on general grids. In *Finite Volumes for Complex Applications V*, R. Eyraud and J.-M. Hérard, Eds. John Wiley & Sons, 2008, pp. 659–692.
- [38] LEMAIRE, S. Bridging the hybrid high-order and virtual element methods. *IMA Journal of Numerical Analysis* (Published online).
- [39] LIPNIKOV, K., AND MANZINI, G. A high-order mimetic method on unstructured polyhedral meshes for the diffusion equation. *Journal of Computational Physics* 272 (2014), 360 – 385.
- [40] MEDJO, T. Mixed formulation of the two-layer quasi-geostrophic equations of the ocean. *Numerical Methods for Partial Differential Equations* 15 (1999), 489–502.
- [41] NEČAS, J. *Introduction to the theory of nonlinear elliptic equations*. A Wiley-Interscience Publication. John Wiley & Sons, Ltd., Chichester, 1986. Reprint of the 1983 edition.
- [42] ORTNER, C., AND SLI, E. Discontinuous galerkin finite element approximation of nonlinear second-order elliptic and hyperbolic systems. *SIAM J. Numerical Analysis* 45 (2007), 1370 – 1397.
- [43] RUESCHENDORF, L. Monge-Kantorovich transportation problem and optimal couplings. *Jahresbericht der Deutschen Mathematiker-Vereinigung (DMV)* 109 (2007).
- [44] VEIGA, B., DRONIOU, J., AND MANZINI, G. A unified approach to handle convection terms in finite volumes and mimetic discretization methods for elliptic problems. *IMA J. Numer. Anal.* 31 (2011), 1357–1401.
- [45] VEIGA, L., BREZZI, F., CANGIANI, A., MANZINI, G., MARINI, L., AND RUSSO, A. Basic principles of virtual element methods. *Mathematical Models and Methods in Applied Sciences* 23 (2012).
- [46] VEIGA, L., LIPNIKOV, K., AND MANZINI, G. *The Mimetic Finite Difference Method for Elliptic Problems*, vol. 11 of *Modeling, Simulation and Applications*. Springer International Publishing, 2014.
- [47] VEIGA, L., LOVADINA, C., AND MORA, D. A virtual element method for elastic and inelastic problems on polytope meshes. *Computer Methods in Applied Mechanics and Engineering* 295 (2015), 327 – 346.
- [48] ŽENÍŠEK, A. *Nonlinear elliptic and evolution problems and their finite element approximations*. Computational Mathematics and Applications. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], London, 1990. With a foreword by P.-A. Raviart.
- [49] ZHANG, Y., AND MEI, L. A hybrid high-order method for a class of quasi-newtonian stokes equations on general meshes. *Applied Mathematics and Computation* (Published online).



# Centro de Investigación en Ingeniería Matemática (CI<sup>2</sup>MA)

## PRE-PUBLICACIONES 2019 - 2020

- 2019-45 MAURICIO MUNAR, FILANDER A. SEQUEIRA: *A posteriori error analysis of a mixed virtual element method for a nonlinear Brinkman model of porous media flow*
- 2019-46 RAIMUND BÜRGER, STEFAN DIEHL, MARÍA CARMEN MARTÍ, YOLANDA VÁSQUEZ: *On dynamic models of flotation with sedimentation*
- 2019-47 RAIMUND BÜRGER, SARVESH KUMAR, DAVID MORA, RICARDO RUIZ-BAIER, NITESH VERMA: *Virtual element methods for the three-field formulation of time-dependent linear poroelasticity*
- 2019-48 GABRIEL R. BARRENECHEA, FABRICE JAILLET, DIEGO PAREDES, FREDERIC VALENTIN: *The multiscale hybrid mixed method in general polygonal meshes*
- 2020-01 SERGIO CAUCAO, GABRIEL N. GATICA, RICARDO OYARZÚA, NESTOR SÁNCHEZ: *A fully-mixed formulation for the steady double-diffusive convection system based upon Brinkman–Forchheimer equations*
- 2020-02 GONZALO A. BENAVIDES, LEONARDO E. FIGUEROA: *Orthogonal polynomial projection error in Dunkl-Sobolev norms in the ball*
- 2020-03 RODOLFO ARAYA, ABNER POZA, FREDERIC VALENTIN: *An adaptative multiscale hybrid-mixed method for the Oseen equations*
- 2020-04 CARLOS PARÉS, DAVID ZORÍO: *Lax Wendroff approximate Taylor methods with fast and optimized weighted essentially non-oscillatory reconstructions*
- 2020-05 JULIO ARACENA, LILIAN SALINAS: *Finding the fixed points of a Boolean network from a positive feedback vertex set*
- 2020-06 TOMÁS BARRIOS, ROMMEL BUSTINZA, CAMILA CAMPOS: *A note on a posteriori error estimates for dual mixed methods*
- 2020-07 RAIMUND BÜRGER, ELVIS GAVILÁN, DANIEL INZUNZA, PEP MULET, LUIS M. VILLADA: *Numerical simulation of forest fires by IMEX methods*
- 2020-08 ROMMEL BUSTINZA, JONATHAN MUNGUÍA: *An a priori error analysis for a class of nonlinear elliptic problems applying the hybrid high-order method*

Para obtener copias de las Pre-Publicaciones, escribir o llamar a: DIRECTOR, CENTRO DE INVESTIGACIÓN EN INGENIERÍA MATEMÁTICA, UNIVERSIDAD DE CONCEPCIÓN, CASILLA 160-C, CONCEPCIÓN, CHILE, TEL.: 41-2661324, o bien, visitar la página web del centro: <http://www.ci2ma.udec.cl>



**CENTRO DE INVESTIGACIÓN EN  
INGENIERÍA MATEMÁTICA (CI<sup>2</sup>MA)  
Universidad de Concepción**



Casilla 160-C, Concepción, Chile  
Tel.: 56-41-2661324/2661554/2661316  
<http://www.ci2ma.udec.cl>

