



**HAL**  
open science

# Families of hybridizable interior penalty discontinuous Galerkin methods for degenerate advection-diffusion-reaction problems

G. Etangsale, M. Fahs, V. Fontaine, A. R. Isa-Abadi

► **To cite this version:**

G. Etangsale, M. Fahs, V. Fontaine, A. R. Isa-Abadi. Families of hybridizable interior penalty discontinuous Galerkin methods for degenerate advection-diffusion-reaction problems. 2021. hal-03247309

**HAL Id: hal-03247309**

**<https://hal.science/hal-03247309>**

Preprint submitted on 3 Jun 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Families of hybridizable interior penalty discontinuous Galerkin methods for degenerate advection-diffusion-reaction problems

Grégory Etangsale<sup>a</sup>, Marwan Fahs<sup>b</sup>, Vincent Fontaine<sup>a,\*</sup>, Ali Raeisi Isa-Abadi<sup>c</sup>

<sup>a</sup>*Department of Building and Environmental Sciences, University of La Réunion - South Campus, France*

<sup>b</sup>*Université de Strasbourg, CNRS, ENGEES, LHYGES UMR 7517, F-67000 Strasbourg, France*

<sup>c</sup>*Department of Water Engineering, Shahrekord University, Shahrekord, Iran*

---

## Abstract

We analyze families of primal high-order hybridizable discontinuous Galerkin (HDG) methods for solving degenerate (second-order) elliptic problems. One major trouble regarding this class of PDEs concerns its mathematical nature, which may be nonuniform over the domain. Due to the local degeneracy of the diffusion term, it can be purely hyperbolic in a subregion and elliptic in the rest. This problem is thus quite delicate to solve since the exact solution is discontinuous at interfaces separating both elliptic and hyperbolic parts. The proposed HDG method is developed in a unified and compact fashion. It can efficiently handle pure diffusive or advective regimes and intermediate regimes that combine the above mechanisms for a wide range of Péclet numbers, including the delicate situation of local evanescent diffusion. To this end, an adaptive stabilization strategy based on addition of jump-penalty terms is then considered. A  $\theta$ -upwind-based scheme is favored for the hyperbolic region, and an inspired Scharfetter–Gummel-based technique is preferred for the elliptic region. The well-posedness of the HDG method is also discussed by analyzing the consistency and discrete coercivity properties. Extensive numerical experiments are finally considered to verify the model’s robustness for all the abovementioned regimes.

*Keywords:* Primal hybridizable discontinuous Galerkin, interior penalty methods, degenerate second-order elliptic problems, adaptive penalty strategy, upwind-based scheme, Scharfetter–Gummel scheme, extensive numerical experiments

---

## 1. Introduction

Degenerate second-order elliptic equations are well-established models to describe a wide variety of phenomena in real-life applications, such as pure diffusion or advection problems, and mixed problems combining the above mechanisms for a wide range of Péclet numbers [24]. A detrimental situation may arise in the context of locally evanescent diffusivity. Indeed, its mathematical nature is nonuniform over the entire domain, as it can be purely hyperbolic in a subregion and elliptic in the rest. Consequently, the state variable can be discontinuous at interfaces separating both subregions according to the wind flow sense [9, 10, 14]. This critical

---

\*Corresponding and Principal author

*Email addresses:* `gregory.etangsale@univ-reunion` (Grégory Etangsale), `fahs@unistra.fr` (Marwan Fahs), `vincent.fontaine@univ-reunion` (Vincent Fontaine), `ali_raeisi@sku.ac.ir` (Ali Raeisi Isa-Abadi)

situation is easily encountered in the context of mass transport in fractured porous media. It is well known that fractures deeply affect the transport phenomena since they represent the preferential fluid flow paths. The large variability of the velocity’s magnitude indicates that the advection mechanism is predominant in the fractures as compared with the rest of the domain. During recent decades, different authors have analyzed this model problem, although mainly in the context of discontinuous Galerkin (DG) methods (see, e.g., [10, 14, 17] and the extensive references therein). The success story of DG methods is because they combine advantages of finite volume and finite element methods, and they are well-suited to capture large gradients or discontinuities of exact solutions [25]. Despite all of these advantages, DG methods are generally more expensive than most other numerical methods due to their high number of coupled degrees of freedom (DOFs) and their large stencils. It is in this context that the HDG methods were initially devised: as a way to circumvent these drawbacks.

The HDG methods were first introduced by Cockburn *et al.* in [7] and have been applied successfully to various physical problems [4, 8, 12, 22, 21, 20, 23, 19]. They can be considered as a new class of DG methods that are eligible for static condensation. For that application, an additional discrete variable is introduced corresponding to the trace approximation of the state variable on the mesh skeleton. Thus, the interior-based DOFs can be easily eliminated by solving a local problem at the element level so that only skeleton-based DOFs remain. The problem is then closed by (weakly) imposing transmission conditions on the mesh skeleton, leading to a smaller and sparser final matrix system. In practice, DG methods and their HDG counterparts do not coincide because the latter use a richer definition of numerical traces. Consequently, HDG methods turn out to be more accurate than their predecessors in many situations, and they are thus more efficiently implementable and highly parallelizable [15, 18]. Despite all of these assets, the literature is relatively scarce concerning the resolution of degenerate elliptic equations by the class of HDG methods, which is the purpose of the present work. To the best of our knowledge, only Di Pietro *et al.* recently designed a primal discontinuous skeletal method based on the hybrid-high order (HHO) formalism for the diffusive part to address this kind of issue [9, 11].

In the present paper, we focus instead on the class of interior penalty HDG methods denoted by H-IP and its three well-known variants, namely, the incomplete (H-IIP), symmetric (H-SIP), and nonsymmetric (H-NIP) schemes [16, 28]. Indeed, the families of interior penalty methods are well-suited for solving degenerate elliptic equations since the diffusion term might not be invertible at every point of the domain; i.e., they belong to the class of  $\kappa$ -methods. For this aim, both diffusive and advective-reactive contributions are discretized separately. The stability of these contributions is ensured by adding jump-penalty terms, which correspond to the discrepancy between interior- and interface-based DOFs, on the mesh skeleton. The stabilization penalty parameters are selected in accordance with the nature of the local cellwise problem reducing to an upwinding-based scheme in the hyperbolic subregion [3] and the Scharfetter–Gummel (SG) scheme elsewhere [2, 9]. Thus, the stated H-IP formalism can treat in an automatic fashion the pure diffusion or advection-reaction processes or (mixed) advection-diffusion-reaction processes characterized by a diffusion- or advection-dominated regime, i.e., a wide range of Péclet numbers - including the delicate situation of local evanescent diffusion. A stability analysis is then investigated by establishing the consistency and discrete coercivity properties. Numerical experiments are also presented to prove the following assertions, such as the high-order accuracy and robustness of the discretization method.

The material is organized as follows. In Section 2, we describe the homogeneous Dirichlet

boundary value problem in the sense of Fichera [24] by introducing some specific notations, and we precisely define the corresponding discrete setting in Section 3. In Section 4, we derive the discrete bilinear and linear operators of the discretization method, briefly discuss the static condensation and stability analysis, and precisely delineate the stabilization strategy. In Section 5, extensive numerical experiments are investigated using  $h$ - and  $p$ -refinement strategies for all abovementioned regimes. We end with some concluding remarks and perspectives.

## 2. Boundary value problem

We consider the stationary linear advection-diffusion-reaction model problem in its conservative form,

$$\nabla \cdot (-\boldsymbol{\kappa} \nabla u + \boldsymbol{\beta} u) + \gamma u = f \quad \text{in } \Omega, \quad (1)$$

where  $\Omega \subset \mathbb{R}^d$  is a bounded polyhedral domain ( $d \geq 2$ ) with boundary  $\partial\Omega \subset \mathbb{R}^{d-1}$ . The boldface fonts are used throughout the paper to characterize any vector- or matrix-valued functions. In the physical context of contaminant transport in a porous media,  $\boldsymbol{\kappa} : \Omega \rightarrow \mathbb{R}^{d,d}$  represents an anisotropic heterogeneous dispersion tensor, which is itself a function of the Darcy velocity field  $\boldsymbol{\beta} : \Omega \rightarrow \mathbb{R}^d$ ,  $\gamma : \Omega \rightarrow \mathbb{R}^+$ , the reaction coefficient, and  $f : \Omega \rightarrow \mathbb{R}$ , a forcing term. We then assume that the constitutive coefficients of (1) satisfy the following minimal regularity requirements:

- \*  $\boldsymbol{\kappa} \in [L^\infty(\Omega)]^{d,d}$  is a symmetric positive semidefinite matrix-valued function verifying that

$$\underline{\kappa} \|\boldsymbol{\zeta}\|^2 \leq \boldsymbol{\zeta}^t \boldsymbol{\kappa}(x) \boldsymbol{\zeta} \leq \bar{\kappa} \|\boldsymbol{\zeta}\|^2, \quad \forall \boldsymbol{\zeta} \in \mathbb{R}^d, \quad \text{a.e. } x \in \Omega. \quad (2)$$

where  $\bar{\kappa} \geq \underline{\kappa} \geq 0$  denote the largest and smallest eigenvalues of  $\boldsymbol{\kappa}$ , respectively. Under the hypothesis (2), we shall assume the existence of a subdomain  $\Omega^{\text{ell}}$  (resp.  $\Omega^{\text{hyp}}$ ) corresponding to the elliptic (resp. hyperbolic) region such that

$$\begin{aligned} \Omega^{\text{ell}} &:= \{x \in \Omega : \boldsymbol{\zeta}^t \boldsymbol{\kappa}(x) \boldsymbol{\zeta} > 0, \quad \forall \boldsymbol{\zeta} \in \mathbb{R}^d\}, \\ \Omega^{\text{hyp}} &:= \{x \in \Omega : \boldsymbol{\zeta}^t \boldsymbol{\kappa}(x) \boldsymbol{\zeta} = 0, \quad \forall \boldsymbol{\zeta} \in \mathbb{R}^d\}, \end{aligned} \quad (3)$$

verifying that  $\Omega = \Omega^{\text{ell}} \cup \Omega^{\text{hyp}}$  and  $\Omega^{\text{ell}} \cap \Omega^{\text{hyp}} = \emptyset$  (nonoverlapping subregions).

- \*  $\boldsymbol{\beta} \in [L^\infty(\Omega)]^d$  is s.t.  $\nabla \cdot \boldsymbol{\beta} \in L^\infty(\Omega)$ ,

- \*  $\gamma \in L^\infty(\Omega)$  verifying that the following standard coercivity condition holds

$$\exists \gamma_0 \in \mathbb{R} \quad \text{s.t.} \quad \gamma(x) + \frac{1}{2} \nabla \cdot \boldsymbol{\beta}(x) \geq \gamma_0 > 0, \quad \text{a.e. } x \in \Omega. \quad (4)$$

Let us now introduce the following disjoint boundary sets as defined by Oleřnik and Radkevič [24] (see, e.g., Di Pietro *et al.* [9] and references therein);

$$\begin{aligned} \Gamma^- &:= \{x \in \partial\Omega : \boldsymbol{n}^t \boldsymbol{\kappa}(x) \boldsymbol{n} > 0 \quad \text{or} \quad \boldsymbol{\beta} \cdot \boldsymbol{n} < 0\}, \\ \Gamma^+ &:= \{x \in \partial\Omega : \boldsymbol{n}^t \boldsymbol{\kappa}(x) \boldsymbol{n} = 0 \quad \text{and} \quad \boldsymbol{\beta} \cdot \boldsymbol{n} \geq 0\}, \end{aligned} \quad (5)$$

where  $\boldsymbol{n}$  denotes the unit outward normal to  $\partial\Omega$ . Thus, the disjoint subsets  $\Gamma^\pm$  will be referred to as the *nondegenerate inflow* and *degenerate outflow/no-flow* parts of the boundary  $\partial\Omega$ , respectively, verifying that  $\partial\Omega = \Gamma^- \cup \Gamma^+$  and  $\Gamma^- \cap \Gamma^+ = \emptyset$ . For clarity of our exposition,

we supplement the partial differential equation (1) with the following homogeneous Dirichlet boundary conditions, namely,

$$u = 0 \quad \text{on} \quad \Gamma^-. \quad (6)$$

Let us remark that the Dirichlet boundary conditions are only prescribed for portions of  $\partial\Omega$  touching the elliptic region or the hyperbolic region, provided that the advective field flows into the domain. Let us now summarize some physical situations commonly encountered in the literature for which Dirichlet boundary conditions (6) adapt automatically:

- ▶ Nondegenerate problems. Here,  $\kappa$  is assumed to be a symmetric positive definite matrix-valued function on the whole domain  $\Omega$ . Hence,  $\Omega^{\text{ell}} = \Omega$  and the Dirichlet boundary condition is automatically enforced for the whole boundary  $\Gamma^- = \partial\Omega$ . This situation includes the pure diffusive regime as well as the mixed advective-diffusive regime. Here, the weak solution  $u$  of the problem (1) respects some regularity requirements i.e.,  $u \in H^1(\Omega)$ , but it may present some sharp fronts along the characteristic direction  $\beta$ , particularly for the advective-dominated regime.
- ▶ Fully degenerate problems. Here,  $\Omega^{\text{ell}} = \emptyset$  and we recover the standard advection-reaction problem defined in  $\Omega^{\text{hyp}}$  respecting the coercivity condition (4) and the usual definition of the inflow  $\Gamma^-$  or outflow  $\Gamma^+$  parts of the boundary  $\partial\Omega$ . Following these assumptions, the problem is well-posed with no smoothing properties i.e.,  $u \in L^2(\Omega)$ , and discontinuities in the solution  $u$  induced by  $f$  will propagate along the flow field  $\beta$ , giving rise to internal layers.
- ▶ Locally degenerate problems. We assume that both  $\Omega^{\text{hyp}}$  and  $\Omega^{\text{ell}}$  are nonempty subsets. The model problem is then purely hyperbolic in  $\Omega^{\text{hyp}}$  and elliptic in the rest. Thus, we now define the common interface  $\mathcal{I} := \{x \in \Omega : \partial\Omega^{\text{hyp}} \cap \partial\Omega^{\text{ell}}\}$ . We emphasize that such problems are particularly delicate to solve since the solution can be discontinuous at the portion  $\mathcal{I}^- := \{x \in \mathcal{I} : \beta(x) \cdot \mathbf{n}_{\mathcal{I}} < 0\}$ , where  $\mathbf{n}_{\mathcal{I}}$  is an (arbitrary) oriented unit normal vector pointing out of the elliptic region. Concretely,  $\mathcal{I}^-$  corresponds to a subset of  $\mathcal{I}$  where the advection field flows from the hyperbolic side to the elliptic side. Thus, we set  $\mathcal{I}^+ := \mathcal{I} \setminus \mathcal{I}^-$ .

For a given forcing term  $f \in L^2(\Omega)$ , the continuous problem reads:

$$\begin{aligned} \nabla \cdot (-\kappa \nabla u + \beta u) + \gamma u &= f & \text{in} & \quad \Omega, \\ \llbracket u \rrbracket &= \mathbf{0} & \text{on} & \quad \mathcal{I}^+, \\ u &= 0 & \text{on} & \quad \Gamma^-, \end{aligned} \quad (7)$$

where  $\llbracket \cdot \rrbracket$  denotes the standard DG-jump trace operator as defined in [1]. The well-posedness of the boundary value problem (7) has been analyzed by Oleřnik and Radkevič in [24]. They proved the existence and uniqueness of a weak solution for homogeneous and nonhomogeneous Dirichlet boundary conditions, respectively. The main objective of the present paper is to propose, in a unified formalism, an inspired interior penalty HDG method that can treat all of these abovementioned physical situations in an automated fashion. Before doing so, let us specify the discrete setting concerning mesh assumptions, the definition of trace operators and the approximation spaces that will be used later in the rest of this paper.

### 3. Discrete setting

#### 3.1. Mesh notations

Let  $h$  be a positive parameter, and assume without loss of generality that  $h \leq 1$ . We denote by  $\mathcal{T}_h^{\text{ell}}$  (resp.  $\mathcal{T}_h^{\text{hyp}}$ ) a conformal partition of the subdomain  $\Omega^{\text{ell}}$  (resp.  $\Omega^{\text{hyp}}$ ) satisfying  $\mathcal{T}_h^{\text{ell}} \cap \mathcal{T}_h^{\text{hyp}} = \emptyset$ . The set of all mesh elements is denoted by  $\mathcal{T}_h$ , i.e.,  $\mathcal{T}_h := \mathcal{T}_h^{\text{ell}} \cup \mathcal{T}_h^{\text{hyp}}$ , where  $h$  stands for the largest diameter of all elements. We precisely state here that the sets of mesh elements  $\mathcal{T}_h^{\text{ell}}$  and  $\mathcal{T}_h^{\text{hyp}}$  can be composed of several types of geometric elements, i.e., hybrid meshes. Following our notation, the generic term *interface* indicates a  $(d-1)$ -dimensional geometric object with a positive measure, i.e., an edge if  $d=2$  and a face if  $d=3$ . The set of boundary interfaces is denoted by  $\mathcal{F}_h^{\text{b}}$ , i.e.,  $F \in \mathcal{F}_h^{\text{b}}$  if there exists  $E$  in  $\mathcal{T}_h$  such that  $F := \partial E \cap \partial \Omega$ . We assume that the set  $\mathcal{F}_h^{\text{b}}$  coincides with the disjoint boundary partition  $\mathcal{F}_h^{\text{b}\pm}$ , i.e., the boundary interfaces lying entirely in one of the subsets  $\Gamma^\pm$ . Likewise, we denote by  $\mathcal{F}_h^{\text{i}}$  the set of interior interfaces, i.e.,  $F \in \mathcal{F}_h^{\text{i}}$  if there exists  $E_1$  and  $E_2$  in  $\mathcal{T}_h$  such that  $F := \partial E_1 \cap \partial E_2$ . The set of all interfaces is denoted by  $\mathcal{F}_h$ , i.e.,  $\mathcal{F}_h := \mathcal{F}_h^{\text{i}} \cup \mathcal{F}_h^{\text{b}}$ , and we set  $\mathcal{F}_h^\pm := \mathcal{F}_h^{\text{i}} \cup \mathcal{F}_h^{\text{b}\pm}$ . In particular, we denote by  $\mathcal{I}_h$  the subset of  $\mathcal{F}_h^{\text{i}}$ , which belongs to  $\mathcal{I}$  ( $\mathcal{I}_h \subset \mathcal{F}_h^{\text{i}}$ ), i.e.,  $\mathcal{I}_h$  corresponds here to the discrete counterpart of  $\mathcal{I}$ . We assume first that for any interface  $F \in \mathcal{I}_h$  (i) there exist  $E_1 \in \mathcal{T}_h^{\text{ell}}$  and  $E_2 \in \mathcal{T}_h^{\text{hyp}}$  such that  $F := \partial E_1 \cap \partial E_2$ , and (ii)  $F$  lies entirely in one of the disjoint subsets  $\mathcal{I}_h^\pm$  corresponding to the discrete counterpart of  $\mathcal{I}^\pm$ , respectively. Moreover, for any mesh element  $E \in \mathcal{T}_h$ , we denote by  $\mathcal{F}_E := \{F \in \mathcal{F}_h : F \subset \partial E\}$  the set of interfaces composing the boundary of  $E$ , and we set  $\eta_E := \text{card}(\mathcal{F}_E)$ . For all elements  $X$  of  $\mathcal{T}_h$  or  $\mathcal{F}_h$ , we denote by  $|X|$  and  $h_X$  the measure and diameter of  $X$ , respectively.

#### 3.2. Approximation spaces

For any polyhedral domain  $D \subset \mathbb{R}^d$  with  $\partial D \subset \mathbb{R}^{d-1}$ , we denote by  $(\cdot, \cdot)_{0,D}$  (resp.  $\langle \cdot, \cdot \rangle_{0,\partial D}$ ) the  $L^2$ -inner product in  $L^2(D)$  (resp.  $L^2(\partial D)$ ) equipped with its natural norm  $\|\cdot\|_{0,D}$  (resp.  $\|\cdot\|_{0,\partial D}$ ). Similarly, we denote by  $H^s(D)$  the usual Hilbert space of index  $s$  on  $D$  equipped with its natural norm  $\|\cdot\|_{s,D}$  and seminorm  $|\cdot|_{s,D}$ , respectively. In particular, if  $s=0$ , then we set  $H^0(D) = L^2(D)$ . We now denote by  $H^s(\mathcal{T}_h)$  the usual broken Sobolev space and by  $\nabla_h$  the broken gradient operator acting on  $H^s(\mathcal{T}_h)$  with  $s \geq 1$ . Let us now introduce compact notation associated with the discrete  $L^2$ -inner scalar product:

$$(\cdot, \cdot)_{0,\mathcal{T}_h} := \sum_{E \in \mathcal{T}_h} (\cdot, \cdot)_{0,E}, \quad \langle \cdot, \cdot \rangle_{0,\partial \mathcal{T}_h} := \sum_{E \in \mathcal{T}_h} \langle \cdot, \cdot \rangle_{0,\partial E} \quad \text{and} \quad \langle \cdot, \cdot \rangle_{0,\mathcal{F}_h} := \sum_{F \in \mathcal{F}_h} \langle \cdot, \cdot \rangle_{0,F}, \quad (8)$$

and we denote by  $\|\cdot\|_{0,\mathcal{T}_h}$ ,  $\|\cdot\|_{0,\partial \mathcal{T}_h}$  and  $\|\cdot\|_{0,\mathcal{F}_h}$  its corresponding norms, respectively. As usual in HDG methods, we consider broken Sobolev spaces:

$$\mathbb{P}_k(\mathcal{T}_h) := \{v_h \in L^2(\mathcal{T}_h) : v_h|_E \in \mathbb{P}_k(E), \forall E \in \mathcal{T}_h\}, \quad (9)$$

and similarly for  $\mathbb{P}_k(\mathcal{F}_h)$ . Here,  $\mathbb{P}_k(X)$  denotes the space of polynomials of at least degree  $k$  on  $X$ , where  $X$  corresponds to a generic element of  $\mathcal{T}_h$  or  $\mathcal{F}_h$ , respectively. For H-IP discretization, two types of discrete variables are necessary to approximate the weak solution  $u$  of problem (7). First, the discrete variable  $u_h \in \mathcal{V}_h$  is defined within each mesh element, and its trace  $\hat{u}_h \in \hat{\mathcal{V}}_h^0$  is defined on the mesh skeleton with respect to the imposed Dirichlet boundary conditions at the boundary part  $\Gamma^-$ . Thus, we set:

$$\mathcal{V}_h := \mathbb{P}_k(\mathcal{T}_h) \quad \text{and} \quad \hat{\mathcal{V}}_h^0 := \{\hat{v}_h \in \mathbb{P}_k(\mathcal{F}_h) : \hat{v}_h|_F = 0, \forall F \in \mathcal{F}_h^{\text{b}-}\}. \quad (10)$$

For clarity, we introduce compact discrete variables  $\mathbf{u}_h := (u_h, \hat{u}_h)$  and  $\mathbf{v}_h := (v_h, \hat{v}_h)$  belonging to the composite approximation space  $\mathbf{V}_h := \mathcal{V}_h \times \hat{\mathcal{V}}_h^0$ , i.e.,  $\mathbf{u}_h, \mathbf{v}_h \in \mathbf{V}_h$ . Finally, let us retain the discrete trace inequalities. For all  $v_h \in \mathcal{V}_h$ , the following holds,

$$\|v_h\|_{0,F} \leq C_{\text{tr}} h_E^{-1/2} \|v_h\|_{0,E}, \quad (11)$$

where  $C_{\text{tr}}$  is a positive constant independent of  $h_E$ .

### 3.3. Discrete trace operators

Let  $[\![\cdot]\!]$  denote the standard DG-jump operator as introduced by Brezzi *et al.* in [3]. For all  $E \in \mathcal{T}_h$  and  $F \in \partial E$ , we now define the HDG-jump operator of the composite discrete variable  $\mathbf{v}_h \in \mathbf{V}_h$  across  $F$  as  $[\![\mathbf{v}_h]\!]_{E,F} := (v_h|_F - \hat{v}_h|_F)\mathbf{n}_{E,F}$ , where  $\mathbf{n}_{E,F}$  denotes the unit normal vector to  $F$  pointing out of  $E$ . To ensure that confusion cannot arise, we voluntarily omit the subscripts  $E$  and  $F$  from the definition, and we simply write  $[\![\mathbf{v}_h]\!] := (v_h - \hat{v}_h)\mathbf{n}$ . Finally, we define the weighted-average operator denoted  $\{\!\!\{ \cdot \}\!\!\}_\omega$  and its conjugate  $\{\!\!\{ \cdot \}\!\!\}_\omega^\omega$ . For all  $F \in \mathcal{F}_h^i$  with  $F := \partial E_1 \cap \partial E_2$  and  $v \in H^s(\mathcal{T}_h)$  with  $s \geq 1$ , we set:

$$\{\!\!\{ v \}\!\!\}_\omega := \omega_1 v_1 + \omega_2 v_2 \quad \text{and} \quad \{\!\!\{ v \}\!\!\}_\omega^\omega := \omega_2 v_1 + \omega_1 v_2, \quad (12)$$

where  $v_i = v|_{E_i,F}$  and  $\omega := (\omega_1, \omega_2)$  is a double-valued function verifying that weights satisfy  $\omega_1 + \omega_2 = 1$ . If  $F \in \mathcal{F}_h^b$ , we then assume that  $\{\!\!\{ v \}\!\!\}_\omega = \{\!\!\{ v \}\!\!\}_\omega^\omega := v$ . If  $\omega = (1/2, 1/2)$ , we then recover the classical average operator, and we will omit the subscript  $\omega$  in their definitions. These definitions (12) are also available for any vector-valued function  $\mathbf{v}$ .

## 4. Hybridizable interior penalty discontinuous Galerkin method

In this section, we describe the primal HDG method for solving the problem (7). First, we derive it intuitively, and we then propose a compact notation of all of these variants. The consistency and coercivity properties are also discussed in order to ensure the well-posedness of the discrete problem. Finally, we suggest an adaptive strategy for selecting suited penalty parameters with respect to the coercivity requirement for all mentioned regimes.

### 4.1. Intuitive derivation

The discrete formulation of the continuous model problem (7) can be intuitively derived with respect to the three following steps:

- **Global weak formulation:** Let  $\mathbf{u}_h, \mathbf{v}_h \in \mathbf{V}_h$ . For all  $E \in \mathcal{T}_h$ , multiplying (1) by a test function  $v_h$ , and integrating by parts over  $E$ , we obtain a local equation:

$$-(\boldsymbol{\sigma}_h(u_h), \nabla_h v_h)_{0,E} + \langle \hat{\boldsymbol{\sigma}}_h(\mathbf{u}_h) \cdot \mathbf{n}, v_h \rangle_{0,\partial E} + (\mu u_h, v_h)_{0,E} = (f, v_h)_{0,E}, \quad (13)$$

where  $\boldsymbol{\sigma}_h(u_h) := -\boldsymbol{\kappa} \nabla_h u_h + \boldsymbol{\beta} u_h$  corresponds to the approximation of the total flux on  $E$  and  $\hat{\boldsymbol{\sigma}}_h(\mathbf{u}_h)$ , its trace approximation on  $\partial E$  that we will precisely define below. By summing (13) over all elements  $E$  of  $\mathcal{T}_h$ , we then obtain the global equation,

$$-(\boldsymbol{\sigma}_h(u_h), \nabla_h v_h)_{0,\mathcal{T}_h} + \langle \hat{\boldsymbol{\sigma}}_h(\mathbf{u}_h) \cdot \mathbf{n}, v_h \rangle_{0,\partial \mathcal{T}_h} + (\mu u_h, v_h)_{0,\mathcal{T}_h} = (f, v_h)_{0,\mathcal{T}_h}. \quad (14)$$

Let us precisely state at this stage that  $\hat{\boldsymbol{\sigma}}_h(\mathbf{u}_h)$  can be evaluated independently on both sides of a given interface. Consequently, an additional equation needs to be included in (14) to ensure the continuity requirements.

- Transmission conditions: By considering that  $\hat{v}_h \in \hat{\mathcal{V}}_h^0$  vanishes on  $\mathcal{F}_h^{b-}$ , we impose the global transmission requirements respecting the outflow boundary condition on  $\mathcal{F}_h^{b+}$ :

$$\langle \hat{\boldsymbol{\sigma}}_h(\mathbf{u}_h) \cdot \mathbf{n}, \hat{v}_h \rangle_{0, \partial \mathcal{T}_h} = \langle (\boldsymbol{\beta} \cdot \mathbf{n}) \hat{u}_h, \hat{v}_h \rangle_{0, \mathcal{F}_h^{b+}}, \quad (15)$$

which is the statement of weak continuity of  $\hat{\boldsymbol{\sigma}}_h(\mathbf{u}_h) \cdot \mathbf{n}$  across the mesh skeleton  $\mathcal{F}_h^+$ . Thus, substituting (15) into (14) leads to

$$-(\boldsymbol{\sigma}_h(u_h), \nabla_h v_h)_{0, \mathcal{T}_h} + (\mu u_h, v_h)_{0, \mathcal{T}_h} + \langle \hat{\boldsymbol{\sigma}}_h(\mathbf{u}_h), \llbracket \mathbf{v}_h \rrbracket \rangle_{0, \partial \mathcal{T}_h} + \langle (\boldsymbol{\beta} \cdot \mathbf{n}) \hat{u}_h, \hat{v}_h \rangle_{0, \mathcal{F}_h^{b+}} = (f, v_h)_{0, \mathcal{T}_h}. \quad (16)$$

- Numerical Flux in the HDG sense: The discrete problem (16) is then closed by setting  $\hat{\boldsymbol{\sigma}}_h(\mathbf{u}_h)$  at the element level. For all  $E \in \mathcal{T}_h$ , we assume the following simple form:

$$\hat{\boldsymbol{\sigma}}_h(\mathbf{u}_h) := \boldsymbol{\sigma}_h(u_h) + \tau \llbracket \mathbf{u}_h \rrbracket \quad \text{on } \partial E, \quad (17)$$

where  $\tau$  denotes the stabilization penalty parameter that we describe precisely below.

Inserting (17) into (16) leads to the incomplete scheme of the hybridizable interior penalty discontinuous Galerkin (H-IIP) method. Several variations of this methodology can then be derived by controlling the introduction of additional consistent terms to the discrete formulation (16) that we have summarized below.

#### 4.2. Compact discrete formulation

Thus, the compact formulation of the H-IP method consists of seeking  $\mathbf{u}_h \in \mathcal{V}_h$  such that

$$a_h^{(\epsilon)}(\mathbf{u}_h, \mathbf{v}_h) = l(\mathbf{v}_h), \quad \forall \mathbf{v}_h \in \mathcal{V}_h, \quad (18)$$

where  $l(\mathbf{v}_h) := (f, v_h)_{0, \mathcal{T}_h}$  and the bilinear form  $a_h^{(\epsilon)}$  can be linearly decomposed following its diffusive, advective-reactive, and stability contributions, respectively:

$$a_h^{(\epsilon)}(\mathbf{u}_h, \mathbf{v}_h) := a_{\boldsymbol{\kappa}, h}^{(\epsilon)}(\mathbf{u}_h, \mathbf{v}_h) + a_{\boldsymbol{\beta}, \mu, h}(\mathbf{u}_h, \mathbf{v}_h) + s_h(\mathbf{u}_h, \mathbf{v}_h). \quad (19)$$

Here,  $a_{\boldsymbol{\kappa}, h}^{(\epsilon)}$  corresponds solely to the H-IP discretization of the diffusive part and is given by:

$$a_{\boldsymbol{\kappa}, h}^{(\epsilon)}(\mathbf{u}_h, \mathbf{v}_h) := (\boldsymbol{\kappa} \nabla_h u_h, \nabla_h v_h)_{0, \mathcal{T}_h} - \langle \boldsymbol{\kappa} \nabla_h u_h, \llbracket \mathbf{v}_h \rrbracket \rangle_{0, \partial \mathcal{T}_h} - \epsilon \langle \boldsymbol{\kappa} \nabla_h v_h, \llbracket \mathbf{u}_h \rrbracket \rangle_{0, \partial \mathcal{T}_h}, \quad (20)$$

where the parameter  $\epsilon \in \{0, \pm 1\}$  controls the introduction of the (consistent) symmetry term  $\langle \boldsymbol{\kappa} \nabla_h v_h, \llbracket \mathbf{u}_h \rrbracket \rangle_{0, \partial \mathcal{T}_h}$ . Similar to the standard IPDG methods,  $\epsilon = 0$  corresponds to the Incomplete scheme denoted as the H-IIP method as described above, while  $\epsilon = +1$  (resp.  $\epsilon = -1$ ) denotes the Symmetric (resp. Nonsymmetric) scheme denoted as the H-SIP (resp. H-NIP) method. Thus, the advective-reactive part is discretized as follows:

$$a_{\boldsymbol{\beta}, \mu, h}(\mathbf{u}_h, \mathbf{v}_h) := -(\boldsymbol{\beta} u_h, \nabla_h v_h)_{0, \mathcal{T}_h} + (\mu u_h, v_h)_{0, \mathcal{T}_h} + \langle \boldsymbol{\beta} u_h, \llbracket \mathbf{v}_h \rrbracket \rangle_{0, \partial \mathcal{T}_h} + \langle (\boldsymbol{\beta} \cdot \mathbf{n}) \hat{u}_h, \hat{v}_h \rangle_{0, \mathcal{F}_h^{b+}}. \quad (21)$$

The last quantity  $s_h$  in (19) is called the discrete stability form based on jump-penalty terms and is given by:

$$s_h(\mathbf{u}_h, \mathbf{v}_h) := \langle \tau \llbracket \mathbf{u}_h \rrbracket, \llbracket \mathbf{v}_h \rrbracket \rangle_{0, \partial \mathcal{T}_h}, \quad (22)$$



where  $\tau$  denotes the total stabilization function accounting for both the diffusive and advective normal effects denoted by  $\tau_\kappa$  and  $\tau_\beta$ , respectively; i.e.,  $\tau := \tau(\tau_\kappa, \tau_\beta)$ . Concretely,  $\tau_\kappa$  (resp.  $\tau_\beta$ ) corresponds to an arbitrary given definition of the penalty parameter in the context of a pure-diffusive (resp. -advective) regime. Let us also precisely state that  $\tau_\kappa$  and  $\tau_\beta$ , and hence  $\tau$  can be distinctively evaluated on both sides of an interface of the mesh skeleton. Different formulations have been proposed in the literature, and its definition is concretely prescribed regarding the predominant normal effect. In what follows, we shall assume some properties and minimal requirements concerning the function  $\tau$ :

**Proposition 4.1.** *The stabilization function  $\tau : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  (22) is chosen such that:*

1. *For all  $s, t \in \mathbb{R}$ , the following holds:*

$$s := \tau(s, 0) \quad \text{and} \quad t := \tau(0, t). \quad (23)$$

*The first assumption (23) allows for an encompassed treatment of both (extremal) configurations, i.e., the pure-diffusive and pure-advective regimes characterized by  $\tau_\beta := 0$  and  $\tau_\kappa := 0$ , respectively.*

2. *We shall assume the existence of a constant  $\tau_0 > 0$  such that, for all  $s, t \in \mathbb{R}$ , then*

$$\min(\tau(s, t) - \tau(s, 0) + \frac{\beta \cdot \mathbf{n}}{2})|_{\partial\mathcal{T}_h} \geq \tau_0 \quad (24)$$

*The second assumption ensures the coercivity and hence the well-posedness of the discrete problem (18) (see Lemma 4.3.2).*

**Definition 4.1** (Diffusive penalty parameter). *For all  $E \in \mathcal{T}_h$  and  $F \in \mathcal{F}_E$ , we assume that the diffusive parameter  $\tau_\kappa|_{E,F}$  has the following form,*

$$\tau_\kappa|_{E,F} := \alpha_0 C_{\text{tr}}^2 \frac{\kappa_{E,F}}{h_E} \quad \text{on} \quad F \in \partial E, \quad (25)$$

*where  $\kappa_{E,F} := \mathbf{n}_{E,F} \kappa_E \mathbf{n}_{E,F}$  corresponds to the normal diffusivity,  $\alpha_0 > 0$  is a user-dependent parameter, and  $C_{\text{tr}}$  is the constant of the discrete trace inequality (11).*

**Remark 4.1** (Static condensation). *As mentioned in [4] (see, e.g., [5]), the technique of static condensation was introduced to reduce the size of the discrete matrix associated with the global problem (18). Indeed, let  $\mathbf{U}_h := [\mathbf{U}_h, \hat{\mathbf{U}}_h]$  denote the vector of DOFs of the composite variable  $\mathbf{u}_h \in \mathcal{V}_h$ , which is composed of interior- and interface-based DOFs denoted by  $\mathbf{U}_h$  and  $\hat{\mathbf{U}}_h$ , respectively. The strategy consists of eliminating interior-based unknowns from the above equations by successively projecting (18) on  $(v_h, 0)$  and  $(0, \hat{v}_h)$ . We thus obtain*

$$\begin{aligned} a_h^{(\epsilon)}(\mathbf{u}_h, (v_h, 0)) &: \mathbf{A}_{uu} \mathbf{U}_h + \mathbf{A}_{u\hat{u}} \hat{\mathbf{U}}_h = \mathbf{F}, \\ a_h^{(\epsilon)}(\mathbf{u}_h, (0, \hat{v}_h)) &: \mathbf{A}_{\hat{u}u} \mathbf{U}_h + \mathbf{A}_{\hat{u}\hat{u}} \hat{\mathbf{U}}_h = 0. \end{aligned} \quad (26)$$

*Due to the discontinuous nature of  $\mathcal{V}_h$ , all computations can be performed cellwise, leading to a block-diagonal matrix  $\mathbf{A}_{uu}$ , which can be easily inverted and eliminated. Finally, we obtain the linear system:*

$$[\mathbf{A}_{\hat{u}\hat{u}} - \mathbf{A}_{\hat{u}u} \mathbf{A}_{uu}^{-1} \mathbf{A}_{u\hat{u}}] \hat{\mathbf{U}}_h = \mathbf{G}. \quad (27)$$

*The matrix on the left-hand-side of (27) is called the Shur complement of  $\mathbf{A}_{uu}$ . Let us precisely state that the original discrete problem (18) and its reduced version (27) are globally nonsymmetric since the continuous problem (26) is itself nonsymmetric: this is due to the advective term.*

### 4.3. Stability analysis

We now must check some favorable properties such as the consistency and stability of all  $\epsilon$ -variants of the H-IP method to ensure the existence and uniqueness of a discrete solution.

#### 4.3.1. Consistency

**Lemma 4.1** (Consistency). *Let  $u$  be the exact solution of the problem (7) and  $\hat{u}$  denote its trace on the mesh skeleton. By setting  $\mathbf{u} := (u, \hat{u})$ , the following holds*

$$a_h^{(\epsilon)}(\mathbf{u}, \mathbf{v}_h) = l(\mathbf{v}_h), \quad \forall \mathbf{v}_h \in \mathbf{V}_h, \quad (28)$$

for any value of the parameter  $\epsilon = \{0, \pm 1\}$ .

*Proof.* The regularity of  $\mathbf{u}$  implies that its jump (in the HDG sense) is null on  $\partial\mathcal{T}_h$ , i.e., for all  $E \in \mathcal{T}_h$  and  $F \in \mathcal{F}_E$ , and then  $\llbracket \mathbf{u} \rrbracket := 0$ , since  $u$  is a single-valued field on  $F \in \mathcal{F}_E$ . Thus, by setting  $\mathbf{v}_h := (v_h, 0)$ , integrating by parts on each element of the mesh, and finally observing that  $\llbracket -\boldsymbol{\kappa} \nabla_h u + \boldsymbol{\beta} u \rrbracket := 0$  on internal interfaces, we immediately have

$$a_h^{(\epsilon)}(\mathbf{u}, (v_h, 0)) := \sum_{E \in \mathcal{T}_h} (\nabla_h \cdot (-\boldsymbol{\kappa} \nabla_h u + \boldsymbol{\beta} u) + \mu u, v_h) = \sum_{E \in \mathcal{T}_h} (f, v_h). \quad (29)$$

Considering now that  $\mathbf{v}_h := (0, \hat{v}_h) \in \mathbf{V}_h$ , we then obtain

$$a_h^{(\epsilon)}(\mathbf{u}, (0, \hat{v}_h)) := - \sum_{E \in \mathcal{T}_h} \langle (-\boldsymbol{\kappa} \nabla_h u + \boldsymbol{\beta} u) \cdot \mathbf{n}, \hat{v}_h \rangle_{0, \partial A} + \langle (\boldsymbol{\beta} \cdot \mathbf{n}) u, \hat{v}_h \rangle_{\mathcal{F}_h^{b+}} = 0, \quad (30)$$

which corresponds to the (imposed) transmission conditions. The proof is then completed by summing (29) and (30).  $\square$

#### 4.3.2. Coercivity

Let us now introduce the natural energy-norm  $\|\cdot\|$  equipping  $\mathbf{V}_h$ . For all  $\mathbf{v}_h \in \mathbf{V}_h$ , it is thus given by

$$\|\mathbf{v}_h\|^2 := \|\boldsymbol{\kappa}^{1/2} \nabla_h v_h\|_{0, \mathcal{T}_h}^2 + \|\mu_0^{1/2} v_h\|_{0, \mathcal{T}_h}^2 + \|\beta^{1/2} \hat{v}_h\|_{0, \mathcal{F}_h^{b+}}^2 + |\mathbf{v}_h|_{\tau_0}^2, \quad (31)$$

which clearly depends on constitutive coefficients  $\boldsymbol{\kappa}$ ,  $\boldsymbol{\beta} := \boldsymbol{\beta} \cdot \mathbf{n}$ , and parameters  $\mu_0$  and  $\tau_0$  as given in (4) and (24), respectively. Here,  $|\cdot|_\gamma$  corresponds to the HDG-jump seminorm, which is defined as follows:

$$|\mathbf{v}_h|_\gamma^2 := \sum_{E \in \mathcal{T}_h} |\mathbf{v}_h|_{\gamma, \partial E}^2 \quad \text{with} \quad |\mathbf{v}_h|_{\gamma, \partial E}^2 := \sum_{F \in \mathcal{F}_E} \|\gamma_{E,F}^{1/2} \llbracket \mathbf{v}_h \rrbracket\|_{0,F}^2, \quad (32)$$

where  $\gamma_{E,F} \geq 0$  is an arbitrary positive constant associated with  $F \in \mathcal{F}_E$ . The norm (31) is well-defined owing to assumptions (4) and (24). To prove the coercivity, let us now introduce the following intermediate result, which is a minor adaptation to the  $\boldsymbol{\kappa}$ -tensor field of the proof given by Wells in [28]:

**Lemma 4.2** (A bound on consistency term). *Assuming (25), for any  $\xi > 0$ , the following holds,*

$$\left| \langle \boldsymbol{\kappa} \nabla_h v_h, \llbracket \mathbf{v}_h \rrbracket \rangle_{0, \partial \mathcal{T}_h} \right| \leq \frac{\xi \eta_0}{2\alpha_0} \|\boldsymbol{\kappa}^{1/2} \nabla_h v_h\|_{0, \mathcal{T}_h}^2 + \frac{1}{2\xi} |\mathbf{v}_h|_{\tau_\kappa}^2, \quad (33)$$

where  $\eta_0 := \max_{\forall E \in \mathcal{T}_h} (\eta_E)$ .

*Proof.* The decomposition of the consistency term yields

$$\left| \langle \boldsymbol{\kappa} \nabla_h v_h, \llbracket \mathbf{v}_h \rrbracket \rangle_{0, \partial \mathcal{T}_h} \right| \leq \sum_{E \in \mathcal{T}_h} \left| \langle \boldsymbol{\kappa}^{1/2} \nabla_h v_h, \boldsymbol{\kappa}^{1/2} \llbracket \mathbf{v}_h \rrbracket \rangle_{0, \partial E} \right|. \quad (34)$$

Successively applying the Cauchy–Schwarz inequality and the discrete trace inequality (11), and using the definition of  $\tau_\kappa$  given in (25), we infer that

$$\left| \langle \boldsymbol{\kappa}^{1/2} \nabla_h v_h, \boldsymbol{\kappa}^{1/2} \llbracket \mathbf{v}_h \rrbracket \rangle_{0, \partial E} \right| \leq \left[ \frac{h_E}{\alpha_0 C_{\text{tr}}^2} \right]^{1/2} \|\boldsymbol{\kappa}^{1/2} \nabla_h v_h\|_{0, \partial E} \|\tau_\kappa^{1/2} \llbracket \mathbf{v}_h \rrbracket\|_{0, \partial E}, \quad (35)$$

$$\leq \left[ \frac{\eta_E}{\alpha_0} \right]^{1/2} \|\boldsymbol{\kappa}^{1/2} \nabla_h v_h\|_{0, E} |\mathbf{v}_h|_{\tau_\kappa, \partial E}, \quad (36)$$

The proof is then completed by applying Young’s inequality for any  $\xi > 0$  and summing over all mesh elements.  $\square$

**Lemma 4.3** (Coercivity). *For any value of the parameter  $\epsilon := \{0, \pm 1\}$ , there exists a positive constant  $C$  independent of  $h$  such that*

$$a_h^{(\epsilon)}(\mathbf{v}_h, \mathbf{v}_h) \geq C \|\mathbf{v}_h\|^2. \quad (37)$$

*Proof.* Setting  $\mathbf{u}_h = \mathbf{v}_h$  in discrete bilinear forms  $a_{\boldsymbol{\kappa}, h}^{(\epsilon)}$  (20) and  $s_h$  (22), we immediately obtain that:

$$a_{\boldsymbol{\kappa}, h}^{(\epsilon)}(\mathbf{v}_h, \mathbf{v}_h) + s_h(\mathbf{v}_h, \mathbf{v}_h) := \|\boldsymbol{\kappa}^{1/2} \nabla_h v_h\|_{0, \mathcal{T}_h}^2 - (1 + \epsilon) \langle \boldsymbol{\kappa} \nabla_h v_h, \llbracket \mathbf{v}_h \rrbracket \rangle_{0, \partial \mathcal{T}_h} + \langle \tau \llbracket \mathbf{v}_h \rrbracket, \llbracket \mathbf{v}_h \rrbracket \rangle_{0, \partial \mathcal{T}_h}. \quad (38)$$

After integration by parts, the advective bilinear form (21) yields that

$$a_{\beta, \mu, h}(\mathbf{v}_h, \mathbf{v}_h) := \|\mu_*^{1/2} v_h\|_{0, \mathcal{T}_h}^2 + \|\beta^{1/2} \hat{v}_h\|_{0, \mathcal{F}_h^{\text{b}+}}^2 + \underbrace{\langle \beta v_h, \llbracket \mathbf{v}_h \rrbracket \rangle_{0, \partial \mathcal{T}_h} - \frac{1}{2} \langle \beta v_h, v_h \rangle_{0, \partial \mathcal{T}_h}}_{=T_1}, \quad (39)$$

where  $\mu_* := \mu + \frac{\nabla \cdot \beta}{2} > 0$  by virtue of (4), and  $\beta := \boldsymbol{\beta} \cdot \mathbf{n}$ . Let us now focus on the (last) quantity  $T_1$  in (39). Considering that  $\hat{v}_h$  and  $\beta$  are single-valued on interfaces of the mesh skeleton, and after some tedious algebraic manipulations, we deduce that

$$\begin{aligned} T_1 &= \langle \beta \llbracket \mathbf{v}_h \rrbracket, \llbracket \mathbf{v}_h \rrbracket \rangle_{0, \partial \mathcal{T}_h} + \langle \beta \hat{v}_h, v_h - \hat{v}_h \rangle_{0, \partial \mathcal{T}_h} - \frac{1}{2} \langle \beta v_h, v_h \rangle_{0, \partial \mathcal{T}_h} \\ &= \langle \beta \llbracket \mathbf{v}_h \rrbracket, \llbracket \mathbf{v}_h \rrbracket \rangle_{0, \partial \mathcal{T}_h} - \frac{1}{2} \langle \beta (v_h - \hat{v}_h), v_h \rangle_{0, \partial \mathcal{T}_h} + \frac{1}{2} \langle \beta (v_h - \hat{v}_h), \hat{v}_h \rangle_{0, \partial \mathcal{T}_h} \\ &= \frac{1}{2} \langle \beta \llbracket \mathbf{v}_h \rrbracket, \llbracket \mathbf{v}_h \rrbracket \rangle_{0, \partial \mathcal{T}_h} \end{aligned} \quad (40)$$

Inserting (40) in (39), and finally collecting (38), we immediately obtain

$$\begin{aligned} a_h^{(\epsilon)}(\mathbf{v}_h, \mathbf{v}_h) &:= \|\boldsymbol{\kappa}^{1/2} \nabla_h v_h\|_{0, \mathcal{T}_h}^2 + \|\mu_*^{1/2} v_h\|_{0, \mathcal{T}_h}^2 + \|\beta^{1/2} \hat{v}_h\|_{0, \mathcal{F}_h^{\text{b}+}}^2 + \langle \tau_* \llbracket \mathbf{v}_h \rrbracket, \llbracket \mathbf{v}_h \rrbracket \rangle_{0, \partial \mathcal{T}_h} \\ &\quad - (1 + \epsilon) \langle \boldsymbol{\kappa} \nabla_h v_h, \llbracket \mathbf{v}_h \rrbracket \rangle_{0, \partial \mathcal{T}_h}, \end{aligned} \quad (41)$$

where  $\tau_* := \tau + \frac{\beta \cdot \mathbf{n}}{2}$ . Following (4) and (24), we deduce that  $\mu_* \geq \mu_0$  and  $\tau_* \geq \tau_0$ , and hence

$$a_h^{(\epsilon)}(\mathbf{v}_h, \mathbf{v}_h) \geq \|\mathbf{v}_h\|^2 - (1 + \epsilon) \langle \boldsymbol{\kappa} \nabla_h v_h, \llbracket \mathbf{v}_h \rrbracket \rangle_{0, \partial \mathcal{T}_h}, \quad (42)$$

proving immediately the coercivity of the H-NIP scheme ( $\epsilon = -1$ ) with  $C = 1$ . Else, by considering Lemmata 4.2 and assuming  $0 < \xi < 1$ , we easily infer that

$$\begin{aligned} a_h^{(\epsilon)}(\mathbf{v}_h, \mathbf{v}_h) &\geq \left(1 - \frac{\xi \eta_0}{\alpha_0}\right) \|\boldsymbol{\kappa}^{1/2} \nabla_h v_h\|_{0, \mathcal{T}_h}^2 + \|\mu_0^{1/2} v_h\|_{0, \mathcal{T}_h}^2 + \|\beta^{1/2} \hat{v}_h\|_{0, \mathcal{F}_h^{\text{b}+}}^2 + |\mathbf{v}_h|_{\tau_0}^2, \\ &\geq C_\xi \|\mathbf{v}_h\| \end{aligned}$$

where  $C_\xi := \min(1, 1 - \frac{\xi \eta_0}{\alpha_0})$ . The proof is ended by choosing  $\alpha_0$  such that  $\alpha_0 > \xi \eta_0$ .  $\square$

**Remark 4.2.** *A straightforward consequence of the consistency and coercivity requirements via the Lax–Milgram Theorem is the well-posedness of (18); i.e., the existence and uniqueness of the discrete solution  $\mathbf{u}_h \in \mathcal{V}_h$  are ensured.*

#### 4.4. Adaptive stabilization strategy

In practice, the choice of  $\tau$  is quite delicate, as it strongly affects the accuracy of the HDG method (18). Indeed, its definition directly impacts the numerical flux approximations on interfaces  $F \in \mathcal{F}_h^+$ . To prove its relevance, let us apply a continuity argument, i.e.,  $\llbracket \hat{\boldsymbol{\sigma}}_h(\mathbf{u}_h) \rrbracket = 0$ , on an interior interface  $F := \partial E_1 \cap \partial E_2$ . We immediately deduce that  $(\hat{u}_h, \hat{\boldsymbol{\sigma}}_h)$  can be expressed only in terms of the discrete variables  $(u_h, \boldsymbol{\sigma}_h)$  on both sides of  $F$ ,

$$\begin{cases} \hat{u}_h = \{ \{ u_h \} \boldsymbol{\omega} + \alpha \llbracket \boldsymbol{\sigma}_h \rrbracket, \\ \hat{\boldsymbol{\sigma}}_h = \{ \{ \boldsymbol{\sigma}_h \} \boldsymbol{\omega} + \eta \llbracket u_h \rrbracket, \end{cases} \quad (43a)$$

$$\quad (43b)$$

where  $\boldsymbol{\sigma}_h := -\boldsymbol{\kappa} \nabla_h u_h + \beta u_h$ , and the parameters  $\boldsymbol{\omega}$ ,  $\alpha$  and  $\eta$  are given below by:

$$\boldsymbol{\omega} := \left( \frac{\tau_1}{\tau_1 + \tau_2}, \frac{\tau_2}{\tau_1 + \tau_2} \right), \quad \alpha := \frac{1}{\tau_1 + \tau_2}, \quad \text{and} \quad \eta := \frac{\tau_1 \tau_2}{\tau_1 + \tau_2}, \quad (44)$$

for any given finite value  $\tau_i := \tau_{E_i, F}$  of the total penalty parameter. To derive a suitable analytical expression of  $\tau$ , we now treat both hyperbolic and elliptic regimes distinctively.

##### 4.4.1. Hyperbolic regime

We assume here that  $\tau_\kappa := 0$ , and hence  $\tau := \tau_\beta$  by virtue of (23). To ensure the minimal requirement (24) in the context of the hyperbolic regime, the advective penalty parameter must be chosen such that  $\tau_\beta > |\beta \cdot \mathbf{n}|/2$  on the mesh skeleton. Furthermore, it is well known that an arbitrary choice of the stability parameter  $\tau_\beta$  can be detrimental in the context of pure-advective problems: discontinuities in the boundary data may trigger large spurious oscillations in the numerical solution. However, these drawbacks can be easily circumvented by adopting an *upwind-based* strategy. To this aim, we now consider the following definition of the advective penalty parameter:

**Definition 4.2** ( $\theta$ -upwind penalty). *For all  $E \in \mathcal{T}_h^{\text{hyp}}$  and  $F \in \partial E$ , we assume the following definition of the advective stabilization penalty parameter:*

$$\tau_{\beta_{E,F}}^\theta := \theta |\beta \cdot \mathbf{n}_{E,F}| \quad \text{on} \quad F \in \partial E, \quad (45)$$

where  $\theta > 1/2$  in order to ensure Proposition 4.1.

Following from the regularity of the Darcy field  $\beta$ , i.e.,  $[\![\beta]\!] = 0$ , we infer that the advective penalty parameter as defined in (45) is single-valued on interior interfaces of the hyperbolic region. Let us now introduce the *signum function*, which is given as follows:

$$\text{sgn}(x) := \frac{x}{|x|} = \begin{cases} -1 & \text{if } x < 0, \\ 0 & \text{if } x = 0, \\ +1 & \text{if } x > 0. \end{cases} \quad (46)$$

**Proposition 4.2** ( $\theta$ -upwind fluxes). *Following Definition 4.2, for all  $E_1, E_2 \in \mathcal{T}_h^{\text{hyp}}$  and  $F := \partial E_1 \cap \partial E_2$ , the corresponding numerical fluxes  $(\hat{u}_h^\theta, \hat{\sigma}_h^\theta)$  on  $F$  are given by*

$$\begin{cases} \hat{u}_h^\theta := \{u_h\} + \frac{\beta}{2\tau_\beta^\theta} [u_h] = \{u_h\}_{\omega_\theta}, \\ \hat{\sigma}_h^\theta := \beta \{u_h\} + \frac{\tau_\beta^\theta}{2} [u_h] = \beta \{u_h\}_{\omega^\theta}, \end{cases} \quad (47a)$$

$$\quad (47b)$$

where  $\omega_\theta := (\frac{1}{2} + \frac{\text{sgn}(\beta \cdot \mathbf{n}_1)}{2\theta}, \frac{1}{2} + \frac{\text{sgn}(\beta \cdot \mathbf{n}_2)}{2\theta})$  and  $\omega^\theta := (\frac{1}{2} + \frac{\theta}{2\text{sgn}(\beta \cdot \mathbf{n}_1)}, \frac{1}{2} + \frac{\theta}{2\text{sgn}(\beta \cdot \mathbf{n}_2)})$ .

*Proof.* The proof is evident by substituting the definition (45) in (43), and assuming that  $\kappa$  is null in  $\mathcal{T}_h^{\text{hyp}}$  and that  $\beta$  is single-valued on the mesh skeleton.  $\square$

**Remark 4.3** (Traditional schemes). *By appropriately selecting the value of  $\theta$  in (45), we can then establish bridges with some well-known stabilization schemes.*

- **Upwind-scheme:** By setting  $\theta = 1$ , we can observe that  $\omega_1 = \omega^1$ , and we recover the standard definition of upwinding fluxes denoted by  $(\hat{u}_h^{\text{up}}, \hat{\sigma}_h^{\text{up}})$ ,

$$\hat{u}_h^{\text{up}} := \begin{cases} u_i & \text{if } \beta \cdot \mathbf{n}_i > 0, \\ \{u_h\} & \text{if } \beta \cdot \mathbf{n}_i = 0, \end{cases} \quad \text{and} \quad \hat{\sigma}_h^{\text{up}} := \beta \hat{u}_h^{\text{up}}. \quad (48)$$

The corresponding upwind penalty parameter is denoted  $\tau_\beta^{\text{up}} := |\beta \cdot \mathbf{n}|$ .

- **Centered-scheme:** Assuming now  $\theta \rightarrow +\infty$ , we then obtain the centered fluxes  $(\hat{u}_h^c, \hat{\sigma}_h^c)$ ,

$$\hat{u}_h^c := \lim_{\theta \rightarrow +\infty} \{u_h\}_{\omega_\theta} = \{u_h\} \quad \text{and} \quad \hat{\sigma}_h^c := \beta \{u_h\} + \frac{\tau_\beta^\infty}{2} [u_h]. \quad (49)$$

However, this situation will be precluded in the rest of the paper as it consists of assigning an infinite value to the penalty parameter since  $\tau_\beta^\infty = \lim_{\theta \rightarrow +\infty} \tau_\beta^\theta = +\infty$ . This choice significantly reduces the accuracy of the discrete solution since it converges to the discrete solution produced by the standard conforming Galerkin method characterized by spurious oscillations.

**Remark 4.4** (Degenerate outflow boundaries). *Let us finally precisely denote the transmission conditions at degenerative outflow boundaries that belong to the hyperbolic region. For all  $F \in \mathcal{F}_h^{\text{b}+}$ , we impose that  $\hat{\sigma}_h(\mathbf{u}_h) \cdot \mathbf{n} = (\beta \cdot \mathbf{n})\hat{u}_h$ , where  $\hat{\sigma}_h(\mathbf{u}_h) := \beta u_h + \tau_\beta [u_h]$ . By combining these expressions, we observe that the role of  $\tau_\beta$  on outflow boundaries is clearly insignificant, since for any finite value of  $\tau_\beta > 0$ , we then obtain  $\hat{u}_h := u_h$ .*

Thereafter, we shall assume that the advective stabilization penalty parameter  $\tau_\beta$  is chosen accordingly with the  $\theta$ -upwind strategy described in Definition 4.2.

#### 4.4.2. Elliptic regime

We assume here that the diffusion is not degenerate, i.e.,  $\underline{\kappa} > 0$ . To ensure the minimal requirement (24) in the context of the elliptic regime, the total penalty parameter is chosen such that  $\tau \geq \tau_{\underline{\kappa}} + |\boldsymbol{\beta} \cdot \mathbf{n}|/2$ . Let us now introduce the Péclet number to locally characterize the regime of the physical process.

**Definition 4.3.** For all  $E \in \mathcal{T}_h^{\text{ell}}$  and  $F \in \partial E$ , we define the local (oriented) Péclet number as follows:

$$\text{Pe}_{E,F}^{\theta} := \theta \frac{\boldsymbol{\beta} \cdot \mathbf{n}_{E,F}}{\tau_{\underline{\kappa}_{E,F}}} \quad \text{on } F \in \partial E, \quad (50)$$

where  $\theta$  is a positive constant that we will precisely define below.

Concretely, the regime is considered as locally (i) diffusion-dominated if  $|\text{Pe}_{E,F}^{\theta}| \sim \theta$ , and (ii) advection-dominated if  $|\text{Pe}_{E,F}^{\theta}| \sim +\infty$ . Following the definition (50), we emphasize here that  $\text{Pe}^{\theta}$  can be distinctively evaluated on both sides of an interface  $F \in \mathcal{F}_h^i$ . For all  $E \in \mathcal{T}_h^{\text{ell}}$  and  $F \in \partial E$ , we now assume the following generic form of the (total) stabilization penalty parameter i.e.,

$$\tau_{E,F} := \tau_{\underline{\kappa}_{E,F}} |\mathcal{A}|(\text{Pe}_{E,F}^{\theta}), \quad (51)$$

where  $|\mathcal{A}| : \mathbb{R} \rightarrow \mathbb{R}$  is a given function respecting the following minimal requirements:

- (R1)  $|\mathcal{A}|$  must be an even convex function, i.e.,  $\forall s \in \mathbb{R}$  then  $|\mathcal{A}|(-s) = |\mathcal{A}|(s)$  respecting  $|\mathcal{A}|(0) = 1$ . The latter condition allows recovery of the definition of the penalty term in the pure diffusive limit, i.e., if  $\text{Pe}_{E,F}^{\theta} = 0$ , then  $\tau_{E,F} := \tau_{\underline{\kappa}_{E,F}}$ . This is in accordance with the first assumption (23) as described in Proposition 4.1.
- (R2) The function  $|\mathcal{A}|$  must respect the following asymptotic behaviors:

$$\lim_{|s| \rightarrow \infty} \frac{|\mathcal{A}|(s)}{s} = \text{sgn}(s), \quad (52)$$

which is in accordance with the parity requirement (R1). It allows recovery of the upwinding-based strategy for purely hyperbolic problems, and consequently the definition of the corresponding fluxes (48) for the advection-dominated regime:

$$\tau_{E,F} := \theta \boldsymbol{\beta} \cdot \mathbf{n}_{E,F} \lim_{|\text{Pe}_{E,F}^{\theta}| \rightarrow \infty} \frac{|\mathcal{A}|(\text{Pe}_{E,F}^{\theta})}{\text{Pe}_{E,F}^{\theta}} = \theta |\boldsymbol{\beta} \cdot \mathbf{n}_{E,F}|. \quad (53)$$

This is in accordance with the second assumption (23) as described in Proposition 4.1.

- (R3) There exists  $\theta_0 \in \mathbb{R}$  such that for all  $\theta \geq \theta_0$  and  $s \in \mathbb{R}$ , then the following holds:

$$|\mathcal{A}|(\theta s) \geq 1 + \frac{|s|}{2}, \quad (54)$$

which is in accordance with the convexity requirement (R1). This argument is crucial to ensure that  $a_h^{(\epsilon)}$  is  $\mathcal{V}_h$ -coercive.

Following the above hypotheses of  $|\mathcal{A}|$ , we can recover some well-known stabilization strategies:

► *The Additive scheme:* By setting  $|\mathcal{A}|_{\text{add}}(s) := 1 + |s|$ , we easily infer that

$$\tau_{E,F} := \tau_{\kappa_{E,F}}(1 + |\text{Pe}_{E,F}^\theta|) = \tau_{\kappa_{E,F}} + \tau_{\beta_{E,F}}^\theta, \quad (55)$$

where  $\tau_{\beta_{E,F}}^\theta := \theta |\boldsymbol{\beta} \cdot \mathbf{n}_{E,F}|$  corresponds to the advective stability parameter chosen in accordance with the  $\theta$ -upwind strategy. Here, we assume that the total stability parameter is simply equal to the sum of its distinctive diffusive and advective contributions, respectively. It is evident that  $|\mathcal{A}|_{\text{add}}$  respects all above criteria (R1-R3) assuming  $\theta \geq 1/2$ .

► *The Scharfetter–Gummel scheme:* By setting  $|\mathcal{A}|_{\text{sg}}(s) := \mathcal{B}(-|s|)$  where

$$\mathcal{B}(s) := \begin{cases} \frac{s}{e^s - 1} & \text{if } s \neq 0, \\ 1 & \text{else.} \end{cases} \quad (56)$$

denotes the (well-known) Bernoulli function. Let us notice that the following holds

$$1 + \frac{|s|}{2} \leq \mathcal{B}(-|s|) \leq 1 + |s|, \quad \forall s \in \mathbb{R}. \quad (57)$$

Thus, by using a scaling argument, we can easily infer that for all  $\theta \in \mathbb{R}$ ,

$$|\mathcal{A}|_{\text{add}}(\theta s/2) \leq |\mathcal{A}|_{\text{sg}}(\theta s) \leq |\mathcal{A}|_{\text{add}}(\theta s), \quad \forall s \in \mathbb{R}, \quad (58)$$

proving immediately the last condition (R3) for any given  $\theta \geq 1$ . In practice, the SG-scheme is particularly interesting since it introduces less artificial diffusion than the Add-scheme in the diffusion-dominated regime.

An illustration of both stabilization functions  $|\mathcal{A}|_{\text{add}}(\theta s)$  and  $|\mathcal{A}|_{\text{sg}}(\theta s)$  is given below in Figure 1 using different admissible values of  $\theta$ .

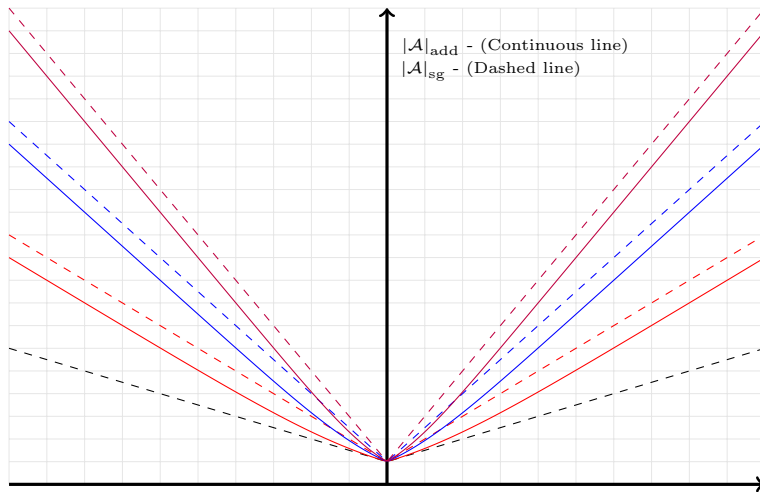


Figure 1: An illustration of both stabilization functions  $|\mathcal{A}|_{\text{add}}(\theta s)$  (dashed-line) and  $|\mathcal{A}|_{\text{sg}}(\theta s)$  (continuous-line) for the set of values  $\theta := \{1/2, 1, 3/2, 2\}$  in black, red, blue and purple, respectively.

## 5. Numerical results

To close this section, we provide some numerical experiments illustrating the robustness and accuracy of the proposed H-IP method for solving degenerate advection-diffusion-reaction model problems. For this aim, we next focus on three distinctive physical situations, namely, nondegenerate, fully degenerate and locally degenerate problems, respectively. All of our numerical experiments are performed using the high-performance finite element library called NGSOLVE [26], and all developed source codes are available for free download from the following Github repository<sup>1</sup>.

### 5.1. Nondegenerate problem

We assume here that both  $\kappa$  and  $\beta$  are nonnull for the whole domain  $\Omega := [0, 1]^2$ . In this first example, as proposed by Egger and Schöberl in [13], we illustrate the ability of the proposed H-IP formalism to deal (efficiently and automatically) with physical processes characterized by a large range of Péclet numbers. Both the velocity field  $\beta$  and the dispersion matrix  $\kappa$  are supposed to be constant on  $\Omega$  – i.e.,  $\beta := (\beta_1, \beta_2)$  with  $\beta_{1,2} \in \mathbb{R}$  and  $\kappa := \kappa \mathbf{I}_d$ , where  $\mathbf{I}_d$  denotes the identity matrix in  $\mathbb{R}^{d,d}$  and  $\kappa > 0$ . The exact solution is given by

$$u(x, y) = [x + (e^{\beta_1 x / \kappa} - 1) / (1 - e^{\beta_1 / \kappa})] \cdot [y + (e^{\beta_2 y / \kappa} - 1) / (1 - e^{\beta_2 / \kappa})], \quad (59)$$

and the right-hand-side  $f$  is chosen such that the exact solution  $u$  is verified. For our numerical study, we then set  $\beta := (2, 1)$ , and we select a large range of  $\kappa := \{5e - 1, 5e - 2, 5e - 3\}$  to control the ratio between the diffusive and advective contributions. The exact solution  $u$  displays sharper fronts on the top and right boundaries of  $\Omega$  as  $\kappa$  becomes smaller. Since the diffusive part is nonnull inside  $\Omega$ , we can distinguish three variants of the H-IP method: the H-IIP, H-NIP and H-SIP schemes, respectively. Here, we analyze the influence of the stabilization strategy, namely, the Add- or SG-schemes as defined in Section 4.4.2, on the behavior of discrete solutions. For clarity of our exposition, we set  $\theta := 1$ . Standard  $h$ - and  $k$ -refinement strategies are used to compute the discrete  $L^2$ -errors and estimated convergence rates (ECRs). A history of convergence of the three variants is presented in Table 1 for different values of  $\kappa := \{5e - 1, 5e - 2\}$  and polynomial degrees  $k := \{1, 2\}$ , and for both Add- and SG-schemes, respectively. First, a brief analysis indicates that the H-IIP and H-NIP schemes behave differently from the H-SIP scheme regardless of the stabilization function. Their convergence orders are (strongly) influenced by the polynomial parity of  $k$ . We observe that the convergence rate is suboptimal (with order  $k$ ) only for even  $k$ , and optimal (with order  $k + 1$ ) for odd  $k$ . The situation is somewhat different for the H-SIP method, since it always converges optimally for all  $k$ . These statements agree with the theoretical results established by Shin *et al.* in [27] (see e.g., [16] for the pure diffusive problem). To pursue our comparative analysis, we then illustrate in Figures 2b and 2d the spatial distribution of the computed  $L^2$ -error on the whole domain for both Add- and SG-schemes. These illustrations indicate that the SG-scheme produces much less artificial diffusion error than the Add-scheme and hence captures sharp fronts much more effectively. Thus, we will favor the H-SIP variant coupled with the Scharfetter-Gummel scheme in elliptic regions in all future experiments.

---

<sup>1</sup><https://github.com/GregoryETANGSALE/HDG-Degenerate-ADR-equation>



$k$	$h^{-1}$	Additive-Upwind				Scharfetter-Gummel			
		$\kappa = 5e-1$		$\kappa = 5e-2$		$\kappa = 5e-1$		$\kappa = 5e-2$	
		$\ u - u_h\ _{0,\mathcal{T}_h}$	ECR	$\ u - u_h\ _{0,\mathcal{T}_h}$	ECR	$\ u - u_h\ _{0,\mathcal{T}_h}$	ECR	$\ u - u_h\ _{0,\mathcal{T}_h}$	ECR
H-NIP									
1	4	$2.2e-03$	–	$9.7e-02$	–	$2.1e-03$	–	$8.3e-02$	–
	8	$5.4e-04$	2.05	$4.1e-02$	1.24	$5.3e-04$	2.02	$3.7e-02$	1.18
	16	$1.3e-04$	2.03	$1.3e-02$	1.72	$1.3e-04$	2.01	$1.2e-02$	1.67
	32	$3.3e-05$	2.02	$3.1e-03$	2.01	$3.2e-05$	2.01	$3.0e-03$	1.97
	64	$8.1e-06$	2.01	$7.4e-04$	2.06	$8.1e-06$	2.00	$7.2e-04$	2.03
2	4	$4.0e-04$	–	$5.1e-02$	–	$4.0e-04$	–	$5.0e-02$	–
	8	$8.8e-05$	2.16	$1.7e-02$	1.61	$8.9e-05$	2.17	$1.7e-02$	1.56
	16	$2.1e-05$	2.05	$4.2e-03$	1.99	$2.1e-05$	2.06	$4.3e-03$	1.98
	32	$5.3e-06$	2.01	$9.4e-04$	2.15	$5.3e-06$	2.01	$9.6e-04$	2.16
	64	$1.3e-06$	2.00	$2.2e-04$	2.09	$1.3e-06$	2.00	$2.2e-04$	2.10
H-IIP									
1	4	$2.2e-03$	–	$9.7e-02$	–	$2.1e-03$	–	$8.3e-02$	–
	8	$5.4e-04$	2.05	$4.1e-02$	1.24	$5.3e-04$	2.02	$3.7e-02$	1.18
	16	$1.3e-04$	2.03	$1.2e-02$	1.72	$1.3e-04$	2.01	$1.2e-02$	1.67
	32	$3.3e-05$	2.02	$3.1e-03$	2.01	$3.2e-05$	2.01	$3.0e-03$	1.97
	64	$8.1e-06$	2.01	$7.4e-04$	2.06	$8.1e-06$	2.00	$7.2e-04$	2.03
2	4	$2.7e-04$	–	$4.6e-02$	–	$2.7e-04$	–	$4.3e-02$	–
	8	$5.2e-05$	2.37	$1.4e-02$	1.75	$5.3e-05$	2.36	$1.3e-02$	1.70
	16	$1.2e-05$	2.14	$3.0e-03$	2.18	$1.2e-05$	2.14	$3.0e-03$	2.15
	32	$2.9e-06$	2.04	$5.9e-04$	2.35	$2.9e-06$	2.04	$5.9e-04$	2.34
	64	$7.2e-07$	2.01	$1.3e-04$	2.23	$7.2e-07$	2.01	$1.3e-04$	2.23
H-SIP									
1	4	$2.1e-03$	–	$6.9e-02$	–	$2.1e-03$	–	$8.3e-02$	–
	8	$5.2e-04$	2.02	$3.3e-02$	1.06	$5.3e-04$	2.03	$3.7e-02$	1.17
	16	$1.3e-04$	2.01	$1.1e-02$	1.59	$1.3e-04$	2.02	$1.2e-02$	1.66
	32	$3.2e-05$	2.01	$2.9e-03$	1.93	$3.2e-05$	2.01	$3.0e-03$	1.97
	64	$8.1e-06$	2.00	$7.2e-04$	2.01	$8.1e-06$	2.00	$7.2e-04$	2.03
2	4	$1.6e-04$	–	$3.5e-02$	–	$1.7e-04$	–	$3.8e-02$	–
	8	$2.1e-05$	2.95	$1.0e-02$	1.78	$2.1e-05$	2.95	$1.1e-02$	1.85
	16	$2.7e-06$	2.99	$1.9e-03$	2.40	$2.7e-06$	2.99	$2.0e-03$	2.44
	32	$3.4e-07$	3.00	$2.8e-04$	2.77	$3.4e-07$	3.00	$2.9e-04$	2.79
	64	$4.2e-08$	3.00	$3.7e-05$	2.93	$4.2e-08$	3.00	$3.7e-05$	2.94

Table 1: Test A - History of convergence of the H-NIP, H-IIP and H-SIP methods on uniform square meshes using the Additive-Upwind and Scharfetter-Gummel scheme with  $\kappa = \{5e-1, 5e-2\}$ .

## 5.2. Fully degenerate problem

In the second example, we analyze the behavior of the H-IP method in the context of hyperbolic problems. Let us notice that the discrete bilinear operator as defined in (18) is reduced to its advective-reactive part since  $\kappa$  is null throughout the whole domain  $\Omega = [0, 1]^2$ . Thus, we set  $\beta := (2, 1)$ ,  $\gamma := 1$  and  $f := 0$ . The following exact solution is prescribed

$$u(x, y) = H(-x + 2y - 1) \quad \text{for all } (x, y) \in \Omega, \quad (60)$$

where  $H(\cdot)$  denotes the Heaviside function. Dirichlet boundary conditions are imposed (only) for the degenerate inflow part  $\Gamma^-$  – i.e., the left and bottom boundaries of  $\Omega$ . The exact

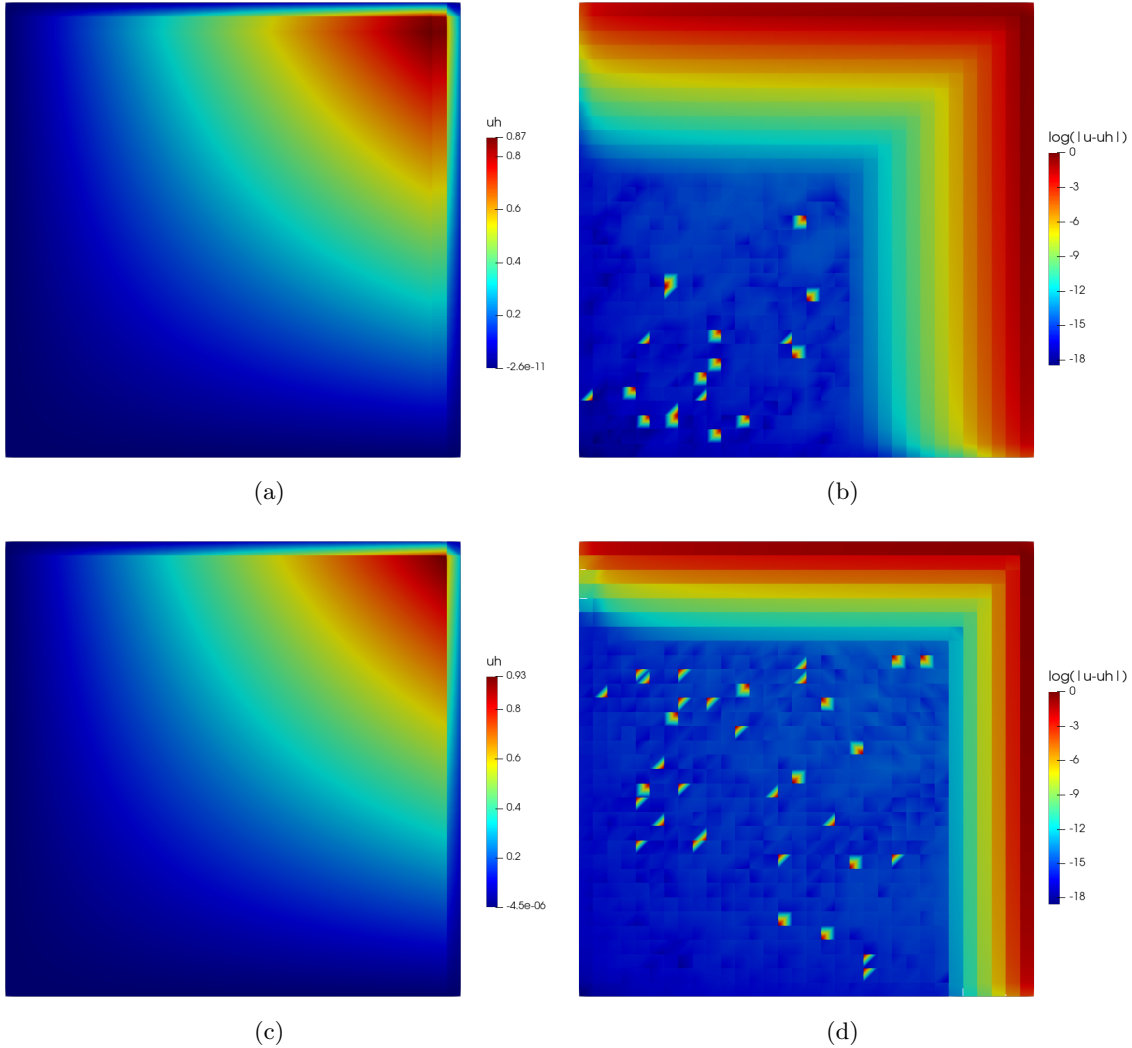


Figure 2: (a) Representation of the linear discrete solution  $u_h$  ( $k = 1$ ) using the Add-scheme for  $\kappa = 5e^{-3}$  and (b) its corresponding  $L^2$ -error spatial distribution. (c) Representation of the linear discrete solution  $u_h$  ( $k = 1$ ) using the SG-scheme for  $\kappa = 5e^{-3}$  and (d) its corresponding  $L^2$ -error spatial distribution.

solution displays a discontinuity along the characteristic direction  $\beta$  due to the jump in imposed boundary conditions at  $x = 0$ . We also assume that the actual location of internal layers is unknown, and an adaptive mesh refinement strategy is investigated to capture them. Let us precisely state that resulting meshes are (generally) not aligned with the characteristic direction  $\beta$ . We analyze here the role of the upwind-parameter  $\theta$  (see e.g., Definition 4.2) with respect to the accuracy of discrete solutions for different polynomial degrees  $k := \{1, 2\}$ . Since the exact solution (60) is only piecewise constant, increasing polynomial degrees can only yield better discrete approximations near internal layers. Following Figure 3, let us notice first that the  $\theta$ -upwind scheme effectively handles outflow boundary conditions for any selected value of  $\theta$ : this observation agrees with Remark 4.4. However, the increase in  $\theta$  significantly deteriorates the discrete solution near the internal layer for all polynomial degrees

since there is more erratic behavior and/or a more considerable artificial numerical diffusion: this observation agrees with Remark 4.3. Thus, we will favor the traditional upwinding scheme obtained by selecting  $\theta = 1$  in hyperbolic regions in all future experiments.

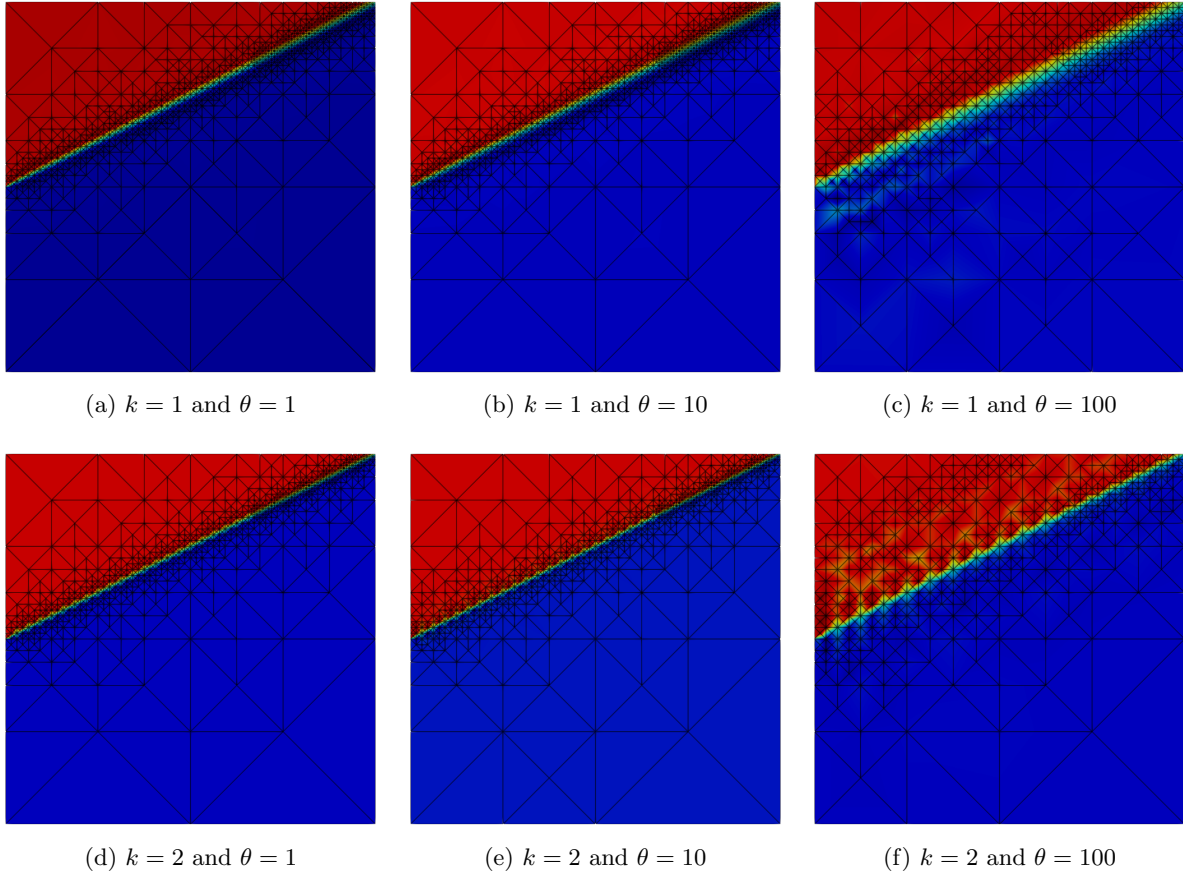


Figure 3: Representation of linear (a)-(b)-(c) and quadratic (d)-(e)-(f) discrete solutions obtained on adaptive meshes for different values of the upwind-parameter  $\theta = \{1, 10, 100\}$ , respectively.

### 5.3. Locally degenerate problem

In the last example, we validated the proposed H-IP formalism in the context of locally degenerate diffusion. To achieve this aim, we consider a slight modification to the test case proposed by Di Pietro *et al.* in [6] (see e.g., [9]). The domain is now taken to be  $\Omega := [-1, 1]^2 \setminus \{x^2 + y^2 < \frac{1}{4}\}$ , which is divided into two disjoint subdomains  $\Omega^{\text{ell}}$  and  $\Omega^{\text{hyp}}$  corresponding to the elliptic and hyperbolic parts, respectively (see e.g., Figure 4a). Denoting by  $(r, \theta)$  the polar coordinates (with azimuth  $\theta$  measured in the anticlockwise sense starting from the positive  $x$ -axis) and by  $\mathbf{e}_\theta$  the (unit) azimuthal vector, the problem coefficients are

$$\kappa = \begin{cases} \pi & \text{if } 0 < \theta \leq \pi, \\ 0 & \text{if } \pi < \theta \leq 2\pi, \end{cases} \quad \boldsymbol{\beta} = \frac{\mathbf{e}_\theta}{r}, \quad \gamma = 1e - 6, \quad (61)$$

and  $\boldsymbol{\kappa} := \kappa \mathbf{I}_2$ . As illustrated in Figure 4b, the exact solution is given by,

$$u(r, \theta) = \begin{cases} (\theta - \pi)^2 & \text{if } 0 < \theta \leq \pi, \\ 3\pi(\theta - \pi) & \text{if } \pi \leq \theta < 2\pi. \end{cases} \quad (62)$$

and is used to infer the forcing term  $f$  and the imposed Dirichlet boundary datum. Here, we focus only on the H-SIP method using the Scharfetter–Gummel scheme in the elliptic region, and the H-IP method using the  $\theta$ -upwind stabilization strategy in the hyperbolic one. For clarity of our exposition, we set  $\theta := 1$  in both regions, but it is quite possible to fix distinctive values in each subdomain. Standard  $h$ - and  $k$ -refinement strategies are used to compute the discrete  $L^2$ -errors and estimated convergence rates (ECRs). A history of convergence is illustrated in Figure 5a for different polynomial degrees  $k := \{1, \dots, 5\}$ . We observe that the convergence rate is optimal with order  $k + 1$  for all  $k$ .

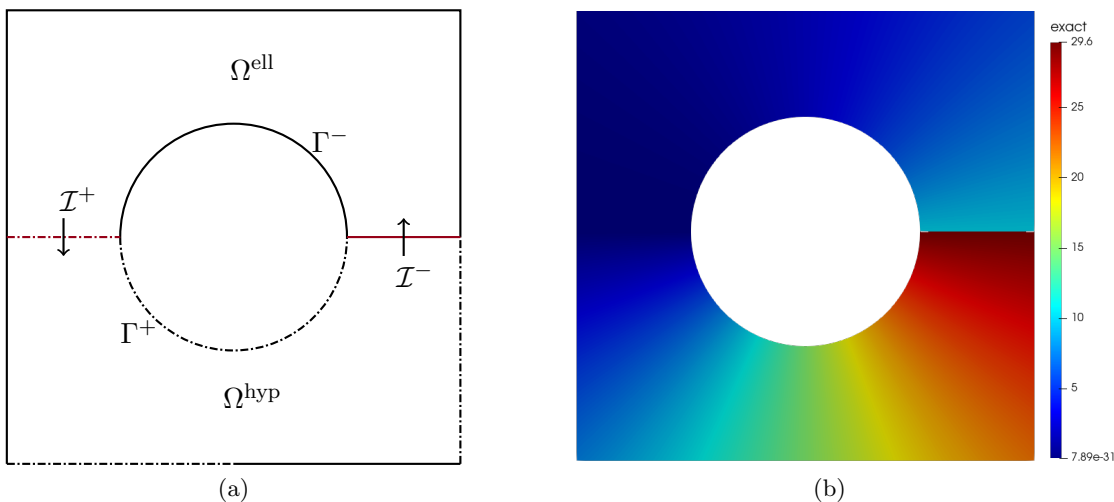


Figure 4: (a) Description of the locally degenerate test case: the continuous black line denotes the nondegenerate inflow boundary  $\Gamma^-$  where Dirichlet conditions are imposed, the dashed black line indicates the degenerate no-flow/outflow boundary  $\Gamma^+$ , and the continuous red line represents the interior interface  $\mathcal{I}^-$  where the exact solution is discontinuous. (b) Illustration of the exact solution (62).

## 6. Conclusion & Perspectives

We have derived a compact interior penalty HDG method for solving degenerate advection-diffusion-reaction problems. The proposed H-IP method can efficiently handle pure diffusive or advective regimes, along with intermediate regimes combining the above mechanisms for a wide range of Péclet numbers, including the delicate situation of local evanescent diffusion. An adaptive stabilization strategy is carried out, automatically accounting for the mathematical nature of (7) and the predominance of the diffusion or advection mechanisms. An upwinding-based scheme was favored for the hyperbolic region, and an inspired Scharfetter–Gummel scheme was preferred for the elliptic region. One undeniable advantage of this strategy is the possibility of tuning the amount of upwind. The stability analysis indicates that all considered variants are consistent and coercive, ensuring the well-posedness of the discrete problem in all regimes. The flexibility and accuracy of the proposed method are confirmed by numerical lines of evidence.

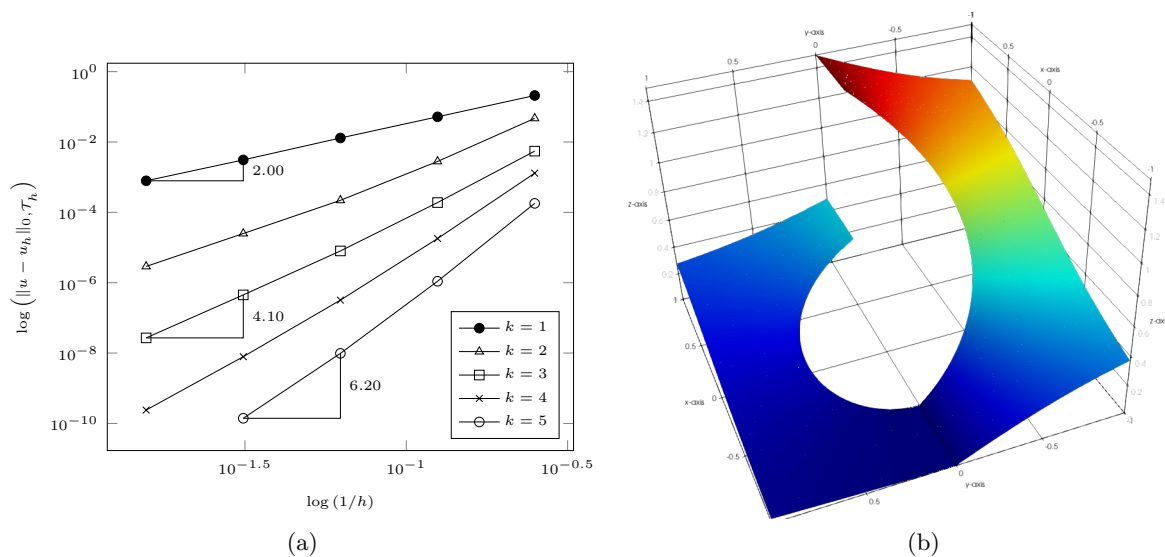


Figure 5: (a) History of convergence error in the  $L_2$ -norm for different polynomial degrees  $k = \{1, \dots, 5\}$ . (b) Representation of  $u_h$  using piecewise linear approximations ( $k = 1$ ) on a fine mesh ( $h = 1/64$ ).

## Acknowledgements

By convention, the names of the authors are listed in alphabetical order. The third author is grateful to Sander Rhebergen for his invitation to the Department of Applied Mathematics at the University of Waterloo (UW) in April 2019. The present work is undoubtedly the fruit of our first discussions and investigations during my visit to the UW.

## References

- [1] Arnold, D.N., Brezzi, F., Cockburn, B., Marini, L.D., 2002. Unified analysis of discontinuous galerkin methods for elliptic problems. *SIAM Journal on Numerical Analysis* 39, 1749–1779. URL: <https://doi.org/10.1137/S0036142901384162>.
- [2] Brezzi, F., Marini, L., Micheletti, S., Pietra, P., Sacco, R., 2006. Stability and error analysis of mixed finite-volume methods for advection dominated problems. *Computers & Mathematics with Applications* 51, 681–696. URL: <https://www.sciencedirect.com/science/article/pii/S0898122106000113>.
- [3] Brezzi, F., Marini, L.D., Süli, E., 2004. Discontinuous galerkin methods for first-order hyperbolic problems. *Mathematical Models and Methods in Applied Sciences* 14, 1893–1903. URL: <https://doi.org/10.1142/S0218202504003866>.
- [4] Cockburn, B., 2016. Static Condensation, Hybridization, and the Devising of the HDG Methods. Springer International Publishing, Cham. chapter 3. pp. 129–177. URL: [https://doi.org/10.1007/978-3-319-41640-3\\_5](https://doi.org/10.1007/978-3-319-41640-3_5).
- [5] Cockburn, B., Di Pietro, D.A., Ern, A., 2016. Bridging the hybrid high-order and hybridizable discontinuous galerkin methods. *ESAIM: M2AN* 50, 635–650. URL: <https://doi.org/10.1051/m2an/2015051>.

- [6] Cockburn, B., Dong, B., Guzmán, J., 2008. A superconvergent ldg-hybridizable galerkin method for second-order elliptic problems. *Mathematics of Computation* 77, 1887–1916. URL: <https://www.jstor.org/stable/40234595>.
- [7] Cockburn, B., Gopalakrishnan, J., Lazarov, R., 2009. Unified hybridization of discontinuous galerkin, mixed, and continuous galerkin methods for second order elliptic problems. *SIAM Journal on Numerical Analysis* 47, 1319–1365. URL: <https://doi.org/10.1137/070706616>.
- [8] Cockburn, B., Qiu, W., Shi, K., 2012. Conditions for superconvergence of hdg methods for second-order elliptic problems. *Mathematics of Computation* 81, 1327–1353. URL: <http://www.jstor.org/stable/23268044>.
- [9] Di Pietro, D.A., Droniou, J., Ern, A., 2015. A discontinuous-skeletal method for advection-diffusion-reaction on general meshes. *SIAM Journal on Numerical Analysis* 53, 2135–2157. URL: <https://hal.archives-ouvertes.fr/hal-01079342>.
- [10] Di Pietro, D.A., Ern, A., 2014. Hybrid High-Order methods for variable diffusion problems on general meshes. *Comptes Rendus Mathématique* 353, 31–34.
- [11] Di Pietro, D.A., Ern, A., Guermond, J.L., 2008. Discontinuous Galerkin Methods for Anisotropic Semidefinite Diffusion with Advection. *SIAM Journal on Numerical Analysis* 46, 805 – 831. URL: <https://hal.archives-ouvertes.fr/hal-01818201>.
- [12] Dijoux, L., Fontaine, V., Mara, T.A., 2019. A projective hybridizable discontinuous galerkin mixed method for second-order diffusion problems. *Applied Mathematical Modelling* 75, 663–677. URL: <https://doi.org/10.1016/j.apm.2019.05.054>.
- [13] Egger, H., Schöberl, J., 2009. A mixed-hybrid-discontinuous galerkin finite element method for convection-diffusion problems. *IMA J. Numer. Anal* 30, 1–2.
- [14] Ern, A., Stephansen, A.F., Zunino, P., 2008. A discontinuous Galerkin method with weighted averages for advection–diffusion equations with locally small and anisotropic diffusivity. *IMA Journal of Numerical Analysis* 29, 235–256. URL: <https://doi.org/10.1093/imanum/drm050>.
- [15] Etangsale, G., Fahs, M., Fontaine, V., Rajaonison, N., 2020. Improved error estimates of hybridizable interior penalty methods using a variable penalty for highly anisotropic diffusion problems. URL: <https://hal.archives-ouvertes.fr/hal-02893064>. (Preprint submitted to CAMWA).
- [16] Fabien, M.S., Knepley, M.G., Riviere, B.M., 2019. Families of interior penalty hybridizable discontinuous galerkin methods for second order elliptic problems. *Journal of Numerical Mathematics* URL: <https://doi.org/10.1515/jnma-2019-0027>.
- [17] Houston, P., Schwab, C., Süli, E., 2002. Discontinuous hp-finite element methods for advection-diffusion-reaction problems. *SIAM Journal on Numerical Analysis* 39, 2133–2163. URL: <http://www.jstor.org/stable/4101054>.

- [18] Kirby, R.M., Sherwin, S.J., Cockburn, B., 2012. To cg or to hdg: A comparative study. *Journal of Scientific Computing* 51, 183–212. URL: <https://doi.org/10.1007/s10915-011-9501-7>.
- [19] Kirk, K.L.A., Horvath, T.L., Cesmelioglu, A., Rhebergen, S., 2019. Analysis of a space-time hybridizable discontinuous galerkin method for the advection-diffusion problem on time-dependent domains. *SIAM Journal on Numerical Analysis* 57, 1677–1696. URL: <https://doi.org/10.1137/18M1202049>.
- [20] Lehrenfeld, C., 2010. Hybrid Discontinuous Galerkin methods for incompressible flow problems. Ph.D. thesis. Diploma thesis, Rheinisch-Westfälischen Technischen Hochschule Aachen.
- [21] Nguyen, N.C., Peraire, J., Cockburn, B., 2009. An implicit high-order hybridizable discontinuous galerkin method for linear convection-diffusion equations. *J. Comput. Phys.* 228, 3232–3254. URL: <https://www.sciencedirect.com/science/article/pii/S0021999109000308>.
- [22] Nguyen, N.C., Peraire, J., Cockburn, B., 2011. An implicit high-order hybridizable discontinuous galerkin method for the incompressible navier-stokes equations. *J. Comput. Phys.* 230, 1147–1170. URL: <https://www.sciencedirect.com/science/article/pii/S0021999110005887>.
- [23] Oikawa, I., 2015. A hybridized discontinuous galerkin method with reduced stabilization. *Journal of Scientific Computing* 65, 327–340. URL: <https://doi.org/10.1007/s10915-014-9962-6>.
- [24] Oleinik, O., Radkevich, E., Radkevic, J., Fife, P., 1973. *Second-Order Equations With Nonnegative Characteristic Form*. Springer US. URL: <https://books.google.com/books?id=fTPvAAAAMAAJ>.
- [25] Raeisi Isa-Abadi, A., Fontaine, V., Ghafouri, H.R., Younes, A., Fahs, M., 2020. A fully interior penalty discontinuous galerkin method for variable density groundwater flow problems. *Computers & Fluids* 213, 104744. URL: <https://www.sciencedirect.com/science/article/pii/S0045793020303145>.
- [26] Schöberl, J., 2014. C++ 11 implementation of finite elements in ngsolve. Institute for Analysis and Scientific Computing, Vienna University of Technology .
- [27] Shin, D., Jeon, Y., Park, E.J., 2015. A hybrid discontinuous galerkin method for advection–diffusion–reaction problems. *Applied Numerical Mathematics* 95, 292–303. URL: <https://www.sciencedirect.com/science/article/pii/S016892741400186X>.
- [28] Wells, G.N., 2011. Analysis of an interface stabilized finite element method: the advection-diffusion-reaction equation. *SIAM Journal on Numerical Analysis* 49, 87–109. URL: <https://www.jstor.org/stable/23074391>.