Les méthodes de pénalisation intérieure hybrides de Galerkine discontinues pour les transferts de masse en milieux poreux

Grégory Etangsale

UNIVERSITÉ DE LA RÉUNION

LAB PIMENT

UFR SHE – Sciences de l'homme et de l'Environnement **ED STS** – Sciences, Technologies et Santé

Les méthodes de pénalisation intérieure hybrides de Galerkine discontinues pour les transferts de masse en milieux poreux

THÈSE

présentée et soutenue publiquement par

Grégory ETANGSALE

pour l'obtention du grade de

Docteur ès Sciences

Spécialité : Mathématiques Appliquées

Composition du jury :

Directeurs :	Vincent FONTAINE	–	Maître de Conférences	Université de La Réunion
	Marwan FAHS	–	Maître de Conférences - HDR	Université de Strasbourg
Rapporteurs :	Thomas GRAF	–	Professeur des Universités	Université de Hanovre
	Abdelaziz RHANDI	–	Professeur des Universités	Université de Saverne
Examinateurs :	Khalid ADDI (Président)	–	Professeur des Universités	Université de La Réunion
	Michel KERN	–	Chargé de Recherche	INRIA, Paris

La Réunion, 7 Décembre 2022

REGION REUNION

l'Europe s'engage à La Réunion avec le FEDER

UNION EUROPÉENNE

Remerciements

Me voici arrivé aux dernières lignes de ce manuscrit, menant à l'aboutissement de ces trois années de thèse. Il m'incombe aujourd'hui d'adresser toute ma gratitude envers les personnes, sans qui, ce travail de recherche ne peut exister...

Tout naturellement, je tiens à remercier profondément M. Vincent Fontaine, Maître de Conférences à l'Université de La Réunion, d'avoir dirigé ce travail de thèse. Sa rigueur, son esprit critique et sa pédagogie sont quelques-unes des nombreuses qualités qui ont rendu nos échanges captivants. À ces démonstrations toujours plus difficiles, auxquels on ne peut pas y échapper. Promis j'y donnerai plus d'énergie ! Malgré celà, je retiendrai la patience et l'écoute dont il a su faire preuve durant les moments les plus difficiles.

Je tiens également remercier Marwan Fahs, Maître de Conférences à l'Université de Strasbourg, qui a su me superviser en dépit des dix-milles kilomètres nous séparant. J'ai été ravi d'être accueilli à l'ITES (Institut Terre et Environnement de Strasbourg) durant deux mois. Bien que ce séjour fut trop court, il a su raviver ma flamme de physicien. Sa curiosité et sa patience m'ont beaucoup apporté. Une dédicace à nos futurs visioconférences durant le week-end.

Que mes directeurs trouvent ici, l'expression de toute ma gratitude et de tout mon respect, pour m'avoir donné le cadre de travail le plus agréable qu'il soit afin de mener à bien ce projet.

Je souhaite maintenant remercier les membres du jury qui me font l'honneur de leur présence, malgré la distance qui nous sépare. En premier lieu, je tiens à remercier M. Thomas Graf, Professeur à l'Université de Hanovre, et M. Abdelaziz Rhandi, Professeur à l'Université de Salerne d'avoir accepté notre invitation, et qui me font l'honneur de rapporter cette thèse. Je suis également reconnaissant que Michel Kern, Chargé de Recherche à l'INRIA, et Khalid Addi, Professeur à l'Université de La Réunion ont accepté de prendre part à ce jury de thèse.

Notamment, je remercie Michel Kern de m'avoir donné l'opportunité de présenter mes travaux à l'INRIA. Au détour d'un bon steak, on a enfin pu se rencontrer et apprendre à se connaître. Mes salutations à M. Vincent Martin, Maître de Conférences à l'UTC, dont on peut dire que ses travaux de recherche m'ont quelque peu fissurés le crâne... À bon entendeur. Merci à vous deux de m'avoir fait l'honneur de participer à mes comités de thèse.

Merci également à ceux qui ont eu le courage de corriger ce manuscrit. Notamment à M. Thierry Mara, pour son aide préciseuse et celà malgré les contraintes de ses enseignements. Mais aussi un grand merci à tous les proches, qui se sont lancés dans la grande aventure de la lecture de ce manuscrit. Mi lé sur ke zot tout la compris ! !

Je remercie également la direction et les membres du laboratoire PIMENT, pour la mise à disposition de tous les moyens physiques nécessaires pour le bon déroulement de cette thèse.

Un chaleureux hommage à tous les doctorants que j'ai pu côtoyer de près ou de loin durant ces longues années. En particulier, aux collègues de l'Open Space, pour ces moments de rires et de détente. J'en profite pour emprunter une célèbre expression du bureau, pour dire qu'on était tous dans la même Galère-(king) !!

Bien évidemment, je tiens à exprimer toute ma gratitude envers toute ma famille, notamment à vous mes parents, de m'avoir accompagné et soutenu dans toutes mes décisions.

Enfin je dédie ces derniers mots, à celle qui aura autant vécu cette thèse que moi, et qui se tient à mes côtés malgré les difficultés. À toi, Mahéva, qui aura su garder toute son attention envers moi comme au premier jour.

À mes Parents
À Mahéva

Résumé

Cette thèse est consacrée à la résolution de l'écoulement et du transport de masse dans les milieux poreux hétérogènes, anisotropes et fracturés par des méthodes Hybrides de Galerkine Discontinues (HDG). Au cours de la dernière décennie, cette méthode connait un succès indéniable auprès de la communauté scientifique. Son formalisme flexible offre de nombreux avantages sur le plan numérique : traitement des discontinuités du milieu ou de la solution du problème, flux localement conservatifs, stratégies de raffinement-h et -k, elligibilité à la condensation statique, et utilisation de maillages non-conformes. L'objectif de ce travail concerne le développement d'une famille de méthodes de Pénalisation Intérieure Hybrides de Galerkine Discontinues (H-IP), pour l'amélioration de la simulation des transferts de masse dans les milieux poreux et fracturés. L'originalité de ce formalisme réside dans la mise en place d'une stratégie de pénalisation adaptative, qui nous permet d'inclure plusieurs situations physiques : (i) les problèmes purement diffusif ou convectif, (ii) les régimes mixtes combinant ces deux processus pour un large éventail du nombre de Péclet, et (iii) les problèmes de diffusion localement évanescente. Enfin, nous présentons l'étude de l'écoulement en milieux poreux fracturés, dans lesquels les fractures sont traitées comme des interfaces de dimension $(d-1)$ immergées dans la roche. Une série d'expériences numériques est réalisée pour évaluer la robustesse, la stabilité et la convergence du schéma pour ces situations physiques délicates. Ces résultats révèlent la souplesse et la pertinence de cette classe pour la modélisation des phénomènes dans les réservoirs naturels.

Mots clés : Milieu poreux fracturé, écoulement, transport de masse, hétérogénéité et anisotropie, nombre de Péclet, méthode hybride de Galerkine discontinue, pénalisation, modèle DFM.

Abstract

This thesis focuses on the numerical modeling of flow and mass transport in heterogeneous, anisotropic, and fractured porous media with Hybrid Discontinuous Galerkin (HDG) methods. During the last decade, this method have received tremendous success from scientific community. Its flexibility provides several benefits regarding the numerical aspects : treatment of discontinuities of the medium and the solution, local conservative fluxes, h- and k-refinement strategies, eligibility to static condensation, and use of non-matching meshes. The objective of this work is to develop families of Hybridizable Interior Penalty (H-IP) schemes, to enhance numerical modeling and simulation of mass transfer in fractured porous media. The originality of the proposed methodology lies in the definition of an adaptative penalty strategy, which allow us to encompass different physical situations : (i) purely diffusive or convective problems, (ii) mixed-regimes combining both processes for a wide range of Péclet number, and (iii) locally evanescent diffusion problems. Finally, we present the study of flow in fractured porous media, in which the fractures are treated as an interface of $(d-1)$-dimension immersed in the bulk. A series of numerical experiments are performed to evaluate the robustness, stability, and convergence of the method for these delicate physical situations. The results prove the flexibility and the suitability of this class for the modeling phenomena in natural reservoirs.

Keywords : Fractured porous media, flow, mass transport, heterogeneity and anisotropy, Péclet number, hybrid discontinuous Galerkin method, penalization, DFM model.

Table des matières

Table des figures

Liste des tableaux

Liste d'abréviations

BO Baumann-Oden 12
BR Bassi-Rebay 12
CI Condition Initiale 9
CL Conditions aux Limites 9
DD Décomposition de Domaine 122
DDL Degrés de Liberté 12
DF Différences Finies 10
DFM Discrete Fracture Matrix 13
EDPs Équations aux Dérivées Partielles 2
EF Éléments Finis 10
DG Galerkine Discontinues 10
HDG Hybrides de Galerkine Discontinues 13
H-IP Pénalisation Intérieure Hybride de Galerkine Discontinue 13
HHO Hybride d'Ordre-Élevé 118
H-LDG Hybride Locale de Galerkine Discontinue 118
IP Pénalisation Intérieure 12
LDG Locale de Galerkine Discontinue 12
VER Volume Elémentaire Représentatif 5
VF Volumes Finis 10

Chapitre 1

Introduction générale

Sommaire

1.1 Cadre général

1.1.1 Contexte de l'étude

Ce manuscrit de thèse est consacré à la simulation des processus d'écoulement et de transport de matières dans les milieux poreux et fracturés. Depuis ces dernières décennies, les enjeux dédiés à l'étude de ces différents phénomènes physiques soulèvent de nombreuses problématiques industrielles, environnementales et sociétales [Diersch 2013, Fetter 2017]. Parmi les diverses applications disponibles, nous retrouvons notamment (*i*) *le stockage et la gestion des déchets toxiques* (lixiviats[1], nucléaire), (*ii*) *la séquestration du carbone*, (*iii*) *la production d'énergie* (géothermie), (*iv*) *la gestion des ressources souterraines* (eau, pétrole, gaz), (*v*) *la contamination des nappes phréatiques* par des sources naturelles (intrusion d'eau salée) et/ou anthropiques (agriculture), pour n'en citer que quelques-unes. Dans les réservoirs naturellement fissurés[2], les questions relatives à la gestion optimale de l'eau sont primordiales. En effet, les eaux souterraines constituent l'une des réserves les plus abondantes et accessibles à la

1. Jus d'ordures dû à la percolation des eaux de pluies.
2. Dans les réservoirs naturels, les fractures sont omniprésentes.

consommation humaine pour de nombreuses régions du monde ; telles que les zones arides, rurales, urbaines et périurbaines. Avec l'accroissement de la population et le changement climatique, cette ressource limitée est continuellement soumise à de nombreux stress anthropiques et/ou naturels pouvant engendrer des risques majeurs de pollution. À ce jour, l'intrusion de l'eau salée dans les nappes phréatiques constitue l'une des menaces les plus préoccupantes pour la qualité des eaux dans les régions franges côtières ([Croucher 1995, Bourhane 2014]). Les résidus des activités industrielles, agricoles et urbaines retiennent également une attention particulière susceptible de constituer des sources de contamination potentielles. Plus précisémment, c'est l'infiltration des eaux de pluie en contact avec ces matières résiduelles, qui peut engendrer un risque de pollution pouvant mettre des décennies à être découvert. Aujourd'hui, ces risques préoccupent les politiques et deviennent un enjeu international pour la gestion durable et pérenne des ressources en eaux souterraines. Dans la région volcanique et insulaire de l'île de La Réunion, ces problématiques sont incontournables en raison de l'inégale répartition spatiale et temporelle de la réserve en eau. À terme, il convient donc de comprendre et quantifier ces différents risques, qui affectent directement ou indirectement le bien-être des populations actuelles et futures.

1.1.2 Les principaux enjeux scientifiques

L'étude couplée de l'écoulement et du transport de masse dans les aquifères[3] reste à l'heure actuelle une tâche trop difficile et délicate pour être menée uniquement par des méthodes experimentales traditionnelles. Ce sont des processus de longues durées (années-décennies) intervenant sur des réservoirs naturels de grandes dimensions (de l'ordre du m-km). Dans la nature, l'eau circule dans ces milieux fracturés, dans lesquels les fractures forment un réseau complexe et inter-connecté dont les propriétés hydrauliques diffèrent de celles du milieu d'accueil [Adler 1999, Dietrich 2005, Koohbor 2020a]. La présence intrinsèque de ces *discontinuités*[4] dans la roche affecte considérablement les transferts en intéraction [Heße 2013], où les fractures agissent comme des voies préférentielles ou obstacles à l'écoulement du fluide [Berkowitz 2002, Martin 2005, Angot 2009]. De nos jours, ces phénomènes physiques sont largement connus des physiciens et mathématiciens sous la forme de *lois de conservation*, qui sont formalisées par des Équations aux Dérivées Partielles (EDPs).

En raison de la complexité de ces lois de comportement, la détermination de solutions analytiques ne peut se limiter qu'aux situations les plus simples. Afin de prédire l'évolution de ces différents phénomènes dans le temps, l'approche numérique apparaît comme la solution la plus économique et adaptée pour décrire ces comportements le plus fidèlement possible. La modélisation se révèle être un outil essentiel pour la simulation de modèles toujours plus réalistes en facilitant la manipulation des géométries complexes, la prise en compte des propriétés physiques de l'aquifère, et le traitement *a-posteriori* des données. Généralement, le choix de la stratégie numérique employée est conditionné par la nature mathématique des équations à résoudre [Fontaine 2008, Dijoux 2019a] : *hyperbolique* (pour l'opérateur de convection) ou *elliptique* (pour l'opérateur de diffusion-dispersion). À titre d'exemple, nous

3. L'aquifère fait référence aux formations souterraines composées de roches perméables, naturellement fissurées, pouvant retenir des quantités considérable d'eau et permettre son déplacement [Adler 1999]

4. En hydrogéologie, les failles ou fissures sont considérées comme des discontinuités dans la roche.

pouvons citer : (*i*) les transferts de diffusion fortement hétérogènes et anisotropes ou (*ii*) de convection pure, et (*iii*) les problèmes mixtes combinant à la fois ces deux processus (e.g., transport de matières) pour un large éventail du nombre de Péclet.

Bien évidemment, la nécessité de disposer d'outils numériques performants, qui respectent les caractéristiques physiques du problème étudié est essentiel dans l'obtention de solutions réalistes [Di Pietro 2014]. Cependant, la simulation fait face à de nombreux obstacles, d'ordre purement numérique, nécessitant *a priori* une réflexion sur le choix et la création d'une méthode de discrétisation adéquate. En effet, la simulation à grande échelle (en 3D) des processus couplés d'écoulement et de transport de masse exigent une architecture numérique flexible pour réduire les temps CPU et le stockage mémoire. Des difficultés supplémentaires s'ajoutent lorsque la nature mathématique des équations varie d'une région à l'autre de l'espace, i.e., *hyperbolique* dans une portion du domaine et *elliptique* dans le reste. Cette situation préjudiciable apparait notamment lorsque la diffusivité du milieu s'estompe dans une portion du domaine, on parle alors de diffusion localement *évanescente*. Dans ce contexte, la dégénerescence du paramètre de diffusion entraine une conséquence fondamentale dans l'approximation de la solution, qui peut être discontinue le long de l'interface *hyperbolique-elliptique*. Cette situation physique se manifeste naturellement lors du transport de matière au sein d'un milieu fissuré, où le mécanisme de convection peut devenir prédominant dans les fractures. Il apparait clairement que la résolution couplée des équations elliptique et hyperbolique représente l'un des verrous majeurs de ce travail de recherche.

L'identification d'une approche numérique adaptée reste une tâche délicate, qui nécessite de garantir plusieurs exigences. Tout d'abord, il s'agit de mettre en oeuvre des techniques numériques robustes et stables, vérifiant la convergence optimale de la solution du problème. Il faut être conscient que la précision du calcul est souvent tributaire de la méthode de discrétisation employée. En effet, l'efficacité du modèle est souvent une source de dépense numérique (dimension des systèmes matriciels), pouvant excéder la capacité des ordinateurs actuels (vitesse du processeur, mémoire vive disponible). Le défi majeur de ce travail de recherche consiste à proposer une méthode de résolution numérique, qui offre le meilleur compromis en terme de précision et temps de calcul pour une large classe de phénomènes physiques.

Dans ce premier chapitre introductif, nous abordons la description des mécanismes régissant l'écoulement et le transport de masse au sein d'un réservoir naturel. Notamment, nous rappelons les principales notions élémentaires caractérisant le milieu poreux et les fractures. Bien que l'étude détaillée des transferts de masse en milieux poreux pourrait faire l'objet de plusieurs ouvrages, nous nous limiterons à une description concise de ces différentes notions. Le lecteur pourra se référer à [Bear 2012] (voir également [Chavent 1986, Dietrich 2005]) pour une étude plus détaillée. De part la complexité des lois de conservation, nous priviliégions dans ce manuscrit de thèse l'approche numérique pour la résolution de ces EDPs. Ainsi, il convient de rappeller les principales orientations et hypothèses méthodologiques qui seront utilisées au cours de ce travail de recherche.

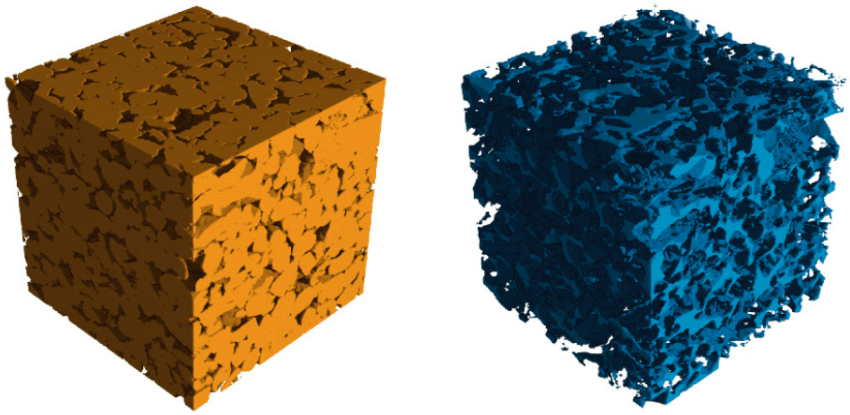
1.2 Hydrodynamique dans les milieux poreux

1.2.1 Le milieu poreux

Par définition, un milieu poreux est une matière qui se compose de pores interconnectés de faible dimension par rapport à l'échelle de l'échantillon étudié [Bear 1988]. La matrice solide représente la portion du squellete comprenant de la matière solide et délimitant les pores - ou volume de vide. On parle alors de *porosité*, pour la capacité du milieu à stocker du fluide et permettre son déplacement [Bear 2018]. La porosité, notée ϕ, donne la proportion du volume de vide sur le volume total du milieu,

$$\phi = \frac{V_p}{V_p + V_s}, \tag{1.1}$$

où V_p et V_s désignent le volume des pores et du solide, respectivement (voir Figure 1.1). Lorsque le milieu poreux est fissuré, la porosité se décompose en deux classes distinctes : *(i)* la porosité d'interstice et *(ii)* la porosité de fissures, qui coexistent naturellement dans les formations géologiques [Adler 1999]. La perméabilité constitue une seconde propriété importante d'un corps poreux. Elle mesure l'aptitude du matériau à se laisser traverser par un fluide et dépend des caractéristiques physiques du milieu [Bear 2012]. Par la suite, nous faisons l'hypothèse que les pores du milieu sont entièrement remplis par un seul fluide monophasique. Dans cette situtation, nous dirons du milieu qu'il est saturé.



(a) (b)

FIGURE 1.1 – Exemple de la structure du grès obtenue par microtomographie à rayon X [Gong 2020] : Volume solide (a) et volume des pores (b).

Dans le but de décrire les différents échanges physiques intervenant dans la roche, tels que l'écoulement du fluide et le transport de matière, nous devons caractériser la distribution des vides et fissures, à toutes les échelles d'observation. Tout d'abord, à l'échelle microscopique (ou des pores), la répartition spatiale et la géométrie des grains [5] sont des paramètres indispensables pour représenter fidèlement le système étudié. Cette description est théoriquement

5. Les grains constituent la partie solide du milieu à l'échelle des pores.

possible, mais irréaliste pour des applications à grande échelle compte tenu de la forte hétérogénéité de la distribution des pores. À l'échelle macroscopique (visible à l'oeil nu), la matrice solide et la fraction de vide sont volontairement confondues. Le système est remplacé par un milieu poreux *équivalent* ou Volume Elémentaire Représentatif (VER), dont les propriétés locales sont moyennées et *continues* dans le VER. La Figure 1.2, qui illustre les variations des propriétés du milieu en fonction de la taille de l'échantillon, montre que le concept de VER n'a de sens que pour des échelles d'observation suffisamment grandes. Des difficultés supplémentaires s'ajoutent lorsque le milieu est fracturé. En effet, les micro-fractures sont généralement moyennées dans la matrice, tandis que les fractures dominantes peuvent être traitées comme un VER indépendant.

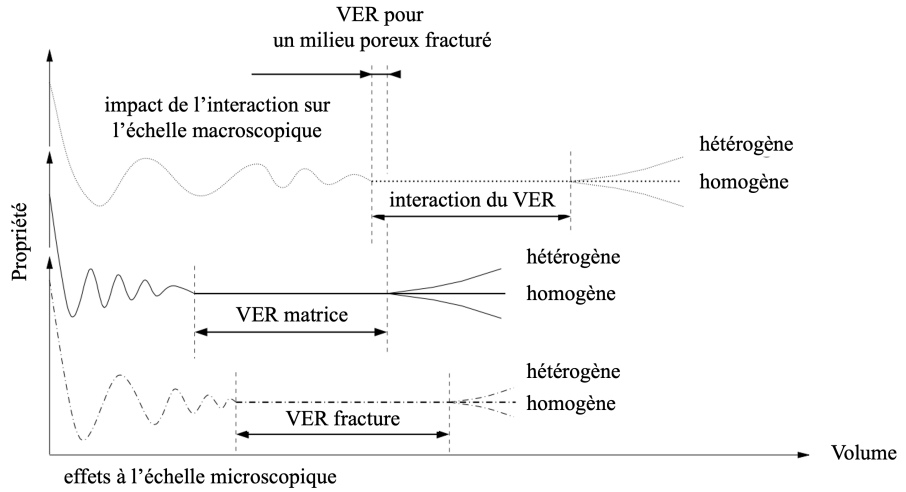


FIGURE 1.2 – Caractéristiques physiques du VER, modifiées par [Schwenck 2015].

En réalité, la plupart des matériaux naturels et artificiels sont considérés comme poreux. Le Tableau 1.1 répertorie des exemples de porosité et perméabilité pour divers matériaux.

Matériau	**Porosité** [%]	**Perméabilité** [6] [mD]
Brique	$12 - 34$	$5 - 220$
Granite	$0.0005 - 1$	$500 - 2000$
Calcaire (dolomite)	$4 - 10$	$2 - 45$
Calcaire (fine)	$25 - 32$	$20 - 200$
Sable	$37 - 50$	$20000 - 180000$
Grès	$8 - 38$	$0.5 - 3000$
Schiste	$0.5 - 5$	$0 - 100$
Terre	$43 - 54$	$300 - 14000$

TABLE 1.1 – Valeurs de porosité et perméabilité pour des matériaux courants, tirés de [Lampe 2013].

Dans la prochaine section nous introduisons les modèles physiques régissant l'écoulement

6. Ici, l'unité utilisée pour la perméabilité est le Darcy [D], mais en pratique nous adopterons le [m^2] dont la conversion est donnée par $1D \approx 9.869 \cdot 10^{-12}$ m^2.

du fluide et le transport de matières dans un milieu poreux. La discussion se concentre sur l'établissement des relations mathématiques qui seront étudiées dans ce manuscrit de thèse.

1.2.2 Écoulement du fluide

Les lois traditionnelles régissant l'écoulement d'un fluide dans un milieu poreux reposent sur le couplage de deux relations fondamentales [Ledoux 2003]. La première exprime la conservation de la matière, dont la construction s'appuie sur le bilan massique du fluide,

$$\frac{\partial(\phi\rho)}{\partial t} + \nabla \cdot (\rho \overrightarrow{q}) = \rho Q, \tag{1.2}$$

doté des grandeurs macroscopiques suivantes : $\phi[-]$ la porosité du milieu, $\rho[M \cdot L^{-3}]$ la masse volumique du fluide, $\overrightarrow{q}[L \cdot T^{-1}]$ la vitesse du fluide traversant le milieu, $Q[T^{-1}]$ un débit volumique représentant un terme source et/ou puits, et $t[T]$ le temps. La seconde relation décrit la vitesse de l'écoulement dans le milieu poreux par la célèbre loi de Darcy. Découverte expérimentalement en 1856 par Henry Darcy [Darcy 1856], elle exprime la densité du flux d'un fluide $\overrightarrow{q}$ à travers la matrice poreuse en fonction du gradient de pression $p[M \cdot L^{-1} \cdot T^2]$,

$$\overrightarrow{q} = -\frac{\overline{\overline{k}}}{\mu} \left(\nabla p + \rho g \nabla z\right), \tag{1.3}$$

où $\overline{\overline{k}}[L^2]$ est le tenseur de perméabilité intrinsèque du milieu, $\mu[M \cdot L^{-1} \cdot T^{-1}]$ la viscosité dynamique du fluide, $z[L]$ la hauteur du fluide et $g[L \cdot T^{-2}]$ l'accélération de la gravité. En pratique, la loi de Darcy est valide pour des écoulements laminaires, dont les valeurs critiques du nombre de Reynolds sont comprises entre 1 et 10 [Chauveteau 1967, Alabi 2011].

Par la suite, nous supposerons que la matrice solide est rigide mais faiblement déformable, [Bear 2018]. En tenant compte de l'hypothèse d'élasticité de l'eau et de la roche, on peut associer la compressibilité du fluide c_ρ et du solide c_ϕ à la pression p de sorte que :

$$\frac{\partial(\phi\rho)}{\partial t} = \rho \left[\phi c_\rho + (1-\phi) c_\phi\right] \frac{\partial p}{\partial t} = \frac{s}{g} \frac{\partial p}{\partial t}, \tag{1.4}$$

avec $c_\rho = \rho^{-1} \partial\rho/\partial p$ et $c_\phi = (1-\phi)^{-1} \partial\phi/\partial p$. Ici, l'*emmagasinement spécifique* $s[L^{-1}]$ exprime la capacité du milieu poreux à libérer un fluide sous l'effet d'une variation de pression.

Dans la plupart des configurations, il est possible de négliger les variations temporelles [7] et spatiales de la masse volumique, i.e., fluide incompressible et/ou écoulement sans transport (voir [Diersch 2013, Bear 2018] pour des exemples d'applications). En remplaçant l'expression de la presssion macroscopique par la hauteur piézométrique, $h = p/(\rho g) + z$, la loi de Darcy (1.3) et l'équation de continuité (1.2) s'expriment en hydrogéologie par,

$$\overrightarrow{q} = -\overline{\overline{K}} \nabla h, \tag{1.5}$$

$$s \frac{\partial h}{\partial t} + \nabla \cdot \overrightarrow{q} = Q, \tag{1.6}$$

7. La dérivée temporelle de la pression s'exprime par $\frac{\partial p}{\partial t} = \rho g \frac{\partial h}{\partial t}$, et l'expression du gradient hydraulique devient $\rho g \nabla h = \nabla p + \rho g \nabla z$.

où $\overline{\overline{K}} = (\rho g/\mu)\overline{\overline{k}}$ désigne le tenseur de conductivité hydraulique qui dépend des propriétés physiques du milieu et du fluide [Chavent 1986] (voir aussi [Younes 2003, Ackerer 2004]). En l'absence du terme source, on décrit l'écoulement du fluide en régime permanent (ou stationnaire) par la conservation de la masse,

$$\nabla \cdot \overrightarrow{q} = 0. \tag{1.7}$$

Cette situation particulière se retrouve pour un fluide incompressible, lorsque la variation temporelle de la hauteur piézométrique est nulle ($\partial/\partial t = 0$).

1.2.3 Le transport de solutés

Dans le milieu poreux, le mécanisme de transport s'opère avant tout par les déplacements du fluide, qui emporte des quantités de matières dissoutes lors de sa migration [Fetter 2017]. C'est un phénomène complexe, que l'on retrouve notamment lors de l'infiltration de l'eau salée dans les nappes phréatiques côtières ([Croucher 1995, Bourhane 2014]), ou encore dans la propagation de polluants dans les sols [Diersch 2013, Fetter 2017]. Dans la suite, nous envisagerons que le transport de solutés [8] se construit autour de deux mécanismes physiques distincts : la *convection* et la *dispersion*.

1.2.3.1 La convection

Le phénomène de convection matérialise l'entrainement d'une espèce en solution par le mouvement du fluide [Rubin 1983]. Dans notre cas, l'équation de transport purement convective est donnée par l'expression

$$\phi \frac{\partial c}{\partial t} + \nabla \cdot (\overrightarrow{q}\, c) = 0, \tag{1.8}$$

avec $\overrightarrow{q}$ la vitesse de Darcy, et $c[M \cdot L^{-3}]$ la concentration en soluté.

1.2.3.2 La dispersion

Le mécanisme de dispersion est un phénomène physique qui se manifeste lorsqu'il existe un contraste du champ de concentration du soluté [Ledoux 2003, Bear 2018]. Sous l'action de l'écoulement du fluide, les matières dissoutes s'étalent progessivement dans le milieu poreux. Il apparaît notamment lors des processus de diffusion moléculaire et de dispersion cinématique :

- La **diffusion moléculaire** prend son origine du mouvement brownien. Il se traduit par l'homogénéisation de la concentration du soluté, qui s'étale et se disperse des zones fortement concentrées vers les zones faiblement concentrées [Bear 1988]. On l'observe toujours en l'absence d'écoulement.
- La **dispersion cinématique** est causée par la variation du champ de vitesse (de Darcy) s'écoulant à travers le milieu [Chavent 1986].

8. On parle de solutés pour les substances chimiques susceptibles de se dissoudre dans l'eau.

Généralement, le transport par dispersion et diffusion est formulé par la loi fickienne

$$\phi\frac{\partial c}{\partial t} + \nabla\cdot(-\overline{\overline{D}}\,(\overrightarrow{q})\,\nabla c) = 0, \tag{1.9}$$

munie du tenseur symétrique de dispersion-diffusion $\overline{\overline{D}}\,(\overrightarrow{q})$, qui s'exprime généralement de la façon suivante (voir [Koohbor 2020a]) :

$$\overline{\overline{D}}\,(\overrightarrow{q}) = \underbrace{\phi D_m \overline{\overline{I}}_d}_{\text{dif. mol.}} + \underbrace{\alpha_T\,|\overrightarrow{q}|\,\overline{\overline{I}}_d + (\alpha_L-\alpha_T)\frac{\overrightarrow{q}\otimes\overrightarrow{q}}{|\overrightarrow{q}|}}_{\text{disp. cin.}} \tag{1.10}$$

où $D_m[L^2\cdot T^{-1}]$ désigne le coefficient de diffusion moléculaire, $\overline{\overline{I}}_d$ le tenseur identité, $\alpha_L[L]$ et $\alpha_T[L]$ les coefficients de dispersivité longitudinale (dans le sens de l'écoulement) et transerversale (orthogonale à l'écoulement) du milieu, respectivement.

1.2.3.3 L'équation de convection-dispersion-réaction

Le modèle mathématique décrivant le transport de solutés est dit *conservatif* en l'absence *(i)* de terme source, *(ii)* d'intéractions entre la matrice solide et le fluide, et *(iii)* de dégradation chimique. En combinant les différents mécanismes de propagation (i.e., convection, diffusion moléculaire et cinématique), ce phénomène s'exprime sous la forme de l'équation de convection-dispersion généralisée

$$\phi\frac{\partial c}{\partial t} + \nabla\cdot(-\overline{\overline{D}}(\overrightarrow{q})\nabla c + \overrightarrow{q}\,c) = 0. \tag{1.11}$$

Il est possible de modifier l'équation (1.11) en ajoutant un opérateur de réaction (notée R), qui comporte des processus chimiques, biologiques ou radioactifs [Diersch 2013, Bear 2018] :

$$\phi\frac{\partial c}{\partial t} + \nabla\cdot(-\overline{\overline{D}}(\overrightarrow{q})\nabla c + \overrightarrow{q}\,c) + R(c) = 0. \tag{1.12}$$

La relation (1.12) est également connue comme l'équation de convection-dispersion-réaction.

En vue de faciliter les notations, nous représentons désormais les grandeurs physiques de nature vectorielle et tensorielle par des caractères gras.

1.3 Modèles mathématiques des lois de conservation

1.3.1 Équation générique

Les modèles physiques régissant l'écoulement et le transport de matières font partie de la grande famille des lois de conservation. En régime transitoire, ces transferts se décomposent sous la forme d'une équation générique de convection-diffusion-réaction :

$$\varphi\frac{\partial u}{\partial t} + \nabla\cdot(-\boldsymbol{\alpha}\nabla u + \boldsymbol{\beta}u) + M(u) = f \quad \text{dans } \Omega\times[0,T], \tag{1.13}$$

où $u(\boldsymbol{x}, t)$ représente la variable d'état, φ le terme de stockage, $\boldsymbol{\beta}$ le champ des vitesses, μ le coefficient de réaction provenant de la décomposition linéaire $M(u) = \mu u$, et f un terme source/puits. Le tenseur de diffusion $\boldsymbol{\alpha} \in \mathbb{R}^d \times \mathbb{R}^d$ est un tenseur symétrique et définit positif, potentiellement anisotrope [9] et/ou hétérogène,

$$\boldsymbol{x}^t \cdot \boldsymbol{\alpha} \cdot \boldsymbol{x} > 0, \quad \forall\, \boldsymbol{x} \in \Omega, \quad \forall\, \boldsymbol{x} \in \mathbb{R}^d, \tag{1.14}$$

dont les coefficients sont donnés suivant les orientations spatiales du milieu. Par analogie avec les modèles d'écoulement (1.5-1.6) et de transport (1.12), l'ensemble des paramètres mathématiques apparaissant dans l'équation (1.13) sont traduits par les grandeurs physiques listées dans le Tableau 1.2.

Modèle mathématique	Écoulement	Transport de solutés
u	h	c
$\boldsymbol{\alpha}$	$\overline{\overline{K}}$	$\overline{\overline{D}}(\overrightarrow{q})$
$\boldsymbol{\beta}$	$\overrightarrow{0}$	$\overrightarrow{q}$
μ	0	r
φ	s	ϕ
f	Q	0

TABLE 1.2 – Description des grandeurs physiques des modèles d'écoulement et de transport de solutés

Pour délimiter le problème (1.13), il est essentiel d'énoncer l'ensemble des conditions décrivant l'état initial dans le milieu et celles imposées sur sa frontière. D'une part, les Conditions aux Limites (CL) imposent les valeurs qui sont attribuées à la solution sur les bordures du domaine $\partial\Omega$. Parmi les contraintes les plus fréquemment utilisées, nous adopterons les conditions de Dirichlet et de Neumann définies sur les portions Γ_D et Γ_N, respectivement, avec $\partial\Omega = \Gamma_D \cup \Gamma_N$ et $\Gamma_D \cap \Gamma_N = \emptyset$, de sorte que :

- La **condition de Dirichlet** impose la valeur de la variable d'état u par,

$$u(\boldsymbol{x}, t) = \bar{u}(\boldsymbol{x}, t) \quad \text{sur } \Gamma_D \times [0, T],$$

- La **condition de Neumann** impose la valeur de la composante normale du flux total $\boldsymbol{\sigma} := -\boldsymbol{\alpha}\nabla u + \boldsymbol{\beta} u$ par,

$$\boldsymbol{\sigma}(\boldsymbol{x}, t) \cdot \boldsymbol{n} = \bar{\sigma}(\boldsymbol{x}, t) \quad \text{sur } \Gamma_N \times [0, T],$$

  où $\boldsymbol{n}$ est la normale unitaire pointant vers l'extérieur du domaine Ω. Le cas particulier $\bar{\sigma}(\boldsymbol{x}, t) = 0$ traduit l'absence d'écoulement sur la frontière de Neumann.

D'autre part, la Condition Initiale (CI) décrit l'état de la variable u à l'instant initial, $t = 0$, dans le domaine Ω,

$$u(\boldsymbol{x}, 0) = u_0(\boldsymbol{x}) \quad \text{dans } \Omega \times \{0\}.$$

Dans le cas d'un écoulement en régime permanent, les variables du système sont indépendantes du temps. Les notations correspondantes seront donc retirées, i.e., $u(\boldsymbol{x}, t) = u(\boldsymbol{x})$.

9. En hydrogéologie, on parle d'hétérogénéité pour les variations de perméabilité des couches sédimentaires. L'anisotropie apparait lorsque la conductivité varie également suivant les directions de l'espace.

1.3.2 Résolution numérique des lois de conservation

Les enjeux environnementaux de ce travail de recherche soulignent clairement les besoins d'établir un modèle numérique optimal visant à reproduire le plus fidèlement possible les phénomènes physiques précédemment exposés. Généralement, le choix de la méthode de discrétisation va dépendre de plusieurs paramètres : *(i)* la nature mathématique des équations à résoudre, *(ii)* les variations du milieu telles que l'hétérogénéité entre les différerentes couches géologiques et les forts ratios d'anisotropie, et *(iii)* la géométrie du problème [Di Pietro 2014]. Aujourd'hui, la demande croissante d'outils numériques plus fiables et robustes vise à intégrer davantage de complexité dans les phénomènes simulés. Face à la multitude de méthodes proposées dans la littérature, il convient de rappeler les principales exigences que devra garantir le modèle numérique qui sera développé dans ce manuscrit :

- La *conservation locale* de la masse, qui est une condition essentielle en hydrogéologie ;
- La *continuité* de la variable d'état u et de la composante normale du flux $\boldsymbol{\sigma} \cdot \boldsymbol{n}$ sur la frontière des éléments de la partition du domaine ;
- L'*approximation optimale* des variables u et $\boldsymbol{\sigma}$, pour limiter les erreurs lors du passage de l'écoulement au transport de solutés ;
- Une *adaptabilité* aux discontinuités du milieu poreux (e.g., hétérogénéité/anisotropie et présence de fractures) ;
- Une *souplesse* dans la manipulation des géométries complexes, et une *flexibilité* dans le traitement des mailles conformes ou non-conformes ;
- Un *faible coût numérique* en terme de temps CPU (i.e., système linéaire compact disposant d'une matrice creuse) ;
- L'*intégrabilité* dans les simulateurs déjà existants.

Au cours des dernières décennies, ce sont principalement les méthodes eulériennes qui se sont imposées dans la résolution des lois de conservation en garantissant la plupart de ces exigences. Parmi les modèles les plus connus, nous mentionnons la méthode des Différences Finies (DF) [Meerschaert 2004], des Volumes Finis (VF) [Eymard 2000], et des Éléments Finis (EF) [Diersch 2013]. Les méthodes de Galerkine Discontinues (DG) [10] font l'objet d'un intérêt croissant depuis la fin du XX^{e} siècle ([Arnold 2002]).

La méthode des Différences Finies

Considérée comme la plus ancienne ([Ramasomanana 2012]), la méthode DF fut fortement employée pour la modélisation des transferts de masse dans les réservoirs. Facile d'implémentation, cette méthodologie repose sur le calcul du bilan massique sur chaque maille de la partition [LeVeque 1998, Meerschaert 2004]. Bien que la méthode DF soit localement conservative, ce formalisme reste inadapté aux géométries les plus complexes, car elle repose sur l'utilisation d'un maillage régulier constitué d'éléments rectangulaires. Par ailleurs, la méthode par Différences Finies se retrouve fortement limitée lorsque le tenseur de perméabilité

10. L'acronyme DG fait directement référence au terme anglo-saxon *Discontinuous Galerkin*, couramment utilisé dans la littérature.

est plein et discontinu. En présence d'un mécanisme de convection dominant, la manipulation de l'opérateur hyperbolique provoque une dégradation non-négligeable de la qualité des résultats [Meerschaert 2004, Ramasomanana 2012].

La méthode des Volumes Finis

Tout comme le schéma DF, la méthode VF établit le bilan massique moyen sur un volume de contrôle, tout en garantissant la conservation locale de la masse et la continuité du flux [Eymard 2000, Fontaine 2008, Di Pietro 2014]. Cette classe d'approximation fait partie des schémas de bas degrés, où la variable recherchée est constante par maille. En fonction du choix de l'approximation du flux à l'interface des éléments, nous pouvons recenser une variété de schémas VF, e.g., Cell-Centered, Vertex-centered, TPFA (pour Two-Point Flux Approximation), et MPFA (pour Multi-Point Flux Approximation) [Eymard 2000, Barth 2003]. L'avantage d'une formulation VF est principalement d'ordre géométrique, puisqu'elle permet d'inclure naturellement les discontinuités du milieu (hétérogénéité/anisotropie). Le traitement de l'opérateur de convection connait de nombreux développements, conduisant notamment aux schémas *upwind* ([Younes 2003, Lehrenfeld 2010] et *Scharfetter-Gummel*. Même si les méthodes VF disposent d'une structure matricielle compacte, le comportement de la solution peut devenir instable en présence de forts ratios d'anisotropie. Pour lutter contre ces oscillations indésirables, on emploie généralement des stratégies de raffinement-h [11] pour le maillage. Cependant, l'augmentation du nombre de mailles peut saturer le stockage mémoire et accroître le temps CPU pour des systèmes complexes de grandes dimensions. Néanmoins, les méthodes VF restent encore à ce jour l'une des classes d'approximation les plus utilisées dans le secteur industriel [Ramasomanana 2012].

La méthode des Éléments Finis

Contrairement aux schémas DF et VF, la méthode des Éléments Finis va chercher à construire une approximation polynomiale de la solution sur des volumes élémentaires. On parle alors de *formulation faible* [Ern 2004, Diersch 2013]. Cette approche se distingue par la nature polynomiale de la solution, dans laquelle la continuité de la variable d'état est fixée par les fonctions de base à l'interface des éléments [Sun 2009, Kirby 2012, Lin 2014, Lin 2015]. La méthode EF se révèle être particulièrement adaptée aux problèmes hétérogènes et anisotropes. Elle facilite l'utilisation des stratégies de raffinement-h et -k [12], et s'adapte aisément aux géométries les plus complexes en négligeant les contraintes liées au maillage. L'inconvénient majeur de cette méthodologie réside dans l'approximation du flux, qui est localement non-conservatif [Fontaine 2008, Sun 2009, Diersch 2013, Lin 2014]. Cette limitation peut entraîner des erreurs non-négligeables lors de la simulation couplée de l'écoulement et du transport, en ajoutant de la diffusion numérique et/ou des oscillations le long des discontinuités du milieu [Ramasomanana 2012]. Généralement, nous avons recours à des procédures de reconstruction *a posteriori* de la vitesse, qui doivent faire l'objet d'une étude à part entière [Cockburn 2007].

11. Le raffinement-h consiste à fractionner finement le maillage en augmentant le nombre d'éléments dans la partition. De ce fait, la constante positive h désigne la dimension des mailles de la partition.
12. Le raffinement-k consiste à augmenter le degré d'approximation polynomial k lors de la résolution.

La méthode de Galerkine Discontinue

Les méthodes DG se distinguent du schéma EF en relaxant entièrement les contraintes liées à la régularité de la solution sur le squelette de la partition (i.e., entre les éléments du maillage) [Arnold 2002, Peraire 2008, Rivière 2008, Di Pietro 2011]. Cette classe d'approximation a été initialement introduite en 1973 par [Reed 1973] pour résoudre le problème hyperbolique du transport de Neutron, et analysée pour la première fois par [Lesaint 1974]. Toutefois, ce n'est qu'à la fin des années 80 et au début des années 90 que la méthode DG connait un développement considérable pour les problèmes hyperboliques [Stephansen 2007, Cockburn 2012a]. Durant la même période, et de manière indépendante, les développements des schémas DG pour l'opérateur elliptique s'inspirent des travaux de Nitsche pour la formulation faible [Nitsche 1972]. Durant la fin du XX^{e} siècle, les nombreux travaux conduisent à une variabilité de méthodes de Galerkine Discontinues, dont un historique détaillé des divers développements est présenté dans [Cockburn 2012a]. Généralement, le choix de la méthodologie dépend de l'expression de la *trace* et du *flux numérique*, qui sont des quantités *uniques* sur le squelette de la partition. Nous nous référons aux travaux d'Arnold [Arnold 2002] qui traitent de façon unifiée plusieurs variations du formalisme de Galerkine Discontinue. Bien qu'il soit impossible de synthétiser l'ensemble des travaux antérieurs sur les méthodes DG, nous rapportons les méthodologies les plus connues : Bassi-Rebay (BR) [Bassi 1997], Baumann-Oden (BO) [Baumann 1999], Locale de Galerkine Discontinue (LDG) [Castillo 2006, Wu 2013] et Pénalisation Intérieure (IP) $\sim$ Interior Penalty en anglais $\sim$ [Douglas 1976, Wheeler 1978, Arnold 1982, Rivière 2008]. Mentionnons également les travaux de [Castillo 2002, Burman 2006, Stephansen 2007, Di Pietro 2011, Mu 2014, Hong 2019] et les références qui s'y trouvent. Le remarquable succès des méthodes DG provient de sa flexibilité, qui combine à la fois les avantages des méthodes VF et EF. Par ailleurs, cette classe s'avère être particulièrement adaptée aux problèmes comportant (i) des fortes hétérogénéités et anisotropies ([Cockburn 1999, Dryja 2003]), (ii) des mécanismes à convection dominante ([Stephansen 2007, Ern 2009, Proft 2009]), ou (iii) un paramètre de diffusion évanescent ([Gastaldi 1989, Di Pietro 2008, Croisille 2005]). Cependant, le formalisme DG présente un inconvénient majeur lié à la taille du système à résoudre, excédant celui de la méthode EF standard (i.e., ratio de 2). Le plus souvent, la forte adaptabilité des méthodes DG aux mailles complexes et non-conformes permet de réduire les coûts additionnels engendrés par les Degrés de Liberté (DDL) du formalisme.

Ces dernières années, les familles de méthodes DG ont connu un essor considérable de la part de la communauté scientifique. La flexibilité de ce formalisme s'étend aisément aux situations physiques et géométriques les plus complexes (e.g., hétérogénéité et anisotropie, mécanisme de convection dominant, discontinuités de la solution, fronts raides, ou maillages non-conformes). Bien que l'approche DG soit très populaire pour la résolution des EDPs (non)-linéaires, cette classe reste généralement plus coûteuse que la plupart des méthodes de discrétisation traditionnelles. En effet, le nombre conséquent de DDLs peut devenir un véritable défi pour les problèmes à grande échelle.

1.4 Motivations

Pour ce travail de recherche, nous abordons principalement la famille des méthodes Pénalisation Intérieure Hybride de Galerkine Discontinue (H-IP) pour améliorer la résolution des lois de conservation. Ce formalisme qui appartient à la classe des méthodes Hybrides de Galerkine Discontinues (HDG), introduit par [Cockburn 2009b], est considéré comme l'homologue *Hybride* de la célèbre méthode de Galerkine Discontinue (voir e.g., [Nguyen 2009, Egger 2010, Wells 2011, Cockburn 2012c, Fabien 2020b]). La méthode HDG introduit principalement la notion d'*hybridation* [Cockburn 2009b, Kirby 2012, Brezzi 2012], qui consiste à rajouter une variable dite *trace* sur la frontière des éléments. En outre, le système linéaire de la méthode HDG bénéficie d'une matrice symétrique plus creuse (faible connectivité inter-éléments) et éparse que le formalisme DG original. Cette classe d'approximation hérite simultanément des avantages des méthodes DG (localement conservatif, stratégie de rafinement-h et -k, flexibilité, approximation polynomiale élevé, fortement parallélisable) et EF (complément de Schur) [Kirby 2012]. Aujourd'hui, il est indéniable que les performances numériques du modèle HDG surpassent celles des méthodes DG standards, en terme de stabilité, de robustesse et d'efficacité [Cockburn 2009b].

Dans ce manuscrit nous avons mené plusieurs études qui ont toutes en commun, la résolution des lois de conservation dans les milieux poreux et fracturés par des méthodes Hybrides de type Pénalisation Intérieure. Ces développements s'orientent autour de trois thématiques principales, que nous proposons de passer en revue dans la suite.

- **Anisotopies et hétérogénéités** - Les discontinuités du tenseur de diffusion sont généralement une source de perturbation pour la résolution du problème. En présence de forts ratios d'anisotropie et d'hétérogénéité, l'approximation de la solution peut devenir irrégulière présentant *a fortiori* des oscillations indésirables (e.g., violation du principe du maximum discret). Afin d'établir un modèle H-IP robuste pour le problème de diffusion hétérogène et anisotrope, nous élaborons une stratégie de pénalisation influencée par la diffusivité du milieu.
- **Diffusion localement évanescente** - La difficulté majeure associée à la résolution de cette EDP concerne la nature mathématique du problème, qui peut devenir non-uniforme d'une région à l'autre de l'espace. Pour contourner cette limitation physique nous proposons d'étendre la méthode H-IP, initialement présentée pour l'opérateur elliptique, dans un contexte de diffusion localement dégénérée. Le formalisme H-IP résultant permettra de traiter simultanément les opérateurs elliptique et hyperbolique, tout en incluant les diférents régimes d'écoulement suivants : (a) diffusion pure, (b) convection pure, (c) régime mixte de diffusion-convection-réaction, et (d) diffusion évanescente.
- **Présence des fractures**- La notion de *fracture*, qui est souvent ignorée, est un aspect fondamental à prendre en compte pour obtenir un comportement plus réaliste des différents phénomènes s'exerçant dans le réservoir. Naturellement, ces formations géologiques sont trop complexes pour être décrites dans leur entièreté [Adler 1999]. Au cours de ce travail de thèse, nous adopterons le modèle de fractures discrètes – Discrete Frac-

ture Matrix (DFM) – pour étudier le mécanisme d'écoulement dans le cas d'un milieu poreux fracturé. L'approche DFM, présentée dans l'annexe A, se distingue par sa représentation géométrique dans laquelle les fractures sont décrites individuellement sous la forme d'interfaces réduites de dimension-$(d-1)$. En employant la méthode H-IP pour la résolution, la localisation des inconnues sur le squelette permet d'inclure naturellement le modèle DFM pour la modélisation des fractures.

1.5 Plan du manuscrit

Ce mémoire de thèse se décline en quatre principaux chapitres, au début desquels se trouve un plan détaillé de leur contenu.

Dans le **chapitre 2**, nous présentons la méthode H-IP pour les problèmes elliptiques du second-ordre fortement hétérogène et anisotrope. Dans notre exposé, nous introduisons un terme de symétrie dans l'expression de la forme bilinéaire du schéma H-IP, qui est directement associé aux trois variations du formalisme, à savoir Incomplète (H-IIP), Non-Symétrique (H-NIP) et Symétrique (H-SIP). En pratique, la méthode H-IP hérite des mêmes propriétés de convergence que les schémas IP classiques. D'une part, l'estimation de l'ordre de convergence suivant la norme énergie est toujours optimale. D'autre part, la convergence des deux variantes non-symétriques suivant la norme-L^2 dépend de la parité du choix de l'approximation polynomiale, i.e., la convergence est optimale uniquement pour des ordres impairs. Afin de restaurer cette propriété de convergence optimale pour des degrés polynomiaux pairs, nous employons une fonction de la pénalisation de la forme $\mathcal{O}(1/h^{1+\delta})$. Puis, nous quantifions l'impact du paramètre δ sur l'étude du caractère *bien-posé* du modèle H-IP en analysant la consistance, la coercivité et la stabilité numérique du schéma. Il en résulte que les estimations de l'erreur suivant la norme-énergie et la norme-L^2 sont dépendantes du coefficient δ. Finalement, nous réalisons une série d'expérimentations numériques pour illustrer ces nouvelles estimations d'erreur, en utilisant des stratégies de rafinement$-h$ et $-k$. Notamment, nous mesurons l'impact du terme de symétrie et des différents paramètres de la fonction de pénalisation sur la convergence optimale de la méthode H-IP. Le chapitre 2 est l'objet d'un article publié dans le journal Computers and Mathematics with Applications.

Dans le **chapitre 3**, nous discutons des familles de méthodes de type Pénalisation Intérieure (IP) pour la résolution des problèmes de diffusion hétérogène et anisotrope. Plus particulièrement, nous nous concentrons sur les schémas Hybrides (H-IP), Encastrés [13] (E-IP) et Pondérés [14] (W-IP), dont la dérivation se structure autour d'un mécanisme de pénalisation unique. Par rapport au formalisme H-IP, la méthode E-IP étend le concept de la trace aux fonctions polynomiales continues. Cette approche Hybride alternative présente une matrice de rigidité réduite et comparable à celle des Éléments Finis. La méthode W-IP se distingue du schéma IP classique en incluant des opérateurs moyennes pondérées, qui sont directement fonction du mécanisme de pénalisation. Bien que les schémas H-IP, E-IP et W-IP sont fondamentalement différentes dans leur représentation (i.e., taille des systèmes matriciels et connectivité inter-élément), nous évaluons les performances numériques de ces trois familles

13. Le terme Encastré fait référence à son équivalent anglo-saxon *Embedded*.
14. Le terme Pondéré fait référence à son équivalent anglo-saxon *Weighted*.

en terme de robustesse, précision et temps CPU. En outre, nous soulignons de suprenantes propriétés d'équivalence totale entre ces différentes classes de modèles IP, établissant de ce fait des passerelles entre les méthodes H-IP, E-IP et W-IP. Pour confirmer le choix de la méthodologie la plus adaptée, nous réalisons des expérimentations numériques pour des tenseurs de diffusion homogène, anisotrope et/ou hétérogène. Le chapitre 3 est l'objet d'un article publié dans le journal Frontiers in Water.

Le **quatrième chapitre** 4 est consacré à l'étude du formalisme H-IP pour la résolution du problème elliptique dégénéré (du second-ordre). Cette EDP comporte des complexités liées à sa nature mathématique, qui peut devenir hyperbolique dans une portion du domaine et elliptique dans le reste. Pour contourner les limitations physiques de ce problème, nous proposons d'enrichir la méthode H-IP afin de traiter simultanément : (*i*) les problèmes purement diffusif ou convectif, (*ii*) les problèmes mixtes combinant à la fois ces deux mécanismes pour une vaste étendue du nombre de Péclet, (*iii*) et la delicate situation de diffusion localement évanescente. Pour ce faire, nous employons une stratégie de pénalisation adaptative, dans le but d'inclure plusieurs approches connues de la littérature, e.g., le schéma upwind, θ-upwind ou encore Scharfetter-Gummel. L'étude du caractère *bien-posé* nous montre que la méthode H-IP dispose des propriétés de consistence, de coercivité et de stabilité. Finalement, pour illustrer les performances et la robustesse du modèle, nous réalisons une série d'expérimentations numériques sur trois situations physiques distinctes : non-dégénérée, entièrement dégénérée et localement dégénérée. Le travail présenté au chapitre 4 a été soumis pour publication dans le journal Applied Mathematics and Computation.

Le **chapitre 5** est dédié à la modélisation de l'écoulement du fluide en présence de fractures. Dans ce contexte, nous orientons la discussion sur le modèle DFM pour la représentation des fissures dans le milieu poreux. L'objectif principal consiste à établir une nouvelle stratégie de résolution basée sur des schémas numériques elligibles à la condensation statique. Plus précisémment, nous combinons la classe des Éléments Finis standards dans les fractures avec la méthode H-IP dans la matrice poreuse, pour résoudre l'écoulement des eaux souterraines dans les milieux poreux fracturés. L'avantage de cette méthodologie couplée réside dans l'assemblage d'un système matriciel, dont les degrés de liberté sont localisés sur le squelette de la partition. Une série d'expérimentations numériques est présentée pour valider la pertinence de ce nouveau formalisme couplé (H-IP/EF). En particulier, nous évaluons la flexibilité et les performances du schéma H-IP/EF dans la manipulation des réseaux de fractures complexes, en incluant des benchmarks réalistes en 2D et 3D, et des comparaisons avec le logiciel commercial COMSOL. Le travail présenté au chapitre 5 a été soumis pour publication dans le journal Advances in Water Ressources.

En **conclusion**, nous proposons un bref récapitulatif des différents travaux effectués dans ce manuscrit. Quelques pistes concernant les perspectives de ce travail de recherche seront mises en avant : par exemple la connexion des méthodes H-IP et DG, ou l'extension de la méthode H-IP pour le transport de masse dans les milieux poreux fracturés. L'annexe A est consacrée à l'établissement des lois de conservation pour le modèle DFM.

Chapitre 2

Error estimates of hybridizable interior penalty methods for diffusion problems

Article published in *Computers and Mathematics with Applications* under the title "Improved error estimates of hybridizable interior penalty methods using a variable penalty for highly anisotropic diffusion problems".
doi:10.1016/j.camwa.2022.05.029

Grégory Etangsale [1], Marwan Fahs [2], Vincent Fontaine[1], Nalitiana Rajaonison[1]

Abstract

In this paper, we derive improved *a priori* error estimates for families of hybridizable interior penalty discontinuous Galerkin (H-IP) methods using a variable penalty for second-order elliptic problems. The strategy is to use a penalization function of the form $\mathcal{O}(1/h^{1+\delta})$, where h denotes the mesh size and δ is a user-dependent parameter. We then quantify its direct impact on the convergence analysis, namely, the (strong) consistency, discrete coercivity and boundedness (with h^{δ}-dependency), and we derive updated error estimates for both discrete energy- and L^2-norms. The originality of the error analysis relies specifically on the use of conforming interpolants of the exact solution. All theoretical results are supported by numerical evidence.

Contents

1. Department of Building and Environmental Sciences, University of La Réunion - South Campus, France
2. Institut Terre et Environnement de Strasbourg, Université de Strasbourg, CNRS, ENGEES, UMR 7063, Strasbourg, France

2.1 Introduction

Hybridizable Discontinuous Galerkin (HDG) methods were first introduced during the last decade by Cockburn *et al.* [Cockburn 2009b] (see e.g., [Egger 2010]), and have since received extensive attention from the research community. They are popular and very efficient numerical approaches for solving a large class of linear and nonlinear partial differential equations. They are still under development and broadly applied in various scientific topics such as groundwater flows [Fabien 2020a, Etangsale 2021a], fluid dynamics [Kirk 2019], solid mechanics [Sevilla 2019], wave propagation [Sánchez 2017], or magneto-hydro-dynamics [Lee 2019], to name but a few. Indeed, they inherit attractive features from both (*i*) discontinuous Galerkin (DG) methods such as local conservation, *hp*-adaptivity and high-order polynomial approximation [Arnold 2002] and (*ii*) standard conforming Galerkin (CG) methods such as the Schur complement strategy [Kirby 2012]. One undeniable additional benefit of the HDG methods is their superconvergence property, obtained through the application of a local postprocessing technique on each element of the mesh [Nguyen 2009]. In the hybrid formalism, additional unknowns are introduced along the mesh skeleton corresponding to discrete trace approximations. Thanks to the specific localization of its additional degrees of freedom (DOFs) and the discontinuous nature of approximation spaces, interior variables can be locally eliminated in favor of its Lagrange multipliers by only static condensation. The problem is then closed, and the algebraic linear system is assembled by imposing *transmission* conditions throughout the mesh skeleton. This strategy is now well-established and -documented in the literature, and we refer the interested reader to the following works for a detailed description [Kirby 2012, Lehrenfeld 2010] (see also, [Etangsale 2021a, Remark 4.1] for the description of the static condensation technique). The resulting matrix system is significantly smaller and sparser than those associated with CG or DG methods for any given mesh and polynomial degree [Kirby 2012]. Several HDG formulations have been derived in the literature and can be classified into two main categories. The first is based on a primal form of the continuous problem, such as the class of interior penalty (IP) methods [Fabien 2020b], whereas the second relies on a dual (often called mixed) form, such as local discontinuous Galerkin (LDG) methods [Cockburn 2009b, Nguyen 2009, Dijoux 2019b]. In the latter formulation, the flux variable is introduced as an additional unknown of the problem.

Our focus is on families of hybridizable Interior Penalty Discontinuous Galerkin (H-IP) methods [Wells 2011]. They are hybridized counterparts of the well-known interior penalty DG (IPDG) methods [Arnold 1982, Rivière 2008, Di Pietro 2011] and have been analyzed

until quite recently by several authors [Kirk 2019, Fabien 2020b]. Specifically, in our exposition, we considered the Incomplete, Non-Symmetric and Symmetric schemes denoted by H-IIP, H-NIP and H-SIP, respectively. The main difference between these schemes concerns the role of the *symmetrization* term in the discrete bilinear form [Rivière 2008]. Fabien *et al.* recently analyzed these schemes using a stabilization function of the form $\mathcal{O}(1/h)$ for solving second-order elliptic problems [Fabien 2020b]. The authors conclude that H-IP methods inherit similar convergence properties to their IPDG equivalents. Notably, they theoretically establish (i) optimal energy error estimates, and because of the lack of symmetry of the associated discrete operator, (ii) only suboptimal L^2-norm error estimates for H-IIP and H-NIP schemes. In addition, they numerically conclude that the L^2-orders of convergence of both Non-Symmetric variants are suboptimal for only even polynomial degrees and are optimal otherwise. Similar conclusions have also been suggested by Oikawa for second-order elliptic problems [Oikawa 2017].

To restore optimal L^2-error estimates for the non-symmetric IPDG method, Rivière *et al.* suggest using a sort of superpenalty on the jumps [Rivière 1999, Guzmán 2009]. In the present paper, we explore a similar idea in the general context of H-IP methods by using a variable penalty function of the form $\tau := \mathcal{O}(1/h^{1+\delta})$, where $\delta \in \mathbb{R}$. Here, we analyze the direct impact of the parameter δ on *a priori* error estimates in different norms. First, we propose a convergence analysis by investigating three key properties : (strong) consistency, discrete coercivity and boundedness. One remarkable feature of this strategy is the h^{δ}-dependency of the coercivity condition and the continuity (or boundedness) constant C_{bnd}, which consequently impacts the error estimates. Improved error estimates are then derived in the spirit of the second Strang lemma [Di Pietro 2011], and we first prove that the order of convergence in the natural energy-norm is linear, δ-dependent, and optimal when $\delta \geq 0$ for any scheme. Then, by using a duality argument, i.e., the so-called Aubin–Nitsche technique, we also prove that the optimal convergence is theoretically reached as soon as $\delta \geq 0$ for the H-SIP scheme only, and when $\delta \geq 2$ for both non-symmetric variants, i.e., H-NIP and H-IIP schemes. Let us underline that we recover theoretical error estimates proposed in the literature for both the energy- and L^2-norms if $\delta = 0$.

The rest of the material is organized as follows : Section 2.2 describes the model problem, mesh notation and assumptions, and recalls some definitions and useful (trace) inequalities, while Section 2.3 derives the discrete H-IP formulation and discusses its stability properties. In Section 2.4, optimal error estimates are provided for both the energy- and L^2-norms by using a standard duality argument. Section 2.5 concerns the numerical experiments that validate our theoretical results. We briefly end with some remarks and perspectives.

2.2 Some preliminaries

2.2.1 The model problem

Let Ω be a bounded (polyhedron) domain in $\mathbb{R}^d$ with Lipschitz boundary $\partial\Omega$ in spatial dimension $d \geq 2$. For clarity, we consider the anisotropic diffusion problem with homogeneous

Dirichlet boundary conditions :

$$-\boldsymbol{\nabla}\cdot(\boldsymbol{\alpha}\boldsymbol{\nabla}u)=f \quad \text{in } \Omega \quad \text{and} \quad u=0 \quad \text{on } \partial\Omega, \tag{2.1}$$

where $\boldsymbol{\alpha}\in[L^\infty(\Omega)]^{d\times d}$ is a bounded, symmetric, uniformly positive-definite matrix-valued function and $f\in L^2(\Omega)$ is a forcing term. Thus, the weak formulation of problem (2.1) is to find $u\in H_0^1(\Omega)$ such that

$$\int_\Omega \boldsymbol{\alpha}\boldsymbol{\nabla}u\cdot\boldsymbol{\nabla}v\mathrm{d}x=\int_\Omega fv\mathrm{d}x \quad \forall v\in H_0^1(\Omega). \tag{2.2}$$

It is well-known that under elliptic regularity assumptions, the variational problem (2.2) is well-posed.

2.2.2 Mesh notation and assumptions

Let h be a positive parameter ; we assume without loss of generality that $h\leq 1$. We denote by $\{\mathcal{T}_h\}_{h>0}$ a family of affine triangulations of the domain Ω, where h stands for the largest diameter : $h_E:=\mathrm{diam}(E)$. We also assume that $\mathcal{T}_h$ is *quasi-uniform*, meaning that for all $E\in\mathcal{T}_h$, there exists $0<\rho_0\leq 1$ independent of h such that $\rho_0 h\leq h_E\leq h$. Following our notation, the generic term *interface* indicates a $(d-1)$–dimensional geometric object, i.e., an edge, if $d=2$ and a face if $d=3$. Thus, we denote by $\mathcal{F}_h^{\mathsf{i}}$ and $\mathcal{F}_h^{\mathsf{b}}$ the set of interior and boundary interfaces, respectively. The set of all interfaces is called the mesh skeleton and is denoted by $\mathcal{F}_h:=\mathcal{F}_h^{\mathsf{i}}\cup\mathcal{F}_h^{\mathsf{b}}$. We denote by $\partial\mathcal{T}_h:=\{\cup\partial E,\forall E\in\mathcal{T}_h\}$, the collection of interfaces of all mesh elements. Let X be a mesh element or an interface ; we then denote by $|X|$ a positive d- or $(d-1)$-dimensional Lebesgue measure of X, 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 ; we define $\eta_E:=\mathrm{card}(\mathcal{F}_E)$ and $\eta_0:=\max\limits_{\forall E\in\mathcal{T}_h}(\eta_E)$.

2.2.3 Broken polynomial spaces

For any polyhedral domain $\mathcal{D}\subset\mathbb{R}^d$ with $\partial\mathcal{D}\subset\mathbb{R}^{d-1}$, we denote by $(\cdot,\cdot)_{0,\mathcal{D}}$ (resp., $\langle\cdot,\cdot\rangle_{0,\partial\mathcal{D}}$) the L^2-inner product in $L^2(\mathcal{D})$ (resp., $L^2(\partial\mathcal{D})$) equipped with its natural norm $\|\cdot\|_{0,\mathcal{D}}$ (resp., $\|\cdot\|_{0,\partial\mathcal{D}}$). Let us now introduce some compact notation associated with the discrete L^2-inner scalar product :

$$(\cdot,\cdot)_{0,\mathcal{T}_h}:=\sum_{E\in\mathcal{T}_h}(\cdot,\cdot)_{0,E},\qquad \langle\cdot,\cdot\rangle_{0,\partial\mathcal{T}_h}:=\sum_{E\in\mathcal{T}_h}\langle\cdot,\cdot\rangle_{0,\partial E},$$

and we denote by $\|\cdot\|_{0,\mathcal{T}_h}$ and $\|\cdot\|_{0,\partial\mathcal{T}_h}$ the corresponding norms. Similarly, we denote by $H^s(\mathcal{D})$ the usual Hilbert space of index s on $\mathcal{D}$ equipped with its natural norm $\|\cdot\|_{s,\mathcal{D}}$ and seminorm $|\cdot|_{s,D}$, respectively. If $s=0$, we then set $H^0(\mathcal{D})=L^2(\mathcal{D})$. We denote by $H^s(\mathcal{T}_h)$ the usual broken Sobolev space and by $\boldsymbol{\nabla}_h$ the broken gradient operator acting on $H^s(\mathcal{T}_h)$ with $s\geq 1$. We then assume an extended regularity requirement of the exact solution u of the weak problem (2.2), i.e., $u\in H_0^s(\Omega)\cap H^2(\mathcal{T}_h)$ with $s>3/2$. We also introduce the additional unknown $\hat{u}\in L^2(\mathcal{F}_h)$ corresponding to the restriction of u on the mesh skeleton ; i.e., $\hat{u}:=u|_{\mathcal{F}_h}$. Let us now introduce the composite variable $\boldsymbol{u}:=(u,\hat{u})$, which belongs to the

continuous approximation space $\boldsymbol{\mathcal{V}} := H_0^s(\Omega) \cap H^2(\mathcal{T}_h) \times L^2(\mathcal{F}_h)$; i.e., $\boldsymbol{u} \in \boldsymbol{\mathcal{V}}$. As usual in HDG methods, we consider broken Sobolev spaces :

$$\mathcal{P}_k(\mathcal{T}_h) := \{v_h \in L^2(\mathcal{T}_h) \ : \ v_h|_E \in \mathcal{P}_k(E), \ \forall E \in \mathcal{T}_h\}, \tag{2.3}$$

and similarly for $\mathcal{P}_k(\mathcal{F}_h)$. Here, $\mathcal{P}_k(X)$ 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 (2.2). First, the discrete variable $u_h \in \mathcal{V}_h$ which is defined within each mesh element, and its trace $\hat{u}_h \in \hat{\mathcal{V}}_h$, defined on the mesh skeleton with respect to the imposed homogeneous Dirichlet boundary conditions. To this aim, we set $\mathcal{V}_h := \mathcal{P}_k(\mathcal{T}_h)$ and $\hat{\mathcal{V}}_h := \mathcal{P}_k^0(\mathcal{F}_h)$, where

$$\mathcal{P}_k^0(\mathcal{F}_h) := \{\hat{v}_h \in \mathcal{P}_k(\mathcal{F}_h) \ : \ \hat{v}_h|_F = 0, \ \forall F \in \mathcal{F}_h^\flat\}. \tag{2.4}$$

Throughout the manuscript, we use the following compact notations. Thus, let $\boldsymbol{\mathcal{V}}_h := \mathcal{V}_h \times \hat{\mathcal{V}}_h$ be the composite approximation space and a generic element of $\boldsymbol{\mathcal{V}}_h$ is denoted by $\boldsymbol{v}_h := (v_h, \hat{v}_h)$. For all $E \in \mathcal{T}_h$ and $F \in \mathcal{F}_E$, we define the jump of $\boldsymbol{v}_h \in \boldsymbol{\mathcal{V}}_h$ across F as $[\![\boldsymbol{v}_h]\!]|_{E,F} := (v_h|_F - \hat{v}_h|_F)\boldsymbol{n}_{E,F}$, where $\boldsymbol{n}_{E,F}$ is the unit normal vector to F pointing out of E. When confusion cannot arise, we omit the subscripts E and F from the definition, and we simply write $[\![\boldsymbol{v}_h]\!] := (v_h - \hat{v}_h)\boldsymbol{n}$. Finally, we introduce the space $\boldsymbol{\mathcal{V}}(h) := \boldsymbol{\mathcal{V}} + \boldsymbol{\mathcal{V}}_h$ to analyze the boundedness of the discrete bilinear form.

2.2.4 Useful inequalities

We recall here some useful inequalities that will be used extensively (see, e.g., [Ciarlet 1991, Di Pietro 2011, Rivière 2008]). For clarity, C denotes a generic constant that is independent of h, h_E and $\boldsymbol{\alpha}$ in the rest of the manuscript. Owing to the shape regularity of $\mathcal{T}_h$, we now introduce multiplicative trace inequalities. Let $E \in \mathcal{T}_h$ and $F \in \mathcal{F}_E$. For all $v \in H^2(E)$, there exists a constant $C_{\mathrm{M}} > 0$ independent of h_E and v such that

$$\|v\|_{0,F}^2 \leq C_{\mathrm{M}}(\|v\|_{0,E}|v|_{1,E} + h_E^{-1}\|v\|_{0,E}^2), \tag{2.5a}$$

$$\|\boldsymbol{\nabla}_h v\|_{0,F}^2 \leq C_{\mathrm{M}}(|v|_{1,E}|v|_{2,E} + h_E^{-1}|v|_{1,E}^2). \tag{2.5b}$$

Let us now remind the discrete and inverse trace inequalities, respectively. For all $v_h \in \mathcal{V}_h$, then the following holds

$$\|v_h\|_{0,F} \leq C_{\mathrm{tr}} h_E^{-1/2}\|v_h\|_{0,E}, \tag{2.6a}$$

$$\|\boldsymbol{\nabla}_h v_h\|_{0,E} \leq C_{\mathrm{inv}} h_E^{-1}\|v_h\|_{0,E}, \tag{2.6b}$$

where C_{tr} and C_{inv} are positive constants independent of h_E.

Remark 2.1. Following [Rivière 2008, Section 2.1.3, p.24], one can obtain an exact expression of the constant C_{tr} used in the discrete trace inequality (2.6a) for a d-simplex mesh element :

$$C_{\mathrm{tr}} := \sqrt{\frac{(k+1)(k+d)}{d}}, \tag{2.7}$$

where k denotes the polynomial degree of $\mathcal{V}_h$ and d denotes the spatial dimension. This expression is particularly important in our analysis since it will be used later in the definition of the penalty parameter.

We are now in a position to introduce the energy-norm used in the stability analysis and error estimations [Wells 2011, Lehrenfeld 2010]. For any given composite function $\boldsymbol{v}_h \in \boldsymbol{\mathcal{V}}_h$, we consider the jump seminorm :

$$|\boldsymbol{v}_h|_\gamma^2 := \sum_{E\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_E} \|\gamma_{E,F}^{1/2}[\![\boldsymbol{v}_h]\!]\|_{0,F}^2, \tag{2.8}$$

where $\gamma_{E,F} \geq 0$ is an arbitrary positive constant associated with $F \in \mathcal{F}_E$. The natural energy-norm equipping the discrete approximation space $\boldsymbol{\mathcal{V}}_h$ is given by

$$\|\boldsymbol{v}_h\|_*^2 := \|\boldsymbol{\alpha}^{1/2}\boldsymbol{\nabla}_h v_h\|_{0,\mathcal{T}_h}^2 + |\boldsymbol{v}_h|_\gamma^2, \tag{2.9}$$

which clearly depends on $\boldsymbol{\alpha}$.

2.3 Hybridizable interior penalty methods

The discrete H-IP problem is to find $\boldsymbol{u}_h \in \boldsymbol{\mathcal{V}}_h$ such that

$$\mathcal{B}_h^{(\varepsilon)}(\boldsymbol{u}_h, \boldsymbol{v}_h) = \mathcal{L}_h(\boldsymbol{v}_h), \quad \forall \boldsymbol{v}_h \in \boldsymbol{\mathcal{V}}_h, \tag{2.10}$$

where $\mathcal{L}_h(\boldsymbol{v}_h) := (f, v_h)_{0,\mathcal{T}_h}$. Here, the bilinear form $\mathcal{B}_h^{(\varepsilon)} : \boldsymbol{\mathcal{V}}_h \times \boldsymbol{\mathcal{V}}_h \to \mathbb{R}$ is given by

$$\begin{aligned}\mathcal{B}_h^{(\varepsilon)}(\boldsymbol{u}_h, \boldsymbol{v}_h) :=& (\boldsymbol{\alpha}\boldsymbol{\nabla}_h u_h, \boldsymbol{\nabla}_h v_h)_{0,\mathcal{T}_h} + \langle \tau[\![\boldsymbol{u}_h]\!], [\![\boldsymbol{v}_h]\!]\rangle_{0,\partial\mathcal{T}_h} \\ &- \langle \boldsymbol{\alpha}\boldsymbol{\nabla}_h u_h, [\![\boldsymbol{v}_h]\!]\rangle_{0,\partial\mathcal{T}_h} - \varepsilon\langle \boldsymbol{\alpha}\boldsymbol{\nabla}_h v_h, [\![\boldsymbol{u}_h]\!]\rangle_{0,\partial\mathcal{T}_h},\end{aligned} \tag{2.11}$$

where $\varepsilon \in \{0, \pm 1\}$. The second, third and fourth terms on the right-hand side of (2.11) are called the jump-penalty, consistency, and symmetry terms, respectively. The discrete bilinear operator $\mathcal{B}_h^{(\varepsilon)}$ is symmetric if $\varepsilon = 1$ and is non-symmetric otherwise. We obtain the Symmetric scheme (H-SIP) if $\varepsilon = 1$, the Incomplete scheme (H-IIP) if $\varepsilon = 0$ and the Non-Symmetric scheme (H-NIP) if $\varepsilon = -1$. For all $E \in \mathcal{T}_h$ and $F \in \mathcal{F}_E$, the penalty term is chosen as follows :

$$\tau_{E,F}\frac{\tau_0 C_{\mathrm{tr}}^2 \alpha_{E,F}}{h_E^{1+\delta}} \quad \text{with} \quad \delta \in \mathbb{R}, \tag{2.12}$$

where τ_0 is a user-dependent parameter, C_{tr} is given by (2.7) and results from the discrete trace inequality (2.6a), and $\alpha_{E,F} := \boldsymbol{n}_{E,F}\boldsymbol{\alpha}_E\boldsymbol{n}_{E,F}$ denotes the normal diffusivity.

Remark 2.2. For simplicity, we assume that $\boldsymbol{\alpha}$ is approximated by piecewise constants on the mesh element $\mathcal{T}_h$; i.e., $\boldsymbol{\alpha}_{|E} \in \mathbb{R}^{d\times d}$ for all $E \in \mathcal{T}_h$.

Lemma 2.3.1 (Consistency)**.** *Let $\boldsymbol{u} \in \boldsymbol{\mathcal{V}}$ be the compact notation of the exact solution of the problem (2.2). For all $\boldsymbol{v}_h \in \boldsymbol{\mathcal{V}}_h$, then the following holds :*

$$\mathcal{B}_h^{(\cdot)}(\boldsymbol{u}, \boldsymbol{v}_h) = \mathcal{L}_h(\boldsymbol{v}_h). \tag{2.13}$$

Proof. The regularity of $\boldsymbol{u} := (u, \hat{u})$ 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 $F \in \mathcal{F}_E$ then $[\![\boldsymbol{u}]\!] := 0$, since u is a single-valued field on the mesh

skeleton. Thus, by setting $\boldsymbol{v}_h := (v_h, 0)$, and integrating by parts on each element of the mesh, the bilinear form $\mathcal{B}_h^{(\varepsilon)}$ yields

$$\mathcal{B}_h^{(\varepsilon)}(\boldsymbol{u}, (v_h, 0)) := \sum_{E \in \mathcal{T}_h} (\boldsymbol{\nabla}_h \cdot (-\boldsymbol{\alpha} \boldsymbol{\nabla}_h u), v_h)_{0,E} = \sum_{E \in \mathcal{T}_h} (f, v_h)_{0,E}. \tag{2.14}$$

Considering now that $\boldsymbol{v}_h := (0, \hat{v}_h) \in \boldsymbol{\mathcal{V}}_h$ and $\hat{v}_h$ vanishes on the boundary skeleton $\mathcal{F}_h^{\mathfrak{b}}$, we then obtain

$$\mathcal{B}_h^{(\varepsilon)}(\boldsymbol{u}, (0, \hat{v}_h)) := \langle (\boldsymbol{\alpha} \boldsymbol{\nabla}_h u) \cdot \boldsymbol{n}, \hat{v}_h \rangle_{0, \partial \mathcal{T}_h} = 0, \tag{2.15}$$

which corresponds to the transmission conditions. The proof is then completed by summing (2.14) and (2.15). ■

A straightforward consequence of Lemma 2.3.1 is the Galerkin orthogonality.

Proposition 2.1 (Galerkin orthogonality)**.** *Let $\boldsymbol{u} \in \boldsymbol{\mathcal{V}}$ be the compact notation of the exact solution of the problem (2.2), and $\boldsymbol{u}_h \in \boldsymbol{\mathcal{V}}_h$, the solution of the discrete problem (2.10). Then,*

$$\mathcal{B}_h^{(\varepsilon)}(\boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{v}_h) = 0 \quad \forall \boldsymbol{v}_h \in \boldsymbol{\mathcal{V}}_h. \tag{2.16}$$

Proof. Subtracting (2.13) and (2.10) yields the assertion. ■

2.3.1 Coercivity and well-posedness

The next step is to prove discrete coercivity of $\mathcal{B}_h^{(\cdot)}$ to ensure the well-posedness of (2.10). To this end, we first need to establish an upper bound of the consistency term using the jump seminorm $|\cdot|_\tau$.

Lemma 2.3.2 (Bound on consistency term)**.** *There exists a constant $C_\delta > 0$ which is h^δ-dependent such that*

$$|\langle \boldsymbol{\alpha} \boldsymbol{\nabla}_h w_h, [\![\boldsymbol{v}_h]\!] \rangle_{0, \partial \mathcal{T}_h}| \leq C_\delta^{1/2} \|\boldsymbol{\alpha}^{1/2} \boldsymbol{\nabla}_h w_h\|_{0, \mathcal{T}_h} |\boldsymbol{v}_h|_\tau, \tag{2.17}$$

for all $(\boldsymbol{w}_h, \boldsymbol{v}_h) \in \boldsymbol{\mathcal{V}}_h \times \boldsymbol{\mathcal{V}}_h$. Here $C_\delta := C_0 h^\delta$ and $C_0 := C\eta_0/\tau_0$ is a constant dependent of the element shape only.

Proof. The decomposition of the consistency term yields

$$\langle \boldsymbol{\alpha} \boldsymbol{\nabla}_h w_h, [\![\boldsymbol{v}_h]\!] \rangle_{0, \partial \mathcal{T}_h} = \sum_{E \in \mathcal{T}_h} \langle \boldsymbol{\alpha} \boldsymbol{\nabla}_h w_h, [\![\boldsymbol{v}_h]\!] \rangle_{0, \partial E}.$$

Applying the Cauchy–Schwarz inequality, using the definition of τ given in (2.12) and finally applying the discrete trace inequality (2.6a), we infer that

$$\begin{aligned}
|\langle \boldsymbol{\alpha} \boldsymbol{\nabla}_h w_h, [\![\boldsymbol{v}_h]\!] \rangle_{0, \partial E}| &\leq \|\boldsymbol{\alpha}^{1/2} \boldsymbol{\nabla}_h w_h\|_{0, \partial E} \|\boldsymbol{\alpha}^{1/2} [\![\boldsymbol{v}_h]\!]\|_{0, \partial E}, \\
&\leq \left[\frac{h_E^{1+\delta}}{\tau_0 C_{\mathrm{tr}}^2} \right]^{1/2} \|\boldsymbol{\alpha}^{1/2} \boldsymbol{\nabla}_h w_h\|_{0, \partial E} |\boldsymbol{v}_h|_{\tau, \partial E}, \\
&\leq \left[\frac{\eta_E h_E^\delta}{\tau_0} \right]^{1/2} \|\boldsymbol{\alpha}^{1/2} \boldsymbol{\nabla}_h w_h\|_{0, E} |\boldsymbol{v}_h|_{\tau, \partial E}.
\end{aligned}$$

Considering now the quasi-uniformity requirement of the partition $\mathcal{T}_h$ – i.e., for all $E \in \mathcal{T}_h$ and $\delta \in \mathbb{R}$ there exists C such that $h_E^\delta \leq C h^\delta$, we thus obtain

$$|\langle \boldsymbol{\alpha} \boldsymbol{\nabla}_h w_h, [\![\boldsymbol{v}_h]\!] \rangle_{0,\partial E}| \leq \left[\frac{C \eta_0 h^\delta}{\tau_0} \right]^{1/2} \|\boldsymbol{\alpha}^{1/2} \boldsymbol{\nabla}_h w_h\|_{0,E} |\boldsymbol{v}_h|_{\tau,\partial E}.$$

The proof is thus completed by summing over all mesh elements, and applying the Cauchy–Schwarz inequality. ■

Lemma 2.3.3 (Coercivity)**.** *Let us first introduce the (minimal) threshold value of the following form* $\underline{\tau}_{\varepsilon,\delta} := C_\varepsilon h^\delta$ *where* $C_\varepsilon \geq 0$ *which is null if* $\varepsilon = -1$*. If the penalty parameter* τ_0 *in* (2.12) *is chosen large enough, i.e.,* $\tau_0 > \underline{\tau}_{\varepsilon,\delta}$*, then the discrete bilinear form* $\mathcal{B}_h^{(\cdot)}$ *is* $\boldsymbol{\mathcal{V}}_h$*-coercive with respect to the energy-norm* $\|\cdot\|_*$*; i.e., for all* $\boldsymbol{v}_h \in \boldsymbol{\mathcal{V}}_h$*, then the following holds*

$$\mathcal{B}_h^{(\cdot)}(\boldsymbol{v}_h, \boldsymbol{v}_h) \geq \frac{1}{2} \|\boldsymbol{v}_h\|_*^2. \tag{2.18}$$

Proof. Setting $\boldsymbol{u}_h = \boldsymbol{v}_h$ in (2.11), we thus obtain

$$\mathcal{B}_h^{(\varepsilon)}(\boldsymbol{v}_h, \boldsymbol{v}_h) = \|\boldsymbol{\alpha}^{1/2} \boldsymbol{\nabla}_h v_h\|_{0,\mathcal{T}_h}^2 + |\boldsymbol{v}_h|_\tau^2 - (1+\varepsilon) \langle \boldsymbol{\alpha} \boldsymbol{\nabla}_h v_h, [\![\boldsymbol{v}_h]\!] \rangle_{0,\partial\mathcal{T}_h},$$

proving immediately the coercivity of H-NIP scheme ($\varepsilon = -1$) for any given value $\tau_0 > 0$. Else, owing to Lemmata 2.3.2 and using Young's inequality, for any $0 < \zeta < 1$, there exists a constant $C_\zeta > 0$ such that

$$\begin{aligned} \mathcal{B}_h^{(\varepsilon)}(\boldsymbol{v}_h, \boldsymbol{v}_h) &\geq \left[1 - \frac{C_\delta}{\zeta} \right] \|\boldsymbol{\alpha}^{1/2} \boldsymbol{\nabla}_h v_h\|_{0,\mathcal{T}_h}^2 + (1-\zeta) |\boldsymbol{v}_h|_\tau^2 \\ &\geq C_\zeta \|\boldsymbol{v}_h\|_*^2, \end{aligned}$$

where $C_\zeta := \min(1 - C_\delta/\zeta, 1-\zeta)$. We now select τ_0 in the definition of C_δ such that $C_\zeta := 1-\zeta$; i.e., $C_\delta < \zeta^2$ or equivalently by assuming that $\tau_0 > \zeta^{-2} C \eta_0 h^\delta$. The proof is thus completed by setting (arbitrary) $\zeta = 1/2$. ■

Remark 2.3. Note here the h^δ-dependency of the coercivity condition of both H-SIP and H-IIP schemes. A straightforward consequence of the consistency and coercivity requirements via the Lax–Milgram Theorem is the well-posedness of the weak problem (2.10); i.e., the existence and uniqueness of $\boldsymbol{u}_h \in \boldsymbol{\mathcal{V}}_h$ are ensured.

2.3.2 Boundedness

We now assume that the bilinear form $\mathcal{B}_h^{(\varepsilon)}$ can be extended to $\boldsymbol{\mathcal{V}}(h) \times \boldsymbol{\mathcal{V}}(h)$, and we assert the boundedness of the product space. To this end, we introduce the enriched energy-norm on $\boldsymbol{\mathcal{V}}(h)$ denoted by $|||\cdot|||$ (which is also a natural norm on $\boldsymbol{\mathcal{V}}_h$) to bound the (normal) derivative terms [Lehrenfeld 2010]. For all $\boldsymbol{v} \in \boldsymbol{\mathcal{V}}(h)$, then we set

$$|||\boldsymbol{v}|||^2 := \|\boldsymbol{v}\|_*^2 + \sum_{E \in \mathcal{T}_h} h_E \|\boldsymbol{\alpha}^{1/2} \boldsymbol{\nabla}_h v\|_{0,\partial E}^2. \tag{2.19}$$

Lemma 2.3.4 (Equivalency of $\|\cdot\|_*$- and $|||\cdot|||$-norms)**.** *The energy norms $\|\cdot\|_*$ and $|||\cdot|||$ given by (2.9) and (2.19), respectively, are uniformly equivalent on $\boldsymbol{\mathcal{V}}_h$; i.e., there exists a constant $\rho > 0$ such that*

$$\forall \boldsymbol{v}_h \in \boldsymbol{\mathcal{V}}_h, \quad \rho^{-1}|||\boldsymbol{v}_h||| \leq \|\boldsymbol{v}_h\|_* \leq |||\boldsymbol{v}_h|||, \tag{2.20}$$

where $\rho := (1 + \eta_0 C_{\mathrm{tr}}^2)^{1/2}$ depends only on the element shape.

Proof. Following the definition (2.19), we first notice that $\|\boldsymbol{v}_h\|_* \leq |||\boldsymbol{v}_h|||$. As $\boldsymbol{v}_h$ is piecewise polynomial, we now can easily bound the difference of both norms by using the discrete trace inequality (2.6a),

$$|||\boldsymbol{v}_h|||^2 - \|\boldsymbol{v}_h\|_*^2 \leq \eta_0 C_{\mathrm{tr}}^2 \|\boldsymbol{\alpha}^{1/2}\boldsymbol{\nabla}_h v_h\|_{0,\mathcal{T}_h}^2 \leq \eta_0 C_{\mathrm{tr}}^2 \|\boldsymbol{v}_h\|_*^2,$$

which yields the assertion. ■

Lemma 2.3.5 (Boundedness with $\boldsymbol{h^\delta}$-dependency)**.** *There exists a constant $C_{\mathrm{bnd}} > 0$ which is h^δ-dependent such that*

$$\forall (\boldsymbol{w}, \boldsymbol{v}) \in \boldsymbol{\mathcal{V}}(h) \times \boldsymbol{\mathcal{V}}(h), \quad \mathcal{B}_h^{(\varepsilon)}(\boldsymbol{w}, \boldsymbol{v}) \leq C_{\mathrm{bnd}} |||\boldsymbol{w}||| \, |||\boldsymbol{v}|||, \tag{2.21}$$

where $C_{\mathrm{bnd}} := \max(2, C_1 h^\delta)$ and $C_1 := (\tau_0 C_{\mathrm{tr}}^2)^{-1}$ is a constant independent of h.

Proof. The bilinear form (2.11) can be decomposed as follows :

$$\left|\mathcal{B}_h^{(\varepsilon)}(\boldsymbol{w}, \boldsymbol{v})\right| \leq |\mathcal{T}_1 + \mathcal{T}_2| + |\mathcal{T}_3| + |\mathcal{T}_4|\,, \tag{2.22}$$

where each terms are given below by

$$\begin{aligned}
\mathcal{T}_1 &:= (\boldsymbol{\alpha}^{1/2}\boldsymbol{\nabla}_h w, \boldsymbol{\alpha}^{1/2}\boldsymbol{\nabla}_h v)_{0,\mathcal{T}_h}, \\
\mathcal{T}_2 &:= \langle \tau^{1/2}[\![\boldsymbol{w}]\!], \tau^{1/2}[\![\boldsymbol{v}]\!]\rangle_{0,\partial\mathcal{T}_h}, \\
\mathcal{T}_3 &:= \langle \boldsymbol{\alpha}^{1/2}\boldsymbol{\nabla}_h w, \boldsymbol{\alpha}^{1/2}[\![\boldsymbol{v}]\!]\rangle_{0,\partial\mathcal{T}_h}.
\end{aligned}$$

The last term $\mathcal{T}_4$ is deduced from $\mathcal{T}_3$ by permuting the role of $\boldsymbol{w}$ and $\boldsymbol{v}$, respectively. Thus, applying the Cauchy–Schwarz inequality, the first two terms can be bounded as follows :

$$|\mathcal{T}_1 + \mathcal{T}_2| \leq \; [\|\boldsymbol{\alpha}^{1/2}\boldsymbol{\nabla}_h w\|_{0,\mathcal{T}_h}^2 + |\boldsymbol{w}|_\tau^2]^{1/2}[\|\boldsymbol{\alpha}^{1/2}\boldsymbol{\nabla}_h v\|_{0,\mathcal{T}_h}^2 + |\boldsymbol{v}|_\tau^2]^{1/2} \leq \; \|\boldsymbol{w}\|_*\|\boldsymbol{v}\|_*.$$

Proceeding as in the proof of Lemmata 2.3.2, the third term can also be bounded as follows :

$$|\mathcal{T}_3| \leq \left[C_1 h^\delta \sum_{E\in\mathcal{T}_h} h_E \|\boldsymbol{\alpha}^{1/2}\boldsymbol{\nabla}_h w\|_{0,\partial E}^2\right]^{1/2} \|\boldsymbol{v}\|_*, \tag{2.24}$$

where $C_1 := (\tau_0 C_{\mathrm{tr}}^2)^{-1}$, and similarly for the fourth term $|\mathcal{T}_4|$. Collecting these estimates, and finally using the Cauchy–Schwarz inequality, we thus obtain

$$\begin{aligned}
\left|\mathcal{B}_h^{(\varepsilon)}(\boldsymbol{w}, \boldsymbol{v})\right| &\leq \left[2\|\boldsymbol{w}\|_*^2 + C_1 h^\delta \sum_{E\in\mathcal{T}_h} h_E \|\boldsymbol{\alpha}^{1/2}\boldsymbol{\nabla}_h w\|_{0,\partial E}^2\right]^{1/2} \times \\
&\qquad \left[2\|\boldsymbol{v}\|_*^2 + C_1 h^\delta \sum_{E\in\mathcal{T}_h} h_E \|\boldsymbol{\alpha}^{1/2}\boldsymbol{\nabla}_h v\|_{0,\partial E}^2\right]^{1/2}, \\
&\leq \max(2, C_1 h^\delta) |||\boldsymbol{w}||| \, |||\boldsymbol{v}|||,
\end{aligned}$$

which yields the assertion. ■

Remark 2.4. Let us emphasize here that $C_{\mathrm{bnd}} \leq Ch^{r_\delta}$, where $r_\delta = \min(0,\delta)$ and $C := 2\max(2,C_1)$ is a positive constant independent of h.

2.4 A priori error analysis

We now derive *a priori* error estimates in both the discrete energy- and $\|\cdot\|_{0,\mathcal{T}_h}$-norms. The first ingredient of our error analysis is a bound on the quantity $\boldsymbol{u} - \pi_h\boldsymbol{u}$ in the energy-norms where $\pi_h\boldsymbol{u}$ denotes a suitable *continuous* interpolant of the compact solution $\boldsymbol{u}$ of the problem (2.2). Different authors have previously employed this trick for the error analysis of DG methods (see, e.g., [Wells 2011] for a detailed description) since it offers several advantages. To this end, we recall standard interpolation estimates that will be used extensively in the rest of the document (see, e.g., [Di Pietro 2011, Ciarlet 1991]). Let us consider $\phi \in H^s(\Omega)$ with $s \geq 2$, and we denote by $\pi_h^{\mathrm{i}}\phi$ its *continuous* interpolant of degree k. Thus, the following estimates hold,

$$|\phi - \pi_h^{\mathrm{i}}\phi|_{q,\mathcal{T}_h} \leq Ch^{\mu-q}|\phi|_{\mu,\mathcal{T}_h}, \quad \forall q \in \{0,\dots,s-1\}, \tag{2.25a}$$

$$\left[\sum_{E\in\mathcal{T}_h} h_E^{\alpha}\|\boldsymbol{\nabla}_h(\phi - \pi_h^{\mathrm{i}}\phi)\|_{0,\partial E}^2\right]^{1/2} \leq Ch^{\mu+\frac{\alpha-3}{2}}|\phi|_{\mu,\mathcal{T}_h}, \tag{2.25b}$$

where $\mu := \min(k+1,s)$ and k denotes the polynomial degree of the approximation space $\mathcal{V}_h$.

Lemma 2.4.1 (Optimal error estimates)**.** *Let $\boldsymbol{u} \in \boldsymbol{\mathcal{V}}$ be the compact notation of the exact solution of the problem (2.2). We denote by $\pi_h\boldsymbol{u} := (\pi_h^{\mathrm{i}}u, \pi_h^{\mathfrak{b}}\hat{u})$ its continuous interpolant, where $\pi_h^{\mathrm{i}}u \in \mathcal{V}_h \cap C^0(\bar{\Omega})$ and $\pi_h^{\mathfrak{b}}\hat{u} := \pi_h^{\mathrm{i}}u|_{\mathcal{F}_h}$ which is contained in $\hat{\mathcal{V}}_h$. Setting $\boldsymbol{e}_\pi^{\boldsymbol{u}} := \boldsymbol{u} - \pi_h\boldsymbol{u}$ then, the following holds*

$$\|\boldsymbol{e}_\pi^{\boldsymbol{u}}\|_* \ \textit{(or equiv.)}\ |||\boldsymbol{e}_\pi^{\boldsymbol{u}}||| \leq C_{\boldsymbol{\alpha}}h^{\mu-1}|u|_{\mu,\mathcal{T}_h}, \tag{2.26}$$

where $\mu := \min(k+1,s)$ and $C_{\boldsymbol{\alpha}} := C\|\boldsymbol{\alpha}^{1/2}\|_{\infty,\Omega}$.

Proof. Setting $\boldsymbol{e}_\pi^{\boldsymbol{u}} := (e_\pi^u, \hat{e}_\pi^u)$ where $e_\pi^u := u - \pi_h^{\mathrm{i}}u$ and $\hat{e}_\pi^u := \hat{u} - \pi_h^{\mathfrak{b}}\hat{u}$, and using the definition of the $\|\cdot\|_*$-norm (2.9) yields

$$\|\boldsymbol{e}_\pi^{\boldsymbol{u}}\|_*^2 = \|\boldsymbol{\alpha}^{1/2}\boldsymbol{\nabla}_h e_\pi^u\|_{0,\mathcal{T}_h}^2 + |\boldsymbol{e}_\pi^{\boldsymbol{u}}|_\tau^2. \tag{2.27}$$

The last term of (2.27), i.e., the jump semi-norm of $\boldsymbol{e}_\pi^{\boldsymbol{u}}$, is null according to the conformity of the (continuous) interpolant $\pi_h\boldsymbol{u}$, i.e., $\|e_\pi^u - \hat{e}_\pi^u\|_{0,\partial E} = 0$ (see, e.g., [Wells 2011, Lemma 5.5, p. 102]). Thus, successively using the Cauchy–Schwarz inequality, and the interpolation estimate (2.25a) yields

$$\|\boldsymbol{e}_\pi^{\boldsymbol{u}}\|_*^2 \leq \|\boldsymbol{\alpha}^{1/2}\|_{\infty,\Omega}^2|e_\pi^u|_{1,\mathcal{T}_h}^2 \leq C^2\|\boldsymbol{\alpha}^{1/2}\|_{\infty,\Omega}^2h^{2\mu-2}|u|_{\mu,\mathcal{T}_h}^2. \tag{2.28}$$

The proof of the second estimate follows by the same arguments. Successively using the definition of the continuity norm (2.19), and the estimates (2.25b) (with $\alpha = 1$) and (2.28),

we thus infer that

$$\begin{aligned}
|||\boldsymbol{e}_\pi^{\boldsymbol{u}}|||^2 &\overset{(2.19)}{:=} \|\boldsymbol{e}_\pi^{\boldsymbol{u}}\|_*^2 + \sum_{E\in\mathcal{T}_h} h_E \|\boldsymbol{\alpha}^{1/2}\boldsymbol{\nabla}_h e_\pi^u\|_{0,\partial E}^2, \\
&\overset{(2.25)}{\leq} \|\boldsymbol{\alpha}^{1/2}\|_{\infty,\Omega}^2 \sum_{E\in\mathcal{T}_h} (|e_\pi^u|_{1,E}^2 + h_E \|\boldsymbol{\nabla}_h e_\pi^u\|_{0,\partial E}^2), \\
&\leq C^2 \|\boldsymbol{\alpha}^{1/2}\|_{\infty,\Omega}^2 h^{2\mu-2} |u|_{\mu,\mathcal{T}_h}^2,
\end{aligned}$$

which yields the assertion. ■

Remark 2.5. We emphasize the optimality of the estimate (2.26) given in Lemma 2.4.1, independently of δ. This is due to the continuous nature of the interpolant $\pi_h\boldsymbol{u}$, which by conformity belongs to the kernel of stabilization terms whose contributions can be suboptimal. Let us specify that this assertion is no longer ensured by using discontinuous interpolants. This estimate is in agreement with the one established by Wells in [Wells 2011] for $\delta = 0$.

2.4.1 Energy-norms error estimates

We now derive an error estimation of the discrete composite variable $\boldsymbol{u}_h$ in the natural energy-norms.

Theorem 2.1 (Energy-norm estimates)**.** *Let $\boldsymbol{u} \in \boldsymbol{\mathcal{V}}$ be the compact notation of the exact solution of the problem (2.2). We denote by $\boldsymbol{u}_h \in \boldsymbol{\mathcal{V}}_h$ the approximate solution of the discrete problem (2.10). Setting $\boldsymbol{e}_h^{\boldsymbol{u}} := \boldsymbol{u} - \boldsymbol{u}_h$ then, for any value of the parameter δ, the following estimates hold :*

$$\|\boldsymbol{e}_h^{\boldsymbol{u}}\|_* \ \textit{(or equiv.)}\ |||\boldsymbol{e}_h^{\boldsymbol{u}}||| \leq C_{\boldsymbol{\alpha}} h^{\mu+r_\delta-1} |u|_{\mu,\mathcal{T}_h}, \tag{2.29}$$

where $\mu := \min(k+1, s)$, $r_\delta := \min(0, \delta)$, and $C_{\boldsymbol{\alpha}} := C\|\boldsymbol{\alpha}^{1/2}\|_{\infty,\Omega}$.

Proof. By using the triangle inequality for the definition of the stability energy-norm (2.9), we easily infer that

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_* \leq \|\boldsymbol{u} - \pi_h\boldsymbol{u}\|_* + \|\pi_h\boldsymbol{u} - \boldsymbol{u}_h\|_*. \tag{2.30}$$

Only an upper bound on the last term of (2.30) remains to be established. Successively using the coercivity, Galerkin orthogonality, and boundedness, we then deduce that

$$\begin{aligned}
\frac{1}{2}\|\pi_h\boldsymbol{u} - \boldsymbol{u}_h\|_*^2 &\overset{(2.18)}{\leq} \mathcal{B}_h^{(\varepsilon)}(\pi_h\boldsymbol{u} - \boldsymbol{u}_h, \pi_h\boldsymbol{u} - \boldsymbol{u}_h), \\
&\overset{(2.16)}{=} \mathcal{B}_h^{(\varepsilon)}(\pi_h\boldsymbol{u} - \boldsymbol{u}, \pi_h\boldsymbol{u} - \boldsymbol{u}_h), \\
&\overset{(2.21)}{\leq} C_{\mathrm{bnd}} |||\boldsymbol{u} - \pi_h\boldsymbol{u}|||\,|||\pi_h\boldsymbol{u} - \boldsymbol{u}_h|||.
\end{aligned}$$

Finally, considering that $\pi_h\boldsymbol{u} - \boldsymbol{u}_h \in \boldsymbol{\mathcal{V}}_h$ and using Lemma 2.3.4, we obtain an upper bound of this term ;

$$\|\pi_h\boldsymbol{u} - \boldsymbol{u}_h\|_* \leq 2\rho C_{\mathrm{bnd}} |||\boldsymbol{u} - \pi_h\boldsymbol{u}|||. \tag{2.31}$$

Inserting into (2.31) into (2.30), we then infer that

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_* \leq (1 + 2\rho C_{\text{bnd}})|||\boldsymbol{u} - \pi_h \boldsymbol{u}|||.$$

The proof of the second estimate (in the continuity-norm (2.19)) follows by the same arguments. Finally proceeding as in Remark 2.4, we can conclude that there exists a constant $C > 0$ such that $1 + 2\rho C_{\text{bnd}} \leq Ch^{r_\delta}$, that we combine with the optimal error estimate (2.26) given in Lemma 2.4.1, hence yielding to the assertion. ■

Corollary 2.1 (Strong-regularity solutions)**.** *Besides the hypotheses of Theorem 2.1, assume* $u \in H_0^{k+1}(\Omega)$*. Then, we have the following estimate*

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_* \leq C_{\boldsymbol{\alpha}} h^{k+r_\delta} |u|_{k+1,\mathcal{T}_h}. \tag{2.32}$$

where $C_{\boldsymbol{\alpha}} := C\|\boldsymbol{\alpha}^{1/2}\|_{\infty,\Omega}$.

Proof. (Evident) ■

Remark 2.6. Following Di Pietro and Ern [Di Pietro 2011, Theorem 4.53, p. 160], since C in Corollary 2.1 is independent of $\boldsymbol{\alpha}$, the discrete method is said to be robust with respect to diffusion heterogeneities (observing that the energy-norms depend on $\boldsymbol{\alpha}$). The given estimate (2.32) indicates that the order of convergence in the energy-norm is linear and δ-dependent, i.e., suboptimal if $\delta < 0$ and optimal otherwise, as summarized in Figure 2.1a. This is in agreement with the estimate given by Fabien *et al.* in the particular case, $\delta = 0$ [Fabien 2020b, Theorem 3.5, p. 8].

2.4.2 L^2-norm error estimate

Using a standard Aubin–Nitsche duality argument, we now derive an improved L^2-error estimate of the H-IP method in terms of the parameter δ. To this end, we define an auxiliary function ψ as the solution of the adjoint problem :

$$-\boldsymbol{\nabla} \cdot (\boldsymbol{\alpha} \boldsymbol{\nabla} \psi) = u - u_h \quad \text{in } \Omega, \quad \text{and} \quad \psi = 0 \quad \text{on } \partial\Omega.$$

By assuming elliptic regularity, the following estimate holds :

$$\|\psi\|_{2,\Omega} \leq C_{\boldsymbol{\alpha}} \|u - u_h\|_{0,\Omega}, \tag{2.33}$$

where $C_{\boldsymbol{\alpha}}$ depends on the shape regularity (i.e., the convexity) of Ω and the distribution of $\boldsymbol{\alpha}$ inside it [Ern 2009]. The weak-adjoint problem is to find $\psi \in H^2(\Omega) \cap H_0^1(\Omega)$ such that

$$(\boldsymbol{\alpha} \boldsymbol{\nabla}_h \psi, \boldsymbol{\nabla}_h v)_{0,\mathcal{T}_h} - \langle \boldsymbol{\alpha} \boldsymbol{\nabla}_h \psi \cdot \boldsymbol{n}, v \rangle_{0,\partial\mathcal{T}_h} = (u - u_h, v)_{0,\mathcal{T}_h}, \tag{2.34}$$

for all $v \in H_0^1(\Omega)$. Let us now introduce the composite error variable $\boldsymbol{e}_h^{\boldsymbol{u}} := \boldsymbol{u} - \boldsymbol{u}_h = (e_h^u, \hat{e}_h^u)$ where $e_h^u := u - u_h$ and $\hat{e}_h^u := \hat{u} - \hat{u}_h$. By setting now $v := e_h^u$ in (2.34), we obtain

$$\|e_h^u\|_{0,\mathcal{T}_h}^2 = (\boldsymbol{\alpha} \boldsymbol{\nabla}_h \psi, \boldsymbol{\nabla}_h e_h^u)_{0,\mathcal{T}_h} - \langle \boldsymbol{\alpha} \boldsymbol{\nabla}_h \psi, e_h^u \boldsymbol{n} \rangle_{0,\partial\mathcal{T}_h}. \tag{2.35}$$

From the regularity of the variables $\hat{u}$, $\hat{u}_h$ and ψ, we deduce that $\langle \boldsymbol{\alpha}\boldsymbol{\nabla}_h\psi, \hat{e}_h^u\boldsymbol{n}\rangle_{0,\partial\mathcal{T}_h} = 0$. By embedding this condition in (2.35), we obtain an equivalent reformulation of the weak-adjoint problem in terms of the discrete bilinear operator $\mathcal{B}_h^{(\varepsilon)}$:

$$\begin{aligned}\|e_h^u\|_{0,\mathcal{T}_h}^2 &= (\boldsymbol{\alpha}\boldsymbol{\nabla}\psi, \boldsymbol{\nabla}e_h^u)_{0,\mathcal{T}_h} - \langle \boldsymbol{\alpha}\boldsymbol{\nabla}\psi, [\![\boldsymbol{e}_h^{\boldsymbol{u}}]\!]\rangle_{0,\partial\mathcal{T}_h},\\ &= \mathcal{B}_h^{(\varepsilon)}(\boldsymbol{\psi}, \boldsymbol{e}_h^{\boldsymbol{u}}),\end{aligned} \tag{2.36}$$

where $\boldsymbol{\psi} := (\psi, \hat{\psi})$. Following the definition of the bilinear form $\mathcal{B}_h^{(\varepsilon)}$ (2.11) and using the Galerkin orthogonality $\mathcal{B}_h^{(\varepsilon)}(\boldsymbol{e}_h^{\boldsymbol{u}}, \pi_h\boldsymbol{\psi}) = 0$, since $\pi_h\boldsymbol{\psi} \in \boldsymbol{\mathcal{V}}_h$ (see Proposition 2.1), we easily infer

$$\begin{aligned}\mathcal{B}_h^{(\varepsilon)}(\boldsymbol{\psi}, \boldsymbol{e}_h^{\boldsymbol{u}}) &= \mathcal{B}_h^{(\varepsilon)}(\boldsymbol{e}_h^{\boldsymbol{u}}, \boldsymbol{e}_\pi^{\boldsymbol{\psi}}) - (1-\varepsilon)\langle \boldsymbol{\alpha}\boldsymbol{\nabla}\psi, [\![\boldsymbol{e}_h^u]\!]\rangle_{0,\partial\mathcal{T}_h},\\ &:= \mathcal{T}_1 - (1-\varepsilon)\mathcal{T}_2,\end{aligned} \tag{2.37}$$

where $\boldsymbol{e}_\pi^{\boldsymbol{\psi}} := \boldsymbol{\psi} - \pi_h\boldsymbol{\psi}$. We will now determine an upper bound of the quantity $\|e_h^u\|_{0,\mathcal{T}_h}^2$. Owing to Lemmas 2.3.5 and 2.4.1 and using the regularity assumption $\psi \in H^2(\Omega)$, we can bound the first term $\mathcal{T}_1$:

$$|\mathcal{T}_1| \leq C_{\text{bnd}}|||\boldsymbol{e}_\pi^{\boldsymbol{\psi}}|||\,|||\boldsymbol{e}_h^{\boldsymbol{u}}||| \leq C_{\boldsymbol{\alpha}}C_{\text{bnd}}h\|\psi\|_{2,\Omega}|||\boldsymbol{e}_h^{\boldsymbol{u}}|||. \tag{2.38}$$

Using the trace inequality $\|\boldsymbol{\nabla}_h\psi\|_{0,\partial\mathcal{T}_h} \leq Ch^{-1/2}\|\psi\|_{2,\Omega}$ [Ciarlet 1991], the second term $\mathcal{T}_2$ can be bounded as follows :

$$|\mathcal{T}_2| \leq C_{\boldsymbol{\alpha}}h^{\frac{1+\delta}{2}}\|\boldsymbol{\nabla}_h\psi\|_{0,\partial\mathcal{T}_h}|\boldsymbol{e}_h^{\boldsymbol{u}}|_\tau \leq C_{\boldsymbol{\alpha}}h^{\frac{\delta}{2}}\|\psi\|_{2,\Omega}|||\boldsymbol{e}_h^{\boldsymbol{u}}|||. \tag{2.39}$$

Combining (2.38) and (2.39), we obtain the estimate

$$\|u - u_h\|_{0,\mathcal{T}_h} \leq C_{\boldsymbol{\alpha}}(C_{\text{bnd}}h + (1-\varepsilon)h^{\frac{\delta}{2}})|||\boldsymbol{e}_h^{\boldsymbol{u}}|||, \tag{2.40}$$

and we can assert the theorem below.

Theorem 2.2 ($\boldsymbol{L^2}$-norm error estimate)**.** *Besides the hypotheses of Theorem 2.1 then, we have the estimate*

$$\|u - u_h\|_{0,\mathcal{T}_h} \leq C_{\boldsymbol{\alpha}}h^{\mu+s_\delta^{(\varepsilon)}}|u|_{\mu,\mathcal{T}_h}, \tag{2.41}$$

where the constant $C_{\boldsymbol{\alpha}}$ depends on the shape regularity of Ω and the distribution of $\boldsymbol{\alpha}$ inside it, and $s_\delta^{(\varepsilon)}$ is only dependent on ε and δ and is given by

$$s_\delta^{(\varepsilon)} := \begin{cases}\min(0, 2\delta) \equiv 2r_\delta & \text{if} \quad \varepsilon = 1,\\ \min(0, \delta/2 - 1) & \text{if} \quad \varepsilon \neq 1 \quad \text{and} \quad \delta \geq 0,\\ \min(2\delta, 3\delta/2 - 1) & \text{if} \quad \varepsilon \neq 1 \quad \text{and} \quad \delta < 0.\end{cases} \tag{2.42}$$

Proof. The estimate (2.41) using (2.42) follows after some algebraic manipulations from the previous equation (2.40), the definition of C_{bnd} given in Lemma 2.3.5 and the error estimate (in the $|||\cdot|||$-norm (2.19)) given in Theorem 2.1. ■

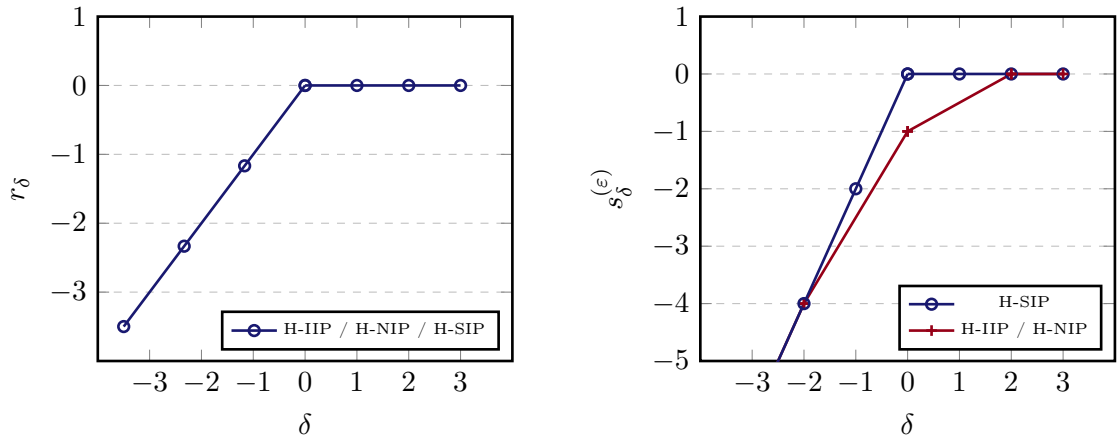


FIGURE 2.1 – Representation of the quantities r_δ and $s_\delta^{(\varepsilon)}$ vs. δ given in Theorems 2.1 and 2.2, respectively.

Let us emphasize that H-IP methods inherit similar asymptotic behaviors than their standard IPDG counterparts. Due to the lack of symmetry of both H-IIP and H-NIP schemes, the a priori error estimates in L^2 are optimal only if $\delta \geq 2$ (see Figure 2.1b). We also point out that the estimate given in Theorem 2.2 is in agreement with previous results established by different authors in the literature in the specific case $\delta = 0$ (see, e.g., [Fabien 2020b, Theorem 3.6, p. 9] and [Wells 2011, Lemma 5.5, p. 103] for the H-SIP method).

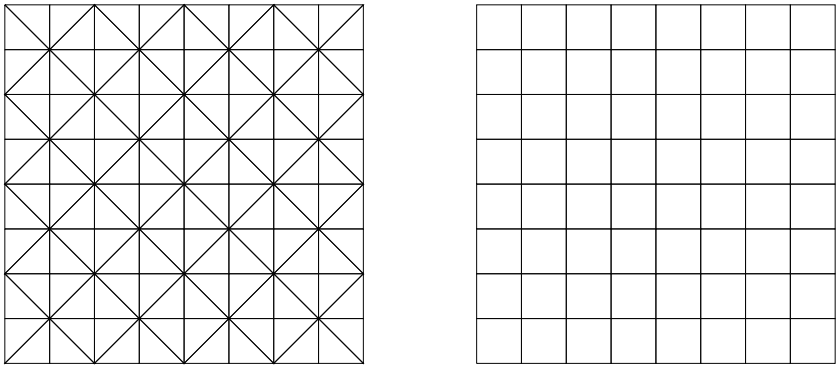
Remark 2.7. The authors are certain that these estimates given in Theorems 2.1 and 2.2 have already been established in the literature, but we have not been able to find them.

2.5 Numerical experiments

In the previous sections, we built families of hybridizable interior penalty methods based on an adaptive definition of the penalty parameter that depends on several coefficients. This section highlights the benefit these methods provide in the approximation of diffusion problems with anisotropic and/or discontinuous coefficients and in the validation of a priori error estimates. All of our numerical experiments are performed using the high-performance finite element library called NGSolve containing a rich Python interface [Schöberl 2014]. All developed source codes (Python's files) are available for free download from the following Github repository [3] and can be easily used by the interested reader to duplicate the numerical experiments presented below. Then, the physical domain is taken to be a unit square –i.e., $\Omega := [0,1]^2 \subset \mathbb{R}^2$– and the right-hand-side f is chosen such that the given exact solution u respecting the homogeneous boundary conditions is verified. We use a sequence of subdivisions $\mathcal{T}_h$, where regular triangles or squares form each partition (see, e.g., Figure 2.2). Standard h- and k-refinement strategies are used to compute the numerical errors and estimated convergence rates (ECRs). To pursue our quantitative analysis, we first measure the impact of the parameter δ on the a posteriori error estimates. Second, we point out the

3. https://github.com/GregoryETANGSALE/Variable_Diffusion_HIP

crucial role of the factor α_n arising in (2.12) for the robustness of the H-IP methods when the medium becomes highly anisotropic and/or discontinuous. Finally, we complete our experiments by pointing out some unexpected benefits of the value of τ_0 for the ECRs of the H-SIP scheme.



(a) (b)

FIGURE 2.2 – Uniform triangular (a) and square (b) meshes with $h = 1/8$, respectively.

2.5.1 Test A : Influence of the parameter δ

We consider the following test case, which was previously proposed in [Fabien 2020b] : the diffusion tensor is homogeneous and isotropic – $\boldsymbol{\alpha} := \boldsymbol{I}_2$ (identity matrix) – and the exact smooth solution is given by $u(x,y) = xy(1-x)(1-y)e^{(-x^2-y^2)}$. Then, for all $E \in \mathcal{T}_h$ and for all $F \in \mathcal{F}_E$, we assume that the penalty parameter has the following simplified form :

$$\tau_{E,F} := \frac{\tau_0}{h_E^{1+\delta}}, \tag{2.43}$$

where $\tau_0 > 0$ is a positive constant chosen to be large enough in accordance with Lemma 2.3.3. The objective here is to measure the impact of the parameter δ on the ECRs in both the L^2- and energy-norms. A history of convergence is shown in Figures 2.3 ($|||\cdot|||$-norm) and 2.4 ($\|\cdot\|_{0,\mathcal{T}_h}$-norm) for uniform triangular meshes and for polynomial degrees $k \in \{1,\ldots,3\}$. As expected, these observations are in agreement with theoretical estimates and underline that the stabilization parameter δ influences the convergence rate. In particular, we recover some well-known estimates if $\delta = 0$. First, we notice that the convergence of the H-IP method in the energy-norm is linearly δ-dependent if $\delta \leq 0$ and optimal if $\delta \geq 0$, which is in accordance with Lemma 2.1 (see Figure 2.3). A brief analysis of the convergence in the L^2-norm indicates that both the H-IIP and H-NIP schemes behave differently from the H-SIP scheme. Non-symmetric variants are strongly influenced by the polynomial parity of k and by the penalty parameter δ. We observe that the convergence rate increases linearly and optimally if $\delta \geq 0$ for odd k and $\delta \geq 2$ for even k. In this last case, let us point out that the optimal convergence is nearly reached once $\delta \geq 1$. As expected, the Symmetric scheme converges optimally when $\delta \geq 0$. These results agree with the theoretical results established in Theorem 2.2.

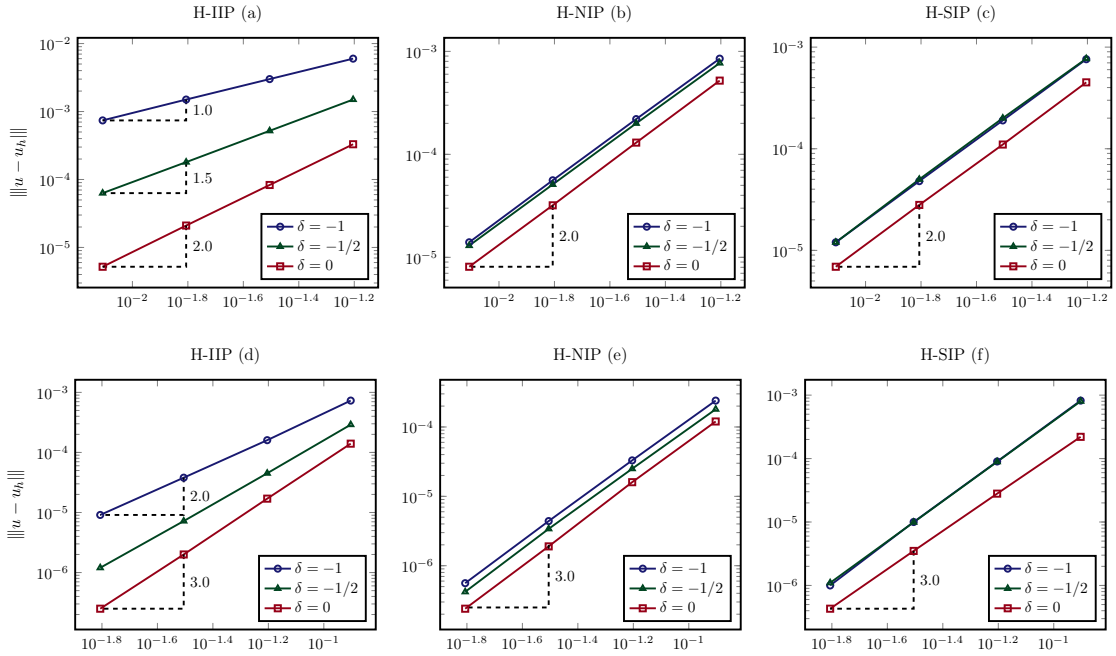


FIGURE 2.3 – Test A : history of convergence in the $|||\cdot|||$-norm (vs. h) of the the H-IIP (a-d), H-NIP (b-e) and H-SIP (c-f) schemes, on uniform triangular meshes with $-1 \leq \delta \leq 0$. In the top images (a-b-c), $k = 2$, and in the bottom images (d-e-f), $k = 3$.

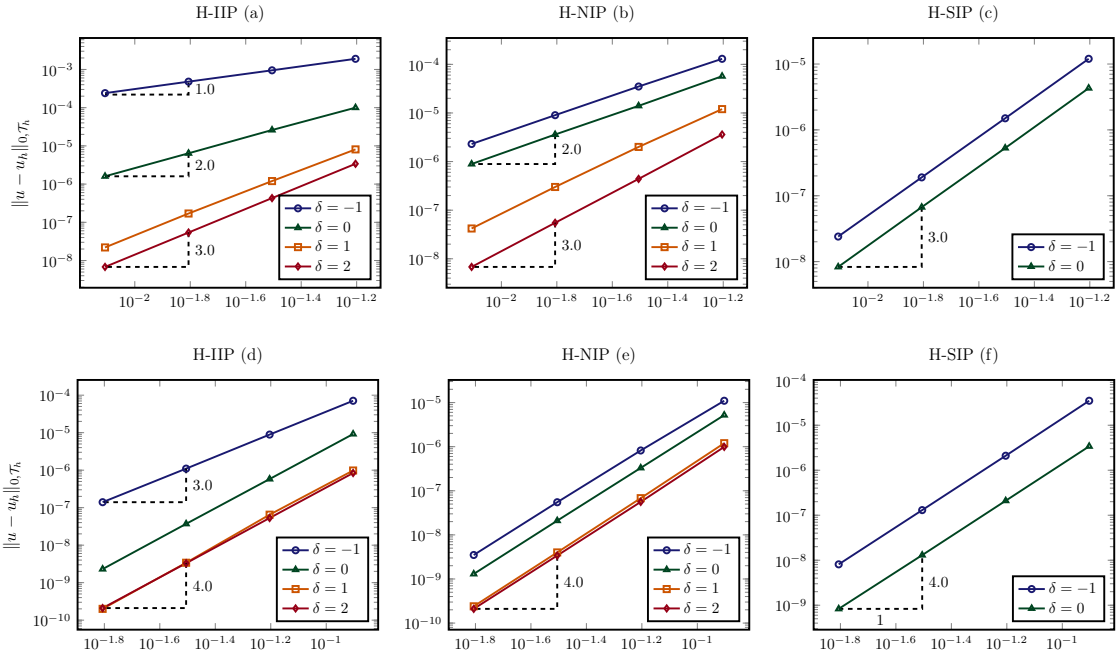


FIGURE 2.4 – Test A : history of convergence in the L^2-norm (vs. h) of the the H-IIP (a-d), H-NIP (b-e) and H-SIP (c-f) schemes, on uniform triangular meshes with $-1 \leq \delta \leq 2$. In the top images (a-b-c), $k = 2$, and in the bottom images (d-e-f), $k = 3$.

2.5.2 Test B : Influence of the parameter $\boldsymbol{\alpha}_{E,F}$

In the second experiment, we analyze the behavior of the discretization method in the context of genuine anisotropic and heterogeneous properties. Then, the unit square Ω is split into four subdomains $\Omega_1 = [0, 1/2]^2$, $\Omega_2 = [1/2, 1] \times [0, 1/2]$, $\Omega_3 = [1/2, 1]^2$ and $\Omega_4 = [0, 1/2] \times [1/2, 1]$, such that $\Omega := \cup_{i=1}^{4} \Omega_i$ as illustrated in Figure 2.5. The exact solution on the whole domain Ω is given by $u(x, y) = \sin(\pi x) \sin(\pi y)$, and the diffusivity tensor takes different values in each subregion :

$$\boldsymbol{\alpha} = \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix} \quad \text{for } (x, y) \in \Omega_1,\ \Omega_3, \quad \text{and,}$$
$$\boldsymbol{\alpha} = \begin{bmatrix} \lambda^{-1} & 0 \\ 0 & 1 \end{bmatrix} \quad \text{for } (x, y) \in \Omega_2,\ \Omega_4,$$

where the parameter $\lambda > 0$ simultaneously controls both the anisotropy and the medium heterogeneity. Here, we focus on the influence of the parameter $\alpha_{E,F}$ on the robustness of the

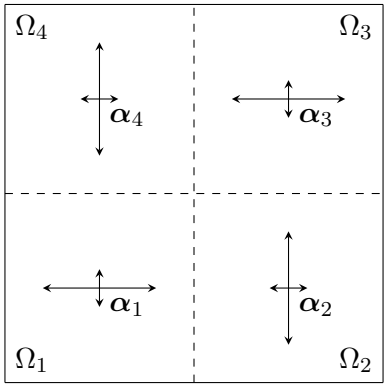


FIGURE 2.5 – Description of test B with genuine anisotropic and heterogeneous properties.

discretization method in the context of highly anisotropic and heterogeneous coefficients, and we choose $\lambda = 10^{-3}$. In this context, the anisotropy and heterogeneity ratios are approximately 10^3 and 10^6, respectively. For the simulations, we consider a conforming triangular mesh ($h = 1/32$) respecting the discontinuities of $\boldsymbol{\alpha}$, we use piecewise linear approximations of the discrete variable u_h, and we set $\delta = 0$ in the definition of the penalty parameter (2.12). Here, the comparisons are only graphical (Figure 2.6). We depict the discrete solutions u_h obtained successively using $\alpha_{E,F} := 1$ (Case 1) and $\alpha_{E,F} := \boldsymbol{n}_{E,F} \boldsymbol{\alpha}_E \boldsymbol{n}_{E,F}$ (Case 2) for all variations of $\varepsilon \in \{0, \pm 1\}$. In the first situation (Figures 2.6a, 2.6b and 2.6c), the discrete solutions exhibit spurious oscillations and erratic behaviors, thus violating the discrete maximum principle. This can be easily explained by observing that the first formulation does not distinguish between the principal directions of the diffusivity tensor. Consequently, a misestimated penalty is applied in directions of low or high diffusivity. In the second situation (Figures 2.6d, 2.6e and 2.6f), the jumps in diffusivity are better captured at the interfaces of discontinuities, and the discrete solutions are significantly more robust, i.e., exhibit less erratic behavior.

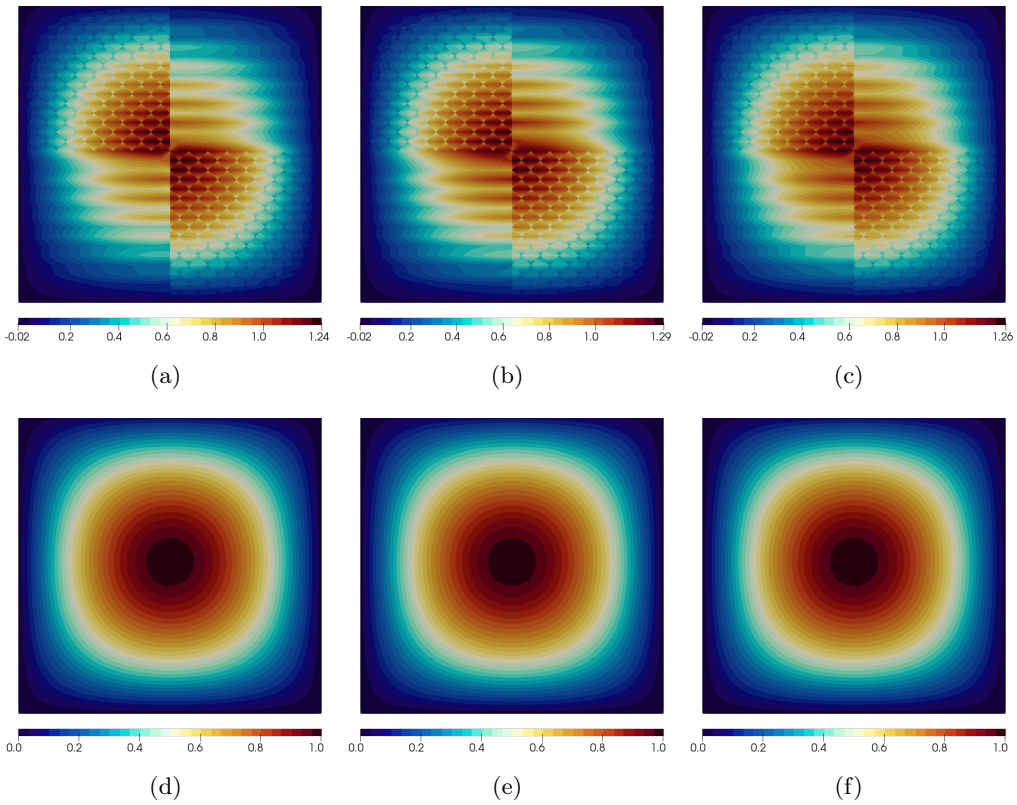


FIGURE 2.6 – Test B : representation of the discrete solution u_h obtained by the H-IIP (a-d), H-NIP (b-e), and H-SIP (c-f) schemes, respectively, on the structured triangular mesh ($h = 1/32$). On the top (a-b-c), the parameter $\alpha_{E,F}$ is chosen as $\alpha_{E,F} := 1$, and on the bottom (d-e-f), $\alpha_{E,F} := \boldsymbol{n}_{E,F}\boldsymbol{\alpha}_E\boldsymbol{n}_{E,F}$.

2.5.3 Test C : Influence of the parameter τ_0

To conclude the sequence of numerical tests, we analyze the influence of the parameter τ_0 on the convergence of the H-SIP method for $\boldsymbol{\alpha}$-orthogonal grids only. For simplicity, we consider the same test case as Test B, (2.5.2), and we set two values of the parameter λ : (i) $\lambda = 1$ for a homogeneous and isotropic media and (ii) $\lambda = 0.1$ for a heterogeneous and anisotropic media. We plot the computed L^2-error of the H-SIP method for a wide range of values of the parameter τ_0 – i.e., $1 \leq \tau_0 \leq 6$ – using a uniform square mesh ($h = 1/32$). The analysis is done for polynomial degrees $1 \leq k \leq 4$, but the results are presented for $k = \{1, 2\}$ only. Analyzing Figure 2.7, we observe that there exists an optimal value of the parameter $\tau_0 := \tau_{\text{opt}}$ that minimizes the L^2-error of the scheme. In the context of $\boldsymbol{\alpha}$-orthogonal grids, this optimal value ($\tau_{\text{opt}} = 2$) is insensitive to the mesh form, the mesh size h, the polynomial degree k, and the heterogeneity and/or anisotropy of the media λ. A history of the convergence of the H-SIP method using $\tau_{\text{opt}} = 2$ is then summarized in Table 2.1, and we note the surprising superconvergence of u_h $(k+2)$ in the discrete L^2-norm obtained without any postprocessing. We emphasize that the *superconvergence* property is not achieved for any triangular mesh or any value of the parameter $\varepsilon \neq 1$, even using the optimal parameter τ_{opt} in (2.12).

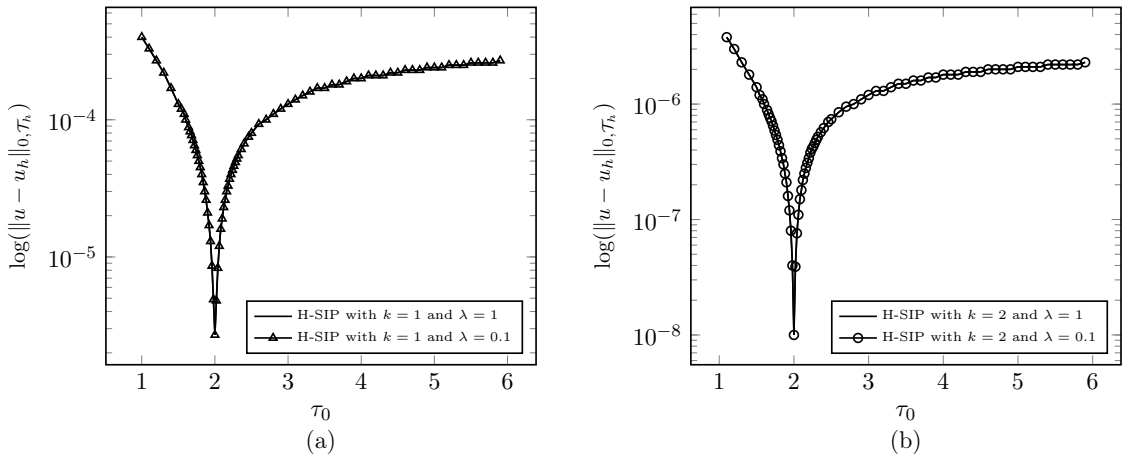


FIGURE 2.7 – Test C : the L^2-error of the H-SIP method vs. τ_0 for a uniform square mesh using piecewise linear (a) and quadratic (b) approximations.

TABLE 2.1 – Test C : history of convergence $\|u-u_h\|_{0,\mathcal{T}_h}$ (vs. h) of the H-SIP method using the optimal parameter α_{opt} on uniform square meshes

	$k=1$				$k=2$			
	$\lambda=1$		$\lambda=0.1$		$\lambda=1$		$\lambda=0.1$	
h^{-1}	$\|u-u_h\|_{0,\mathcal{T}_h}$	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
8	$1.7e-04$	–	$1.7e-04$	–	$2.6e-06$	–	$2.6e-06$	–
16	$2.1e-05$	3.00	$2.1e-05$	3.00	$1.6e-07$	3.99	$1.6e-07$	3.99
32	$2.7e-06$	3.00	$2.7e-06$	3.00	$1.0e-08$	4.00	$1.0e-08$	4.00
64	$3.4e-07$	3.00	$3.4e-07$	3.00	$6.4e-10$	4.00	$6.4e-10$	4.00

2.6 Conclusion

We derive improved *a priori* error estimates of families of hybridizable interior penalty discontinuous Galerkin methods using a variable penalty to solve highly anisotropic diffusion problems. The convergence analysis highlights the h^{δ}-dependency of the coercivity condition and the boundedness requirement that strongly impacts the derived error estimates in terms of both energy- and L^2-norms. The optimal convergence of the energy-norm is proven for any penalty parameter $\delta \geq 0$ and $\varepsilon \in \{0, \pm 1\}$. The situation is somewhat different in L^2, and distinctive features can be found between the three schemes. Indeed, the Symmetric method theoretically converges optimally if $\delta \geq 0$, and Non-Symmetric variants converge only if $\delta \geq 2$ independently of the polynomial parity. All of these estimates are corroborated by numerical evidence. Notably, the superconvergence of the H-SIP scheme is achieved for $\boldsymbol{\alpha}$-orthogonal grids without any postprocessing but only if an appropriate τ_0 is selected.

Link to the Reproducible Capsule

https://codeocean.com/capsule/7350536/tree/v1

Chapitre 3

Performances of Hybridized-, Embedded-, and Weighted-Interior Penalty Discontinuous Galerkin methods for Diffusion Problems

---

Article published in *Water and Hydrocomplexity*, a section of the journal *Frontiers in Water* under the title
Performances of Hybridized-, Embedded-, and Weighted-Interior Penalty Discontinuous Galerkin Methods for Heterogeneous and Anisotropic Diffusion Problems
doi:10.3389/frwa.2021.716459

Grégory Etangsale [1], Vincent Fontaine[1], Nalitiana Rajaonison[1]

Abstract

The present paper discusses families of Interior Penalty Discontinuous Galerkin (IP) methods for solving heterogeneous and anisotropic diffusion problems. Specifically, we focus on distinctive schemes, namely the Hybridized-, Embedded-, and Weighted-IP schemes, leading to final matrixes of different sizes and sparsities. Both the Hybridized- and Embedded-IP schemes are eligible for static condensation, and their degrees of freedom are distributed on the mesh skeleton. In contrast, the unknowns are located inside the mesh elements for the Weighted-IP variant. For a given mesh, it is well-known that the number of degrees of freedom related to the standard Discontinuous Galerkin methods increases more rapidly than those of the skeletal approaches (Hybridized- and Embedded-IP). We then quantify the impact of the static condensation procedure on the computational performances of the different IP classes in terms of robustness, accuracy and CPU time. To this aim, numerical experiments are investigated by considering strong heterogeneities and anisotropies. We analyze the fixed error tolerance versus the run time and mesh size to guide our performance criterion. We also outlined some relationships between these Interior Penalty schemes. Eventually, we confirm the superiority of the Hybridized- and Embedded-IP schemes, regardless of the mesh, the polynomial degree, and the physical properties (homogeneous, heterogeneous and/or anisotropic).

1. Department of Building and Environmental Sciences, University of La Réunion - South Campus, France

Contents

3.1 Introduction

The Discontinuous Galerkin (DG) methods were firstly introduced by Reed and Hill in [Reed 1973] for the neutron transport phenomenon. Since their introduction, the DG methods have become a relevant class of finite element schemes for modeling physical processes. They provide several advantages : they are locally conservative and eligible to hk-refinement strategies and they consider a discontinuous piecewise (polynomial) approximation of the exact solution [Arnold 2002, Rivière 2008, Di Pietro 2008]. During the 1980s, Arnold proposed the famous Interior Penalty Discontinuous Galerkin (IPDG) method for solving the second-order elliptic problem [Arnold 1982]. Even if the stability and robustness of the IP method have been proven for homogeneous media [Wheeler 1978, Rivière 2008], these benefits are no longer ensure in the presence of high heterogeneous, and/or anisotropic ratios [Burman 2006, Ern 2009]. To restore the robustness, these authors replace the standard average operator, in the IP formalism, with a weighted one accounting for the diffusivity in the normal direction. The resulting discretization scheme is well-known as the Weighted-Interior Penalty (W-IP) method. However, despite all these beneficial features, the DG methods have been recognized to be generally more expensive than the standard Conforming Finite Element methods such as the Continuous Galerkin (CG) and the Mixed Finite Element [Peraire 2008]. The final matrix system of DG methods leads to a larger stencil with a higher number of coupled degrees of freedom (DOFs), which is quite challenging for large-scale problems [Rivière 2008]. Following these observations, Cockburn et al. [Cockburn 2009b] introduced a new DG discretization scheme to overcome these drawbacks called the Hybridized Discontinuous Galerkin (HDG) methods.

The HDG methods can be considered as a DG methods that are eligible for static condensation. An additional trace variable is introduced to approximate the exact solution on the mesh skeleton. This original unknown also belongs to the set of piecewise discontinuous functions [Nguyen 2009, Fabien 2020b, Cockburn 2012c, Wells 2011]. Thus, the global coupled linear system can be reduced by static condensation only to the interface-based DOFs located on the mesh skeleton [Cockburn 2009a, Nguyen 2009, Lehrenfeld 2016]. Besides, the HDG methods provide a smaller and sparser matrix system, and they inherit the benefits of the traditionally DG methods. They have proven to be more robust and efficient than the standard DG schemes for many situations. Moreover, superconvergence of discrete variables can be attained by employing an appropriate local postprocessing technique [Nguyen 2009, Cockburn 2012c, Dijoux 2019b]. Recently, an alternative version of the HDG method has been developed to reduce the size of the final matrix system [Cockburn 2009c]. The corresponding discretization scheme is called the Embedded Discontinuous Galerkin method (EDG) and extends the concept of the trace variable to the set of continuous functions. This alternative approach providing less CPU time and a tighter stiffness matrix compared to the original HDG framework [Zhang 2019]. However, by enforcing the continuity of the discrete trace, the EDG method may have to afford a loss of accuracy for approximating the discrete solution [Cockburn 2009c, Fabien 2020b, Zhang 2019].

This manuscript compares the Hybridized-, Embedded-, and Weighted-Interior Penalty formalisms in terms of accuracy, efficiency, and computational cost. Previous works offered comparisons between the standard CG or DG versus the HDG methods (see, e.g., [Woopen 2014, Fidkowski 2016, Fidkowski 2019]) favoring the latest formalism at different scales. However, these studies are limited to basic situations that do not include the high variations of the physical properties of the medium (i.e., heterogeneity and/or anisotropy). The present work focuses specifically on the primal class of Interior Penalty method for heterogeneous and/or anisotropic diffusion problems. We provide an appropriate stabilization function inspired by [Etangsale 2022], which employs the normal diffusivity coefficient to improve the robustness of the IP method. Then, by quantifying the computational cost of the H-IP, E-IP, and W-IP schemes, we prove the effects of the static condensation on the global matrix sparsity and CPU time for the Hybridized- and Embedded-IP methods. Clearly and without ambiguity, we demonstrate the superiority of both skeletal methods (Hybridized- and Embedded-IP) over the traditional DG techniques (standard IP and W-IP). Besides, we established some relationships between the standard IP, W-IP, and H-IP schemes, whose results are similar to those discussed by [Fabien 2020b] for a homogeneous permeability tensor. We also emphasize the total equivalence properties between the Incomplete H-IP and W-IP schemes assuming a specific definitions of (i) the weighting function and (ii) the penalization function. At least, we provided numerical experiments to corroborate the observations mentioned above, and to prove the robustness of each schemes considering heterogeneous and/or anisotropic characteristics.

The material is organized as follows. In section 3.2, we describe the model problem and precise our notations and the discrete settings. In section 3.3, we briefly describe the Hybridized-, Embedded-, and Weighted-Interior Penalty formalisms. We then quantify the number of globally coupled degrees of freedom of each schemes for a given mesh and a fixed polynomial degree. In section 3.4, numerical experiments are investigated using h- and k-refinement stra-

tegies, and we compare the computational performances of the IP methods for a wide range of heterogeneity and anisotropy. We conduct discussions to identify the most suitable IP formalism in terms of robustness, accuracy and computing time. Finally, we end with some concluding remarks and perspectives.

3.2 Problem statement and notations

3.2.1 Model problem

Let Ω be a two-dimensional convex polygonal domain, and we denote by $\Gamma := \partial\Omega$ its boundary such that, $\Gamma := \Gamma_D \cup \Gamma_N$ and $\Gamma_D \cap \Gamma_N = \emptyset$. Here, Γ_D and Γ_N represent the boundary parts of the domain where Dirichlet and Neumann boundary conditions are provided, respectively. The governing equation is described by the second-order elliptic problem

$$\boldsymbol{\nabla} \cdot (-\boldsymbol{\alpha}\boldsymbol{\nabla} u) = f \quad \text{in } \Omega, \tag{3.1}$$

with the following boundary conditions,

$$u = \bar{u} \quad \text{on } \Gamma_D, \quad \text{and} \quad (-\boldsymbol{\alpha}\boldsymbol{\nabla} u) \cdot \boldsymbol{n} = \bar{\sigma} \quad \text{on } \Gamma_N,$$

where f represents a source or sink term, $\bar{u}$ and $\bar{\sigma}$ the prescribed Dirichlet and Neumann boundary data, respectively. In the context of heterogeneous and anisotropic processes, the permeability tensor $\boldsymbol{\alpha}$ is assumed to be a symmetric positive-definite matrix-valued function.

3.2.2 Mesh assumptions

Let us introduce some discrete notations associated to the partition of the domain. We consider the family $\mathcal{T}_h := \{\cup E\}$, which consists of a collection of polygonal elements. We specify here that the set $\mathcal{T}_h$ refers to an affine triangulation of the domain Ω. For all $E \in \mathcal{T}_h$, we denote by $|E|$ its measure and by h_E its diameter such that $h = \max_{E\in\mathcal{T}_h} h_E$. Let ∂E be the boundary of an element of the mesh $\mathcal{T}_h$. In the rest of the document, we will prefer the generic term *interface* in place of an *edge* of the triangulation. Let us denote by $\mathcal{F}_h^{\mathfrak{i}}$ and $\mathcal{F}_h^{\mathfrak{b}}$ the set of interior and boundary interfaces, respectively. We further assume that $\mathcal{F}_h^{\mathfrak{b}}$ coincides with the Dirichlet and Neumann boundary parts such that, $\mathcal{F}_h^{\mathfrak{b}} := \mathcal{F}_h^D \cup \mathcal{F}_h^N$ and $\mathcal{F}_h^D \cap \mathcal{F}_h^N$. The subsets $\mathcal{F}_h^D$ and $\mathcal{F}_h^N$ refers to the set of all interfaces lying entirely into the Dirichlet and Neumann boundary parts, respectively. The set of all interfaces is provided by the mesh skeleton as $\mathcal{F}_h := \mathcal{F}_h^{\mathfrak{i}} \cup \mathcal{F}_h^{\mathfrak{b}}$. In the same way, we introduce the collection of all boundary interfaces for each elements of the triangulation $\mathcal{T}_h$ such that, $\partial\mathcal{T}_h := \{\cup\partial E, \forall E \in \mathcal{T}_h\}$. For any given elements $E \in \mathcal{T}_h$, and for any edges $F \in \partial A$, we denote $\boldsymbol{n}_{E,F}$ as the unit normal vector to F pointing out of E.

3.2.3 Functional spaces

We consider the polygonal domain $\mathcal{D} \subset \mathbb{R}^2$, with boundary $\partial\mathcal{D}$. We denote by $(\cdot,\cdot)_{0,\mathcal{D}}$ and $\langle\cdot,\cdot\rangle_{0,\partial\mathcal{D}}$ the standard L^2-inner product in $L^2(\mathcal{D})$ and $L^2(\partial\mathcal{D})$, respectively. Thus, we

introduce the compact notations related to the discrete L^2-inner scalar product :

$$(\cdot,\cdot)_{0,\mathcal{T}_h} := \sum_{E\in\mathcal{T}_h} (\cdot,\cdot)_{0,E}, \quad \langle\cdot,\cdot\rangle_{0,\partial\mathcal{T}_h} := \sum_{E\in\mathcal{T}_h} \langle\cdot,\cdot\rangle_{0,\partial E} \quad \text{and} \quad \langle\cdot,\cdot\rangle_{0,\mathcal{F}_h} := \sum_{F\in\mathcal{F}_h} \langle\cdot,\cdot\rangle_{0,F}, \tag{3.2}$$

with $\|\cdot\|_{0,\mathcal{T}_h}$, $\|\cdot\|_{0,\partial\mathcal{T}_h}$ and $\|\cdot\|_{0,\mathcal{F}_h}$ corresponds to their respective norms. As usual in families of Discontinuous Galerkin methods, we consider the following broken space inside the elements for the discretization :

$$\mathcal{V}_h := \{v_h \in L^2(\mathcal{T}_h) \,:\, v_h|_E \in \mathcal{P}_k(E),\ \forall E \in \mathcal{T}_h\}, \tag{3.3}$$

where the discrete variable $v_h \in \mathcal{V}_h$ is defined within each elements of the triangulation $\mathcal{T}_h$. For the Hybridized discretization of the DG methods, the approximation requires an auxiliary variable $\hat{v}_h \in \hat{\mathcal{V}}_h^{\bar{u}}$ which is called the *trace* variable. The numerical trace is defined on the skeleton with respect to the imposed Dirichlet boundary conditions such that :

$$\hat{\mathcal{V}}_h^{\bar{u}} := \{\hat{v}_h \in L^2(\mathcal{F}_h) \,:\, \hat{v}_h|_F \in \mathcal{P}_k(F),\ \forall F \in \mathcal{F}_h \ \text{ and } \ \hat{v}_h|_F = \pi_h\bar{u},\ \forall F \in \mathcal{F}_h^D\}, \tag{3.4}$$

where π_h denotes the L^2-orthogonal projection onto $\mathcal{P}_k(F)$. We refer to $\hat{\mathcal{V}}_h^0$ for the set of polynomial functions defined on the skeleton, and vanishing on the Dirichlet boundary Γ_D. Here, $\mathcal{P}_k(X)$ denotes the space of polynomials at least degree $k \geq 1$ on X, where X corresponds to a generic element of $\mathcal{T}_h$ or $\mathcal{F}_h$.

3.2.4 Trace operators

We denote by $\boldsymbol{v}_h := (v_h, \hat{v}_h)$ the composite variable associated to the pair of discrete spaces $\mathcal{V}_h \times \hat{\mathcal{V}}_h^{\bar{u}}$. For all $E \in \mathcal{T}_h$ and $F \in \partial E$, we define the HDG-jump operator of the composite variable $\boldsymbol{v}_h \in \mathcal{V}_h \times \hat{\mathcal{V}}_h^{\bar{u}}$ across F as $[\![\boldsymbol{v}_h]\!]|_{E,F} := (v_h|_F - \hat{v}_h|_F)\boldsymbol{n}_{E,F}$. Since no confusions can arise, we omit the subscripts E and F from the definition, and we simply write $[\![\boldsymbol{v}_h]\!] := (v_h - \hat{v}_h)\boldsymbol{n}$. For any scalar– $v \in H^1(\mathcal{T}_h)$ and vector– $\boldsymbol{\varphi} \in [H^1(\mathcal{T}_h)]^2$ valued function, we introduce the standard DG-jump and weighted-average operators for each faces $F \in \mathcal{F}_h$ as follows,

$$[\![v]\!] := \begin{cases} v_1\boldsymbol{n}_1 + v_2\boldsymbol{n}_2, & \text{if} \quad F \in \mathcal{F}_h^{\mathrm{i}}, \\ v\boldsymbol{n}, & \text{if} \quad F \in \mathcal{F}_h^{\mathfrak{b}}, \end{cases} \quad \text{and} \quad \{\!\{\boldsymbol{\varphi}\}\!\}_{\boldsymbol{\omega}} := \begin{cases} \omega_1\boldsymbol{\varphi}_1 + \omega_2\boldsymbol{\varphi}_2, & \text{if} \quad F \in \mathcal{F}_h^{\mathrm{i}}, \\ \boldsymbol{\varphi} & \text{if} \quad F \in \mathcal{F}_h^{\mathfrak{b}}, \end{cases} \tag{3.5}$$

where $\boldsymbol{\omega} := (\omega_1, \omega_2)$ is a weighting function verifying that the weights satisfy $\omega_1 + \omega_2 = 1$. Similarly, we introduce the conjugate weighting function $\overline{\boldsymbol{\omega}} := (\omega_2, \omega_1)$ associated with $\boldsymbol{\omega}$. In particular, for the case $\boldsymbol{\omega} = (1/2, 1/2)$, we recover the classical average operator, and we will omit the subscript $\boldsymbol{\omega}$ in its definition.

3.3 Discretization of the Interior Penalty methods

This section describes three primal DG schemes, namely, the Hybridized-, Embedded-, and Weighted-Interior Penalty methods. We provide a compact formulation of the corresponding discrete bilinear forms and evaluates the computational cost of the schemes. For readers unfamiliar with the Discontinuous Galerkin approach, we recommend the studies proposed by [Arnold 2002], and [Cockburn 2009b], which outline the concepts and ideas behind the standard and Hybridized DG methods in a unified framework.

3.3.1 Hybridized- and Embedded-IP formulations

The HDG formalism as proposed in [Cockburn 2009b, Cockburn 2009a] supposes that the discrete trace $\hat{u}_h \in \hat{\mathcal{V}}_h^{\bar{u}}$ is evaluated distinctly with a piecewise polynomial functions on the boundaries of the elements. This specific definition of the discrete trace will produce a significant number of degrees of freedom, due to the discontinuous nature of its approximation on the mesh skeleton. To reduce the trace-based degrees of freedom, an alternative possibility consists of the restriction of the trace space to the set of continuous functions on the skeleton [Cockburn 2009c]. By considering the approximation $\hat{u}_h \in \tilde{\mathcal{V}}_h^{\bar{u}}$ with the continuous space $\tilde{\mathcal{V}}_h^{\bar{u}} = \hat{\mathcal{V}}_h^{\bar{u}} \cap \mathcal{C}^0(\mathcal{F}_h)$, then the discretization corresponds to the Embedded DG method [Cockburn 2009c, Fabien 2020b]. For approximating the problem (3.1), we introduce the two discrete composite spaces $\boldsymbol{\mathcal{V}}_{h,\mathrm{H}}^{\bar{u}} = \mathcal{V}_h \times \hat{\mathcal{V}}_h^{\bar{u}}$ and $\boldsymbol{\mathcal{V}}_{h,\mathrm{E}}^{\bar{u}} = \mathcal{V}_h \times \tilde{\mathcal{V}}_h^{\bar{u}}$ associated to the composite discrete variable $\boldsymbol{u}_h$. For clarity, Figure 3.1 gives an illustration of the distribution of degrees of freedom for the Hybridized- and Embedded-IP method with polynomial degree $k = 1$. Notice that the only difference between the Hybridized and Embedded scheme lies in the selection of the discrete trace space. Thus, the discretization consists to seek $\boldsymbol{u}_h \in \boldsymbol{\mathcal{V}}_{h,\mathrm{X}}^{\bar{u}}$ such that

$$\mathcal{B}_{h,\mathrm{X}}^{(\varepsilon)}(\boldsymbol{u}_h, \boldsymbol{v}_h) = (f, v_h)_{0,\mathcal{T}_h} + \langle \bar{\sigma}, \hat{v}_h \rangle_{0,\mathcal{F}_h^N} \quad \text{for all } \boldsymbol{v}_h \in \boldsymbol{\mathcal{V}}_{h,\mathrm{X}}^{0}, \tag{3.6}$$

where the index X corresponds to the Hybridized- (H) or Embedded- (E) IP method, according to the discrete composite space selected (i.e., discontinuous or continuous). The bilinear form $\mathcal{B}_{h,\mathrm{X}}^{(\varepsilon)}$ can be linearly decomposed as :

$$\begin{aligned}\mathcal{B}_{h,\mathrm{X}}^{(\varepsilon)}(\boldsymbol{u}_h, \boldsymbol{v}_h) :=& (\boldsymbol{\alpha}\boldsymbol{\nabla}_h u_h, \boldsymbol{\nabla}_h v_h)_{0,\mathcal{T}_h} - \langle \boldsymbol{\alpha}\boldsymbol{\nabla}_h u_h, [\![\boldsymbol{v}_h]\!] \rangle_{0,\partial\mathcal{T}_h} \\ &- \varepsilon \langle \boldsymbol{\alpha}\boldsymbol{\nabla}_h v_h, [\![\boldsymbol{u}_h]\!] \rangle_{0,\partial\mathcal{T}_h} + \langle \tau [\![\boldsymbol{u}_h]\!], [\![\boldsymbol{v}_h]\!] \rangle_{0,\partial\mathcal{T}_h}.\end{aligned}$$

Here, we reported the parameter ε, which is associated to the variations of the X-IP discretization, as mentioned in [Rivière 2008, Fabien 2020b]. For $\varepsilon = 1$, we retrieve the Symmetric scheme (X-SIP), and the cases $\varepsilon = \{-1, 0\}$ corresponds to both non-symmetric schemes, where $\varepsilon = -1$ (*resp.* $\varepsilon = 0$) denotes the Non-Symmetric (X-NIP) (*resp.* Incomplete (X-IIP)) variants. In particular, the last contribution of the bilinear form $\mathcal{B}_{h,\mathrm{X}}^{(\varepsilon)}$ refers to the penalization term with the penalization function τ. Naturally, the selection of the parameter τ is quite delicate, since it affects strongly the stability and the accuracy of the method (see [Etangsale 2022] for the investigation of the well-posedness and error estimates of the H-IP method). In the following, we give a specific definition of the stabilization function that is influenced by the diffusivity parameter $\boldsymbol{\alpha}$ such that :

$$\tau_{E,F} := \frac{\tau_0 \alpha_{E,F}(k+1)(k+2)}{h_E}, \quad \forall\, F \in \partial E, \tag{3.7}$$

where $\alpha_{E,F} = \boldsymbol{n}_{E,F}\boldsymbol{\alpha}_E\boldsymbol{n}_{E,F}$ is the normal diffusivity coefficient for any element $E \in \mathcal{T}_h$, and τ_0 a positive user-dependent constant. It is important to note that the penalization function is piecewise constant on $\partial\mathcal{T}_h$ and double-valued for any interfaces $F \in \mathcal{F}_h^{\mathrm{i}}$.

Remark 3.1 (Links with CG method). Assuming $\hat{u}_h^{\mathrm{E}}$ corresponds to the approximation of the discrete trace of the E-IP method, and u_h^{CG} is the solution provided by the Continuous

Galerkin method. Then, only for $k = 1$ and $\varepsilon = 1$, the equivalence $\hat{u}_h^{\mathrm{E}} = u_h^{\mathrm{CG}}$ on $\mathcal{F}_h$ is verified for all data $\bar{u}$ and f, independently from the penalization function $\tau_{E,F}$ and the physical properties of the media $\boldsymbol{\alpha}$ (see Appendix in [Cockburn 2009c] for more details). The equivalence is not valid anymore for both non-symmetric variants E-NIP and E-IIP, as listed in Table 3.1.

Degree	E-NIP	E-IIP	E-SIP
$k = 1$	✗	✗	✓
$k \geq 2$	✗	✗	✗

TABLE 3.1 – Summary of the equivalence between the E-IP and CG methods for various polynomial degrees k and different values of $\varepsilon = \{0, \pm 1\}$.

Since both the Hybridized- and the Embedded-IP methods are described identically through the definition (3.6), these schemes comprise the same benefit by applying the well-known static condensation technique [Cockburn 2009a, Cockburn 2012c]. This procedure was introduced for reducing the size of the global problem (3.6) to the only trace unknowns. Let us denote by U_h and $\hat{\mathrm{U}}_h$ the vector of DOFs associated to the element- and trace-based unknowns, respectively. Thus, the following global matrix system is obtained from the linear system (3.6) :

$$\begin{bmatrix} \mathcal{B}_{uu} & \mathcal{B}_{u\hat{u}} \\ \mathcal{B}_{\hat{u}u} & \mathcal{B}_{\hat{u}\hat{u}} \end{bmatrix} \cdot \begin{bmatrix} \mathrm{U}_h \\ \hat{\mathrm{U}}_h \end{bmatrix} = \begin{bmatrix} F_u \\ G_{\hat{u}} \end{bmatrix}, \tag{3.8}$$

where the vectors F_u and $G_{\hat{u}}$ are directly related to the source term and boundary conditions, respectively. Owing to the discontinuous nature of the space $\mathcal{V}_h$, the matrix $\mathcal{B}_{uu}$ have a block-diagonal structure that allows its inversion. We are now able to substitute the global coupled matrix system (3.8) in terms of trace unknowns $\hat{\mathrm{U}}_h$, in a such way that :

$$[\mathcal{B}_{\hat{u}\hat{u}} - \mathcal{B}_{\hat{u}u}\mathcal{B}_{uu}^{-1}\mathcal{B}_{u\hat{u}}]\hat{\mathrm{U}}_h = G_{\hat{u}} - \mathcal{B}_{\hat{u}u}\mathcal{B}_{uu}^{-1}F_u. \tag{3.9}$$

We notice that the left-hand side of the reduced linear system (3.9) is called the Shur complement of $\mathcal{B}_{uu}$. Finally, the last step consists in the reconstruction of the discrete variable U_h element-by-element with the use of the relation :

$$\mathrm{U}_h = \mathcal{B}_{uu}^{-1}(F_u - \mathcal{B}_{u\hat{u}}\hat{\mathrm{U}}_h).$$

The reduced form (3.9) leads to a smaller and symmetric matrix system easing the resolution.

3.3.2 Weighted-IP formulation

The Weighted Interior Penalty (W-IP) method were firstly introduced by [Burman 2006] in the context of advection-diffusion-reaction problems. The principal difference with the classical SIP discretization proposed by [Arnold 1982] resides in the use of a weighted average operators. In the W-IP formulation proposed in [Dryja 2003, Burman 2006, Di Pietro 2011], the weighting function only accounts for the variations of the diffusivity tensor $\boldsymbol{\alpha}$ from both sides of the mesh interfaces. In the following, we will favor a modified version of the weighting

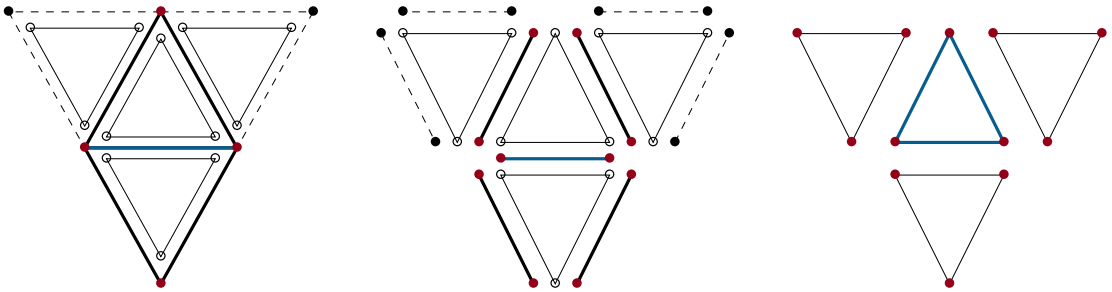


FIGURE 3.1 – Stencil for a generic element (triangle or edge in blue) for the Embedded- (left), Hybridized- (middle), and Weighted-IP (right) discretization with polynomial degree $k = 1$. For all methods, the nodes corresponding to degrees of freedom for the generic element are marked as solid red dots.

function depending on the penalization function of the Hybridized-IP methods as stated in equation (3.7). Thus, the discretization of the W-IP method consists to seek $u_h \in \mathcal{V}_h$ such that

$$\mathcal{B}^{(\varepsilon)}_{h,\mathrm{W}}(u_h, v_h) = \mathcal{L}_h(v_h) \quad \text{for all } v_h \in \mathcal{V}_h, \tag{3.10}$$

where the W-IP bilinear form $\mathcal{B}^{(\varepsilon)}_{h,\mathrm{W}}$ is given by :

$$\begin{aligned}\mathcal{B}^{(\varepsilon)}_{h,\mathrm{W}}(u_h, v_h) :=& (\boldsymbol{\alpha}\boldsymbol{\nabla}_h u_h, \boldsymbol{\nabla}_h v_h)_{0,\mathcal{T}_h} - \langle \{\boldsymbol{\alpha}\boldsymbol{\nabla}_h u_h\}_{\overline{\boldsymbol{\omega}}}, [\![v_h]\!]\rangle_{0,\mathcal{F}_h^{\mathrm{i}}\cup\mathcal{F}_h^D} \\ &- \varepsilon\langle \{\boldsymbol{\alpha}\boldsymbol{\nabla}_h v_h\}_{\overline{\boldsymbol{\omega}}}, [\![u_h]\!]\rangle_{0,\mathcal{F}_h^{\mathrm{i}}\cup\mathcal{F}_h^D} + \langle \eta [\![u_h]\!], [\![v_h]\!]\rangle_{0,\mathcal{F}_h^{\mathrm{i}}\cup\mathcal{F}_h^D},\end{aligned}$$

and the linear form

$$\mathcal{L}_h(v_h) := (f, v)_{0,\mathcal{T}_h} + \langle \eta\bar{u}, v_h\rangle_{0,\mathcal{F}_h^D} - \varepsilon\langle \bar{u}\boldsymbol{n}, \boldsymbol{\alpha}\boldsymbol{\nabla}_h v_h\rangle_{0,\mathcal{F}_h^D} + \langle \bar{\sigma}, v_h\rangle_{0,\mathcal{F}_h^N},$$

contains the weakly imposed Dirichlet and Neumann boundary data. In Figure 3.1c, we depict the stencil of the W-IP method, which is similar to the standard IP method. As in section 3.3.1, the parameter ε is used to control the introduction of the consistent symmetry term. Following its values, these methods are referred to as the Symmetric ($\varepsilon = 1$), the Non-Symmetric ($\varepsilon = -1$), and the Incomplete ($\varepsilon = 0$) schemes and are denoted W-SIP, W-NIP, and W-IIP, respectively. Let us now introduce the corresponding weighting function taking into account the values of the penalization from both sides of the interface F as :

$$\boldsymbol{\omega} := \left(\frac{\tau_1}{\tau_1+\tau_2}, \frac{\tau_2}{\tau_1+\tau_2}\right) \quad \forall\, F \in \mathcal{F}_h^{\mathrm{i}}, \tag{3.11}$$

where the stabilization function $\tau_i = \tau_{E_i,F}$, for $i = 1, 2$, is defined by equation (3.7). For all $E \in \mathcal{T}_h$ and $F \in \partial E$, we give the penalization function η_F as follows :

$$\eta_F := \begin{cases} \dfrac{\tau_1\tau_2}{\tau_1+\tau_2} & \forall\, F \in \mathcal{F}_h^{\mathrm{i}}, \\ \tau_{E,F} & \forall\, F \in \mathcal{F}_h^{\mathfrak{b}}. \end{cases} \tag{3.12}$$

Remark 3.2 (Equivalence between the Hybridized- and Weighted-IP schemes). The numerical experiments performed in section 3.4 provided evidences for the total equivalence between

the W-IIP and the H-IIP schemes, and for any values of $\tau_{E,F}$ satisfying the weights (3.11) and the penalization function (3.12). Consequently, the Incomplete variations of the W-IP and H-IP methods conduct to the same robustness, accuracy, and stability. We would mention that these observations will be reported in a forthcoming work, where the total equivalence property will be mathematically demonstrated. Similar results can be found in [Fabien 2020b] for the homogeneous case, where the H-IIP scheme coincides with the IIP scheme (see Remark 3.3 for the characteristics of the corresponding IP scheme). However, this equivalence is not valid anymore for both Symmetric and Non-Symmetric IP variants.

Remark 3.3 (Relations with the standard IP method). Assuming now $\tau_1 = \tau_2$ for any interior interfaces $F \in \mathcal{F}_h^{\mathrm{i}}$, we then observe that $\boldsymbol{\omega} = (1/2, 1/2)$ and the penalty function is simplified as $\eta_F = \tau_{E,F}/2$. In this case, the W-IP method is reduced to the standard IP method, where the well-known SIP scheme was analyzed by [Arnold 1982] and the IP scheme in [Rivière 2008]. From the definition (3.7) of the stabilization function, the W-IP formulation is simplified to the standard IP method, if and only if the media is homogeneous and the partition $\mathcal{T}_h$ is uniform.

3.3.3 Computational cost

The purpose of this section is to illustrate the computational cost of the Hybridized- and Embedded-IP compared to the Weighted-IP method. By applying the static condensation procedure, the total number of degrees of freedom (DOFs) is reduced to the trace unknowns for both the H-IP and E-IP discretization. It appears that this quantity is proportional to the number of interfaces of the skeleton. For the WIP method, the global number of degrees of freedom is proportional to the number of elements of the partition $\mathcal{T}_h$. In order to ease the description, we consider a regular Cartesian domain $\Omega = [0, 1]^2$ partitioned into $N \times N$ squares. We distinguish two cases :

- **Case 1** : The triangulation consists of a uniform rectangular grid made of N^2 elements ;
- **Case 2** The triangulation consists of a uniform triangular grid made of $2N^2$ elements, which is generated by divided each square into two triangles.

Now, we are able to compute the number of DOFs related to the methods under consideration. Let R_{H} and R_{E} be the ratio of DOFs of the H-IP and E-IP method versus the W-IP method such that,

$$R_{\mathrm{H}} = \frac{\mathrm{DOF}_{\text{H-IP}}}{\mathrm{DOF}_{\text{W-IP}}} = \frac{N_e N_{\mathrm{dof}/e}}{N_E N_{\mathrm{dof}/E}}, \quad \text{and} \quad R_{\mathrm{E}} = \frac{\mathrm{DOF}_{\text{E-IP}}}{\mathrm{DOF}_{\text{W-IP}}} = \frac{N_v + N_e(N_{\mathrm{dof}/e} - 2)}{N_E N_{\mathrm{dof}/E}}, \tag{3.13}$$

where N_v, N_e and N_E refers to the number of vertices, edges and elements (triangular or quadrilateral), respectively. The number of DOFs associated to the local reference element (interface or polygonal) can be defined by

$$N_{\mathrm{dof}/e} = (k+1), \quad N_{\mathrm{dof/trigs}} = \frac{1}{2}(k+1)(k+2), \quad \text{and} \quad N_{\mathrm{dof/quads}} = (k+1)^2. \tag{3.14}$$

In Figure 3.2, we display the computed ratios R_{H} and R_{E} for various polynomial degrees k and element sizes. Figures 3.2a and 3.2c show the decrease of these two ratios for each values

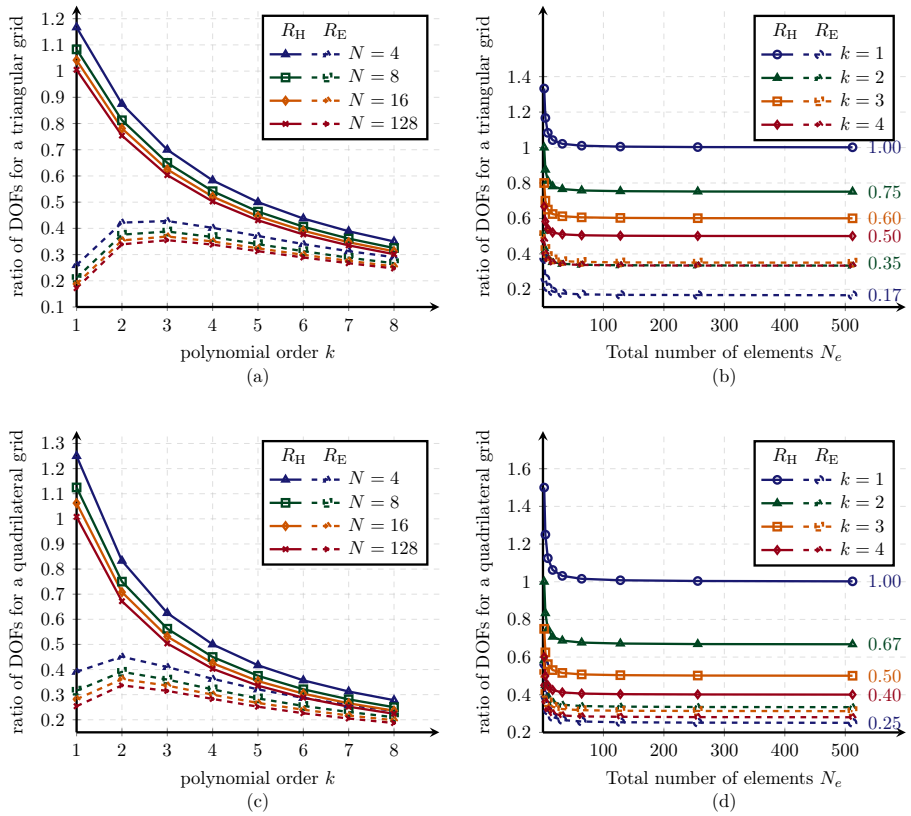


FIGURE 3.2 – Total degrees of freedom of the Weighted-IP compared to Hybridized- and Embedded-IP methods for both uniform triangular (top) and rectangular (bottom) grids.

of N and for any given polynomial degrees $k \geq 2$. These results are considerably stressed by Figures 3.2b and 3.2d, which emphasize the reduction of both ratios with the increase of k. Globally, the W-IP method contains a larger number of degrees of freedom compared to the H-IP and E-IP schemes, and for any values of $k \geq 2$. We note that for $k \geq 3$ and $k \geq 4$, the number of degrees of freedom of the H-IP method is reduced by half compared to those of the W-IP method for the rectangular and triangular meshes, respectively. For a given polynomial degree $1 \leq k \leq 8$, the calculated ratio R_E always promotes the E-IP method that contains fewer degrees of freedom than the W-IP method, as well as the H-IP method. This is particularly true for the first polynomial degrees smaller than $k = 4$. However, for any polynomial degree k sufficiently large, both ratios R_H and R_E appear to converge on the same value making both H-IP and E-IP methods identical in terms of degrees of freedom.

Remark 3.4 (Global DOFs of the HDG method). Without the static condensation technique, the total number of DOFs of the HDG method is significantly higher than the classical DG scheme. Indeed, the global linear system comprises the unknowns of the state variable increased by those of the discrete trace.

Remark 3.5 (CPU time). The size of the global matrix system is strongly influenced by the number of elements, inter-element connectivities, and polynomial degree k. These parameters

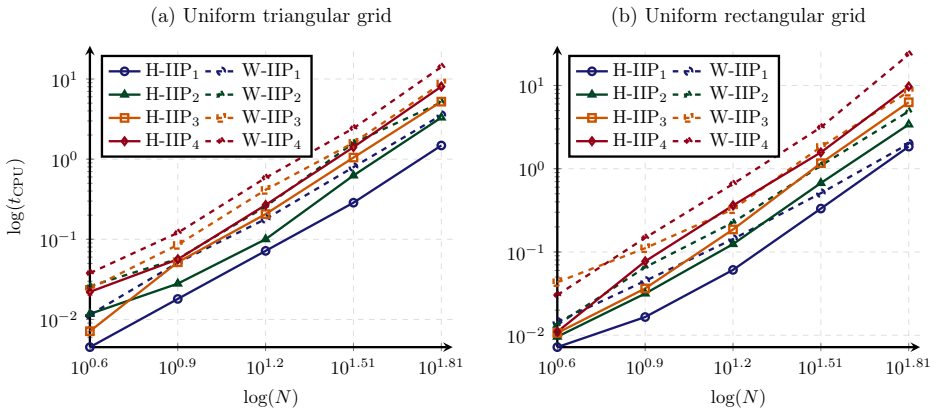


FIGURE 3.3 – Computed CPU time (s) for the H-IIP and W-IIP methods with a triangular (a) and rectangular (b) uniform meshes for different polynomial degree $k = \{1, 2, 3, 4\}$.

play a fundamental role in the running time of the simulations. To understand this purpose, we give an illustration of the computed CPU time for both H-IIP and W-IIP schemes in Figure 3.3. By choosing the Incomplete variants, the dependence on the error estimate is disabled since the Hybridized- and Weighted-IP methods are fully equivalent (see Remark 3.2). Due to the number of degrees of freedom involved in the resolution, we observe that the H-IP method is always faster than its W-IP counterpart.

3.4 Numerical results and discussion

This section provided several numerical experiments, which aims to investigate the performances of the E-IP, H-IP and W-IP methods in terms of stability, accuracy and efficiency. The ability of the IP schemes to handle heterogeneity and anisotropy is also evaluated. For this reason, we measured the CPU time that requires the approximation of the discrete variable u_h, and computed the error following the L^2-norm estimates $\|u - u_h\|_{0,\mathcal{T}_h}$. The numerical experiments are performed using the high-performance multiphysics finite element library called NGSolve [Schöberl 2014]. Then, we consider the homogeneous Dirichlet boundary problem in the unit square $\Omega = [0, 1]^2$, where the exact solution is given by $u(x, y) = \sin(\pi x)\sin(\pi y)$ and the right hand side is chosen such that the exact solution is verified. Thereafter, the domain Ω is split into four subdomains $\Omega_1 = [0, 1/2]^2$, $\Omega_2 = [1/2, 1] \times [0, 1/2]$, $\Omega_3 = [1/2, 1]^2$, and $\Omega_4 = [0, 1/2] \times [1/2, 1]$, i.e., $\Omega := \cup_{i=1}^4 \Omega_i$ (see Figure 3.4a), and the diffusivity tensor is defined separately for each subregions :

$$\boldsymbol{\alpha} = \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix} \quad \text{for } (x, y) \in \Omega_1, \Omega_3, \qquad \text{and} \qquad \boldsymbol{\alpha} = \begin{bmatrix} 1/\lambda & 0 \\ 0 & 1 \end{bmatrix} \quad \text{for } (x, y) \in \Omega_2, \Omega_4.$$

Here λ represents the strongest anisotropy ratio, which controls both the heterogeneity and anisotropy of the media. In the following, we consider two families of structured meshes (triangular/rectangular) respecting the discontinuity of $\boldsymbol{\alpha}$, as illustrated in Figure 3.4. For instance, the rectangular grid is achieved by discretizing the unit domain into $N \times N$ uni-

form quadrilaterals, with length $h = 1/N$. The triangular grid is obtained by divided each quadrilateral of the regular rectangular meshes into two triangles. Notice that, the parameter N refers to the employed mesh refinement. Moreover, we established some discussions (i) to summarize the main results that arise immediately from the numerical experiments, and (ii) to identify the most appropriate schemes according to the characteristics of the problem.

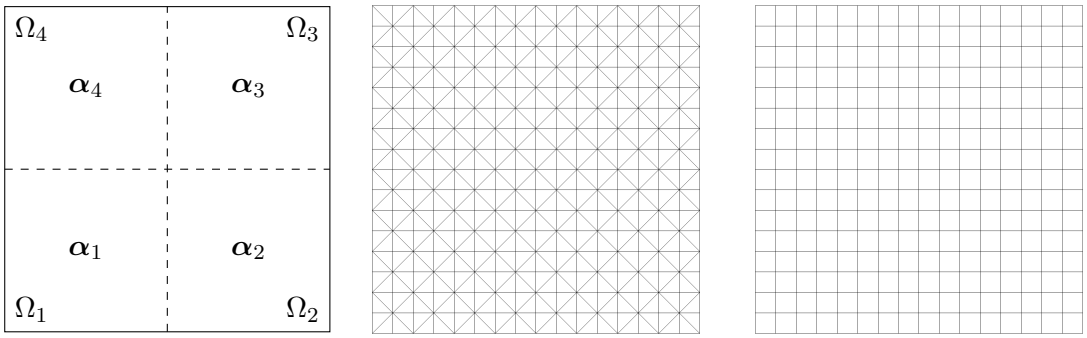


FIGURE 3.4 – Description of the test case with genuine anisotropy and heterogeneity properties with the four subdomains (a). Uniform triangular (b) and rectangular (c) meshes with $h = 1/16$.

3.4.1 Homogeneous isotropic flow

In the first experiment, we evaluate the performances of the Embedded-, Hybridized- and Weighted-Interior Penalty methods for the simplest model problem, i.e., the Poisson's equation. In this context, the material is supposed to be homogeneous and isotropic $\boldsymbol{\alpha} = \boldsymbol{I}_2$ on the whole domain ($\lambda = 1$), where $\boldsymbol{I}_2$ refers to the 2×2 identity tensor. Since the solution is smooth enough, we begin with the analysis of the convergence of the E-IP, H-IP and W-IP discretizations for its respective variants : Non-Symmetric ($\varepsilon = -1$), Incomplete ($\varepsilon = 0$) and Symmetric ($\varepsilon = 1$). To this aim, we computed all methods for both triangular and rectangular regular meshes with $N = \{4, 8, 16, 32, 64\}$ and various polynomial degrees $k = \{2, 3\}$. In this context, the W-IP scheme is reduced to the standard IP method for both grid representations (see Remark 3.2), and the penalization function is supposed to be constant for each interfaces of the skeleton : $\tau_{E,F} = \tau_0(k+1)(k+2)/h$, $\forall\, E \in \mathcal{T}_h$, and $\forall\, F \in \partial E$, with the positive constant $\tau_0 = 2$. In Figure 3.5, we plot the history of convergence of the L^2-error for all variations of the IP methods as a function of the CPU time. Through these results, we observe the benefits of the static condensation procedure with the Hybridized- and Embedded-IP discretizations in terms of CPU time. This can be easily explained by the values of the ratios R_{H} and R_{E} (see Figure 3.2) that are globally smaller than one for each polynomial degrees $k \geq 2$. Owing to the restriction of the discrete trace to the set of continuous functions, the E-IP method is always faster than the H-IP and W-IP methods for any given refined meshes (triangular or rectangular). Additionally, we observe the similar behavior of the Embedded-, Hybridized- and Weighted-IP schemes, and we recover some well-known estimates from the standard IP method. On one side, the L^2-norm estimate confirm its dependence to the parity of the

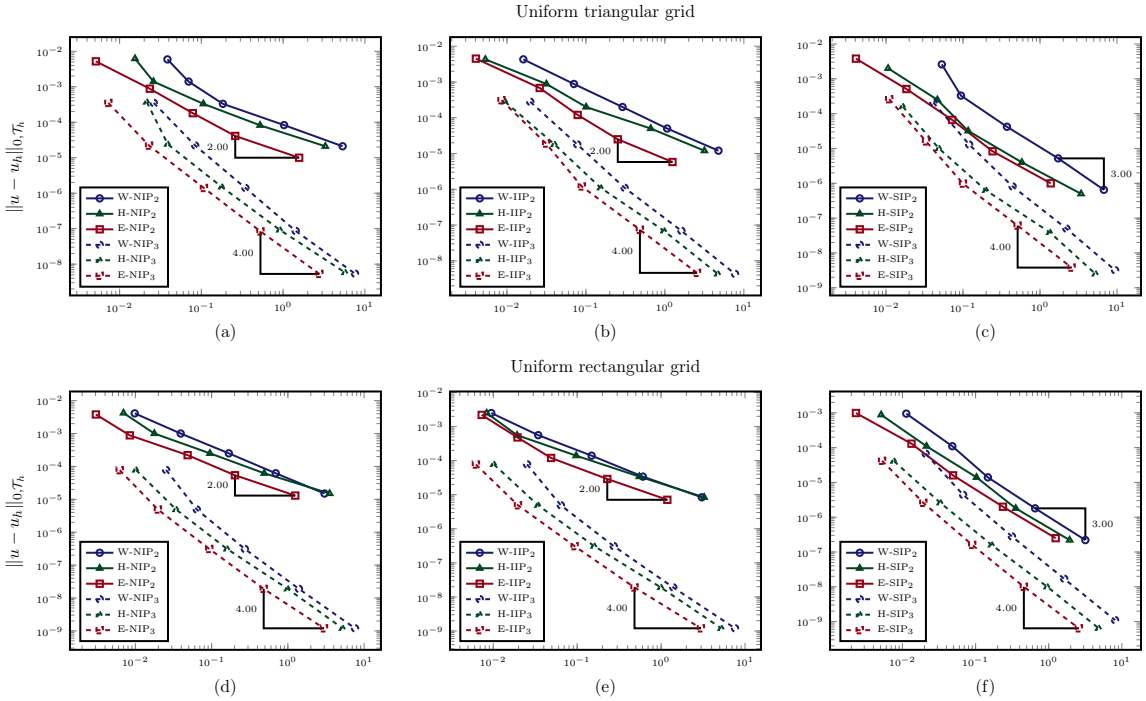


FIGURE 3.5 – Homogeneous case - History of convergence in the $\|\cdot\|_{0,\mathcal{T}_h}$-norm (vs. t_{CPU}) of the W-IP, H-IP and E-IP methods for the Non-Symmetric (a-d), the Incomplete (b-e) and the Symmetric variant (c-f), respectively. The grid is triangular in the top (a-b-c) and rectangular in the bottom (d-e-f).

polynomial degree k. Indeed, the Symmetric variant of the E-IP, H-IP and W-IP methods converge optimally with order $k+1$ for each values of k, while this optimal convergence rate is achieved for both non-symmetric variants ($\varepsilon = \{-1, 0\}$) and for each even degree k. As a contrast, there is a loss of accuracy of the non-symmetric variants with odd degree k, where the convergence rate decreases to k. We point out the advantage of the stabilization function as stated earlier, which ensures the preservation of the accuracy and efficiency of the E-IP method that has the same error magnitude as its counterpart Hybridized. We should

		Uniform triangular meshes						**Uniform rectangular meshes**					
k	N	W-IIP		H-IIP		E-IIP		W-IIP		H-IIP		I-EIP	
		$\|u-u_h\|_{0,\mathcal{T}_h}$	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	$\|u-u_h\|_{0,\mathcal{T}_h}$	ECR	$\|u-u_h\|_{0,\mathcal{T}_h}$	ECR
2	4	$4.3e-03$	–	$4.3e-03$	–	$4.5e-03$	–	$2.5e-03$	–	$2.5e-03$	–	$2.2e-03$	–
	8	$8.8e-04$	2.31	$8.8e-04$	2.31	$6.8e-04$	2.72	$5.6e-04$	2.15	$5.6e-04$	2.15	$4.8e-04$	2.21
	16	$2.0e-04$	2.11	$2.0e-04$	2.11	$1.2e-04$	2.54	$1.4e-04$	2.04	$1.4e-04$	2.04	$1.2e-04$	2.07
	32	$5.0e-05$	2.03	$5.0e-05$	2.03	$2.5e-05$	2.25	$3.4e-05$	2.01	$3.4e-05$	2.01	$2.9e-05$	2.02
	64	$1.2e-05$	2.01	$1.2e-05$	2.01	$5.8e-06$	2.08	$8.4e-06$	2.00	$8.4e-06$	2.00	$7.1e-06$	2.00
3	4	$2.8e-04$		$2.8e-04$	–	$3.1e-04$	–	$7.8e-05$	–	$7.8e-05$	–	$7.7e-05$	–
	8	$1.8e-05$	3.99	$1.8e-05$	3.99	$3.1e-04$	3.56	$4.9e-06$	3.98	$4.9e-06$	3.98	$4.9e-06$	3.98
	16	$1.1e-06$	4.00	$1.1e-06$	4.00	$1.2e-06$	4.01	$3.1e-07$	4.00	$3.1e-07$	4.00	$3.1e-07$	4.00
	32	$6.9e-08$	4.00	$6.9e-08$	4.00	$7.4e-08$	4.00	$1.9e-08$	4.00	$1.9e-08$	4.00	$1.9e-08$	4.00
	64	$4.3e-09$	4.00	$4.3e-09$	4.00	$4.6e-09$	4.00	$1.2e-09$	4.00	$1.2e-09$	4.00	$1.2e-09$	4.00

TABLE 3.2 – Homogeneous case - History of convergence $\|u-u_h\|_{0,\mathcal{T}_h}$ of the W-IIP, H-IIP and E-IIP methods on uniform triangular and square meshes for $k = \{2, 3\}$.

recognize the main benefit of the Embedded-IP scheme, which is more accurate than both H-IP and W-IP schemes for any fixed number of DOFs. On the other hand, we retrieve here

the total equivalence between the W-IIP and H-IIP scheme as stated in section 3.3.2. In order to support this property, we listed in Table 3.2 the L^2-norm error estimation for all W-IIP, H-IIP and E-IIP schemes. Both W-IIP and H-IIP methods coincide, while this assumption is no longer verified for the E-IIP method. Due to the reduction of the W-IP to the standard IP scheme, we highlight the equivalence between the H-IIP and standard IIP methods that appears instantaneously.

3.4.2 Heterogeneous anisotropic flow

For the second test case, we analyze the capability of the Embedded-, Hybridized-, and Weighted-Interior Penalty methods to capture strong heterogeneity and anisotropy. For illustrating the benefits of the penalization function (3.7), we first focus on the Symmetric variant of the E-IP, H-IP and H-IP methods. For the simulation, we consider both structured triangular and rectangular meshes with $N = \{4, 8, 16, 32, 64\}$. We summarized in Figure 3.6 the history of convergence of all schemes for various polynomial degrees $k = \{2, 3\}$ and different values of $\lambda = \{10^1, 10^3, 10^6\}$. Despite the presence of heterogeneity and anisotropy, all conclusions established for the homogeneous test case (section 3.4.1) are verified. In most situations,

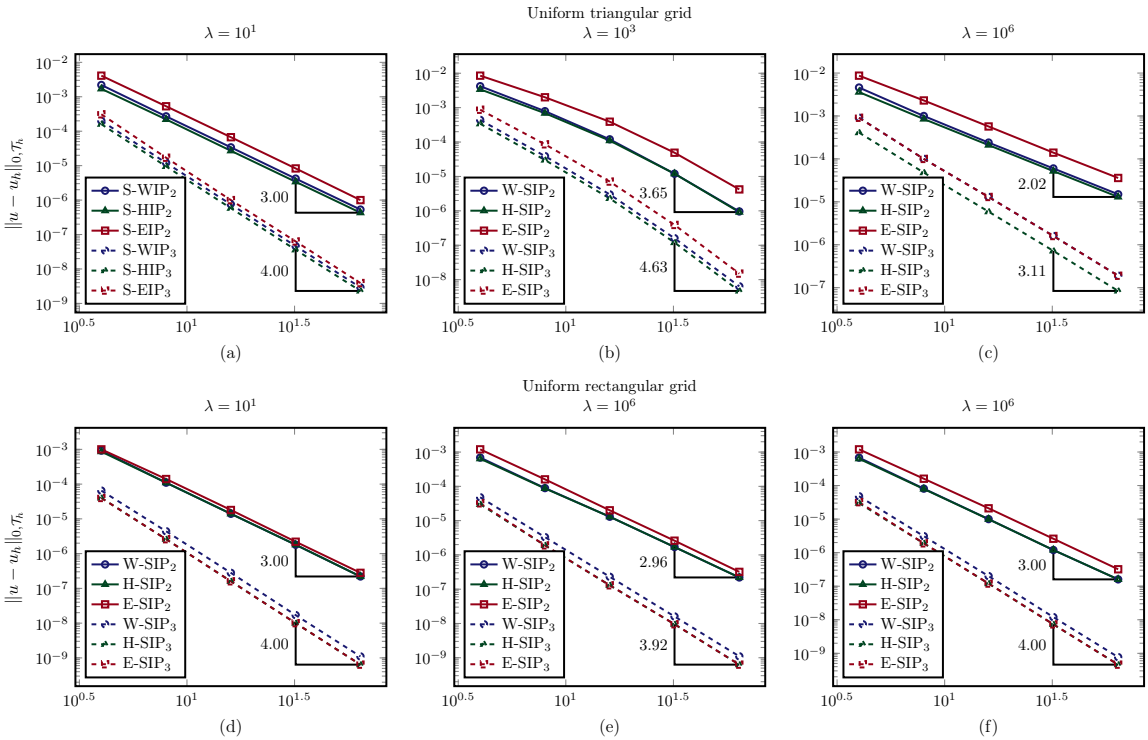


FIGURE 3.6 – Heterogeneous and anisotropic case - History of convergence in the $\|\cdot\|_{0,\mathcal{T}_h}$-norm (vs. $1/h$) of the W-SIP, H-SIP and E-SIP methods for different values of the diffusivity parameter $\lambda = \{10^1, 10^3, 10^6\}$, respectively. The grid is triangular in the top and rectangular in the bottom.

the Embedded-, Hybridized- and Weighted-IP methods converge optimally with order $k + 1$ for each polynomial degrees k and for any values of λ. Particularly, the E-SIP, H-SIP and

W-SIP schemes are more sensitive with the triangular grid, which affects the convergence rate of the scheme for high values of λ. For $\lambda = 10^6$, all Symmetric methods are sub-optimal with order k for each polynomial degrees k. We place emphasis on the normal diffusivity parameter $\alpha_{E,F}$, which plays a fundamental role in the robustness of the respective schemes. It appears that without the use of $\alpha_{E,F}$ the discrete solution may exhibit spurious oscillations and the performances of the methods should decrease. Afterwards, we carry out an analysis of the E-IIP, H-IIP and W-IIP methods. In Table 3.3, we listed a history of convergence of the Incomplete variation of the three IP schemes. Even if the results of the Non-Symmetric variant are not presented, we must notice the behavior of the E-NIP, H-NIP and W-NIP that are quite similar to those established in section 3.4.1. Here again, we wish to outline the total equivalence between the H-IIP and W-IIP that is always verified in presence of heterogeneity and anisotropy, and for any values of λ. These results are in agreement with the homogeneous case, where the convergence rate depending on the parity of the polynomial degree k.

		Uniform triangular meshes						**Uniform rectangular meshes**					
k	N	W-IIP		H-IIP		E-IIP		W-IIP		H-IIP		E-IIP	
		$\|u-u_h\|_{0,\mathcal{T}_h}$	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	$\|u-u_h\|_{0,\mathcal{T}_h}$	ECR	$\|u-u_h\|_{0,\mathcal{T}_h}$	ECR
$\lambda = 10^1$													
2	4	$4.2e-03$	–	$4.2e-03$	–	$4.4e-03$	–	$2.5e-03$	–	$2.5e-03$	–	$2.4e-03$	–
	8	$8.8e-04$	2.25	$8.8e-04$	2.25	$6.4e-04$	2.77	$5.6e-04$	2.14	$5.6e-04$	2.14	$5.3e-04$	2.19
	16	$2.1e-04$	2.07	$2.1e-04$	2.07	$1.0e-04$	2.62	$1.4e-04$	2.04	$1.4e-04$	2.04	$1.3e-04$	2.06
	32	$5.2e-05$	2.02	$5.2e-05$	2.02	$2.1e-05$	2.33	$3.4e-05$	2.01	$3.4e-05$	2.01	$3.1e-05$	2.02
	64	$1.3e-05$	2.00	$1.3e-05$	2.00	$4.8e-06$	2.11	$8.4e-06$	2.00	$8.4e-06$	2.00	$7.8e-06$	2.00
3	4	$2.8e-04$	–	$2.8e-04$	–	$3.3e-04$	–	$7.6e-05$	–	$7.6e-05$	–	$7.5e-05$	–
	8	$1.7e-05$	4.06	$1.7e-05$	4.06	$1.9e-05$	4.13	$4.9e-06$	3.95	$4.9e-06$	3.95	$4.9e-06$	3.95
	16	$1.0e-06$	4.02	$1.0e-06$	4.02	$1.1e-06$	4.06	$3.1e-07$	3.99	$3.1e-07$	3.99	$3.1e-07$	3.99
	32	$6.4e-08$	4.01	$6.4e-08$	4.01	$7.1e-08$	4.02	$1.9e-08$	4.00	$1.9e-08$	4.00	$1.9e-08$	4.00
	64	$4.0e-09$	4.00	$4.0e-09$	4.00	$4.4e-09$	4.01	$1.2e-09$	4.00	$1.2e-09$	4.00	$1.2e-09$	4.00
$\lambda = 10^3$													
2	4	$6.7e-03$	–	$6.7e-03$	–	$8.4e-03$	–	$2.4e-03$	–	$2.4e-03$	–	$2.6e-03$	–
	8	$1.5e-03$	2.19	$1.5e-03$	2.19	$1.9e-03$	2.11	$5.5e-04$	2.10	$5.5e-04$	2.10	$5.7e-04$	2.17
	16	$2.8e-04$	2.39	$2.8e-04$	2.39	$3.6e-04$	2.45	$1.4e-04$	2.03	$1.4e-04$	2.03	$1.4e-04$	2.06
	32	$5.5e-05$	2.36	$5.5e-05$	2.36	$4.6e-05$	2.94	$3.4e-05$	2.01	$3.4e-05$	2.01	$3.4e-05$	2.02
	64	$1.3e-05$	2.06	$1.3e-05$	2.06	$6.5e-06$	2.83	$8.4e-06$	2.00	$8.4e-06$	2.00	$8.4e-06$	2.00
3	4	$6.5e-04$	–	$6.5e-04$	–	$8.6e-04$	–	$6.1e-05$	–	$6.1e-05$	–	$6.2e-05$	–
	8	$6.4e-05$	3.34	$6.4e-05$	3.34	$8.4e-05$	3.35	$3.9e-06$	3.96	$3.9e-06$	3.96	$3.9e-06$	3.97
	16	$5.2e-06$	3.63	$5.2e-06$	3.63	$6.9e-06$	3.61	$2.7e-07$	3.87	$2.7e-07$	3.87	$2.7e-07$	3.88
	32	$2.7e-07$	4.24	$2.7e-07$	4.24	$3.8e-07$	4.20	$1.8e-08$	3.87	$1.8e-08$	3.87	$1.8e-08$	3.87
	64	$1.1e-08$	4.66	$1.1e-08$	4.66	$1.5e-08$	4.65	$1.2e-09$	3.94	$1.2e-09$	3.94	$1.2e-09$	3.95
$\lambda = 10^6$													
2	4	$7.1e-03$	–	$7.1e-03$	–	$8.7e-03$	–	$2.4e-03$	–	$2.4e-03$	–	$2.6e-03$	–
	8	$1.7e-03$	2.05	$1.7e-03$	2.05	$2.2e-03$	1.96	$5.5e-04$	2.10	$5.5e-04$	2.10	$5.7e-04$	2.17
	16	$4.3e-04$	2.01	$4.3e-04$	2.01	$5.6e-04$	1.99	$1.3e-04$	2.03	$1.3e-04$	2.03	$1.4e-04$	2.06
	32	$1.1e-04$	2.01	$1.1e-04$	2.01	$1.4e-04$	2.00	$3.4e-05$	2.01	$3.4e-05$	2.01	$3.4e-05$	2.02
	64	$2.6e-05$	2.01	$2.6e-05$	2.01	$3.5e-05$	2.01	$8.4e-06$	2.00	$8.4e-06$	2.00	$8.4e-06$	2.00
3	4	$7.3e-04$	–	$7.3e-04$	–	$9.1e-04$	–	$6.1e-05$	–	$6.1e-05$	–	$6.2e-05$	–
	8	$8.8e-05$	3.06	$8.8e-05$	3.06	$1.0e-04$	3.17	$3.8e-06$	3.99	$3.8e-06$	3.99	$3.8e-06$	4.01
	16	$1.1e-05$	3.02	$1.1e-05$	3.02	$1.2e-05$	3.05	$2.4e-07$	4.00	$2.4e-07$	4.00	$2.4e-07$	4.00
	32	$1.3e-06$	3.02	$1.3e-06$	3.02	$1.5e-06$	3.02	$1.5e-08$	4.00	$1.5e-08$	4.00	$1.5e-08$	4.00
	64	$1.6e-07$	3.07	$1.6e-07$	3.07	$1.9e-07$	3.02	$9.3e-10$	4.00	$9.3e-10$	4.00	$9.3e-10$	4.00

TABLE 3.3 – Heterogeneous and anisotropic case - History of convergence $\|u-u_h\|_{0,\mathcal{T}_h}$ of the W-IIP, H-IIP and E-IIP methods on uniform triangular and square meshes for $k = \{2, 3\}$ and various diffusivity parameter $\lambda = \{10^1, 10^3, 10^6\}$.

3.4.3 Discussion

The numerical results reported in sections 3.4.1 and 3.4.2 reveal a series of similarities, which connects the three Interior Penalty methods (i.e., the Hybridized-, Embedded- and Weighted-IP). First, for both test cases, all IP schemes converge optimally for (i)

the Symmetric variant and (*ii*) the Incomplete and Non-Symmetric variants, according to the parity of the polynomial degree k. Moreover, it has been proven that by forcing the continuity of the discrete trace, the E-IP scheme can suffer from a loss of accuracy (see [Cockburn 2009c, Fabien 2020b] for more details). However, thanks to the specific definition of the stabilization function (3.7), we obtain a very high level of error estimates as well as for the Embedded-IP method. Indeed, this enriched penalization function provides substantial benefits for the Hybridized-, Embedded- and Weighted-IP schemes (i.e., strong robustness, accuracy, and efficiency) that restrict the different results in a very closed area. Especially, these properties are still valid for highly heterogeneous and/or anisotropic variations of the medium. Secondly, we highlight the total equivalence property between the W-IIP and H-IIP methods that bridges the gap between these two approaches. This characteristic is valid whether the medium is homogeneous or heterogeneous and/or anisotropic. To the best of our knowledge, this is the first attempt to establish a Weighted-IP scheme as efficient as a Hybridized-IP scheme for highly perturbated materials properties. Since their introduction, the HDG methods have demonstrated their superiority over the traditionally DG frameworks in terms of efficiency, accuracy, and robustness [Cockburn 2009b]. Nevertheless, the Weighted-IP scheme, described in this manuscript, exhibits the ability to recover some performances comparable to those of traces formalism (i.e., Hybridized- and/or Embedded) at most. Finally, there is a significant contrast between all IP approaches, which is undeniable when focusing on CPU time. Owing to the static condensation, we measured running times, which are almost reduced by half and quarter for the H-IP and E-IP methods, respectively. This latest point enables us to recognize the leadership of the Hybridized- and Embedded-IP schemes for all numerical aspects. Even if the E-IP approach is faster than its H-IP counterpart, we prefer to outline the accuracy and the robustness of the H-IP schemes, which surpasses those of the E-IP and W-IP methods with heterogeneity and/or anisotropy.

3.5 Conclusion

In the present paper, we compare the computational performances of several variations of the IP methods, namely, the Hybridized-, Embedded- and Weighted-Interior Penalty schemes for solving heterogeneous and anisotropic diffusion problems. Specifically, they lead to a final matrix systems of different sizes and sparsities, which strongly impacts the computing time of each methods. We then quantify their total numbers of degrees of freedom for a given mesh and a fixed polynomial degree. This comparative analysis clearly indicated that both the E-IP and the H-IP methods led to smaller and sparser final matrix systems requiring less CPU time to compute. Due to the regularity requirement of the discrete trace approximation, the E-IP method is slightly more favorable since it generates fewer DOFs than its H-IP counterpart. Particularly, for the Symmetric E-IP variant, it produces more accurate results for the homogeneous case, with a trace approximation as robust as the Continuous Galerkin method. Moreover, we recovered some well-known error estimates in the L^2-norm of IP methods : for both non-symmetric variants obtained by selecting $\varepsilon = 0$ or -1, the estimated convergence rates are influenced by the parity of the polynomial degree k (sub-optimally for even k and optimally for odd k). The situation is quite different for Symmetric variants ($\varepsilon = 1$) since they

always converge optimally. Nevertheless, we must recognize the robustness of the Hybridized- and Embedded-IP methods, which exceeds that of the Weighted-IP schemes independently of the mesh, the polynomial degree, and physical properties (homogeneity, heterogeneity and/or anisotropy). To conclude, let us emphasize on the numerical equivalence between the Incomplete H-IP and W-IP methods, which is achieved by applying a specific definition of the weighting function and the penalty parameter. In a forthcoming paper, we will discussed and theoretically proved this property for the Incomplete, Non-Symmetric, and Symmetric variants to recover a Weighted-IP scheme as accurate as the Hybridized-IP method.

Chapitre 4

Families of hybridized interior penalty methods for locally degenerate problems

Article in evaluation in the Journal of *Applied Mathematics and Computation* under the title Families of hybridized interior penalty discontinuous Galerkin methods for locally degenerate advection-diffusion-reaction problems
arXiv:2106.00226

Grégory Etangsale [1], Marwan Fahs [2], Vincent Fontaine[1], Nalitiana Rajaonison[1]

Abstract

We analyze families of primal high-order hybridized discontinuous Galerkin (HDG) methods for solving degenerate second-order elliptic problems. One major problem regarding this class of PDEs concerns its mathematical nature, which may be nonuniform over the whole domain. Due to the local degeneracy of the diffusion term, it can be purely hyperbolic in a subregion and elliptical in the rest. This problem is thus quite delicate to solve since the exact solution can be discontinuous at interfaces separating the elliptic and hyperbolic parts. The proposed hybridized interior penalty DG (H-IP) method is developed in a unified and compact fashion. It can handle pure-diffusive or -advective regimes as well as intermediate regimes that combine these two mechanisms for a wide range of Péclet numbers, including the tricky case of local evanescent diffusivity. To this end, an adaptive stabilization strategy based on the addition of jump-penalty terms is considered. An upwind-based scheme using a Lax–Friedrichs correction is favored for the hyperbolic region, and a Scharfetter–Gummel-based technique is preferred for the elliptic one. The well-posedness of the H-IP method is also briefly discussed by analyzing the strong consistency and the discrete coercivity condition in a self-adaptive energy-norm that is regime dependent. Extensive numerical experiments are performed to verify the model's robustness for all the abovementioned regimes.

Contents

1. Department of Building and Environmental Sciences, University of La Réunion - South Campus, France
2. Institut Terre et Environnement de Strasbourg, Université de Strasbourg, CNRS, ENGEES, UMR 7063, Strasbourg, France

4.1 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 [Oleinik 1973]. 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 [Di Pietro 2008, Ern 2009, Di Pietro 2015b]. This critical 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 framework of discontinuous Galerkin (DG) methods (see, e.g., [Gastaldi 1989, Houston 2002, Di Pietro 2008, Ern 2009] 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 [Raeisi Isa-Abadi 2020]. Despite all of these assets, DG methods are generally more expensive than most other numerical methods due to their high number of coupled degrees of freedom (DOFs). It is in this context that the hybridized DG methods were initially devised : as a way to circumvent these drawbacks.

The HDG methods were first introduced in [Cockburn 2009b] and have been applied successfully to various physical problems [Nguyen 2009, Lehrenfeld 2010, Nguyen 2011,

Cockburn 2012b, Oikawa 2015, Cockburn 2016a, Dijoux 2019b, Kirk 2019]. 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 [Kirby 2012]. 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 introduced in [Di Pietro 2015a] for the diffusive part to address this kind of issue [Di Pietro 2015b]. We would like to refer the readers to the seminal paper [Cockburn 2016b], where the authors recast the HHO method into the general HDG framework for the diffusion model problem, establishing hence bridges between both classes of discontinuous skeletal methods.

In the present paper, we focus instead on the class of hybridized interior penalty DG methods denoted by H-IP and its three well-known variants, namely, the Incomplete (H-IIP), Symmetric (H-SIP), and Non-Symmetric (H-NIP) schemes [Fabien 2020b, Wells 2011]. The starting point of our investigation is the H-IP formulation initially designed and analyzed in [Etangsale 2022] for purely diffusive problems, and extended here for degenerate elliptic equations. Indeed, the families of interior penalty methods are well-suited since the diffusion term might not be invertible at every point of the domain ; i.e., they belong to the class of $\boldsymbol{\alpha}$-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 automatically and appropriately in accordance with the mathematical structure of the local cellwise problem reducing to an upwinding-based scheme in the hyperbolic subregion [Brezzi 2004] and the Scharfetter–Gummel (SG) scheme elsewhere [Brezzi 2006, Di Pietro 2015b]. 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. In the latter case, our discrete formulation does not require any ad hoc techniques at the separative interface between elliptic/hyperbolic regions, which is a significant numerical improvement. A stability analysis is then investigated by establishing the (strong) consistency and the discrete coercivity in a self-adaptive energy-norm which is regime-dependent. 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 4.2, we describe the homogeneous Dirichlet boundary value problem in the sense of Fichera [Oleinik 1973] by introducing some specific

notations, and we precisely define the corresponding discrete setting in Section 4.3. In Section 4.4, we derive the discrete bilinear and linear operators of the discretization method, briefly discuss the static condensation and stability analysis, and precisely delineate the adaptive stabilization strategy. In Section 4.5, extensive numerical experiments are investigated using h- and k-refinement strategies for all abovementioned regimes. We end with some concluding remarks and perspectives.

4.2 Boundary value problem

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

$$\boldsymbol{\nabla}\cdot(-\boldsymbol{\alpha}\boldsymbol{\nabla}u+\boldsymbol{\beta}u)+\mu u=f \quad \text{in } \Omega, \tag{4.1}$$

where $\Omega\subset\mathbb{R}^d$ is a bounded polyhedral domain ($d\geq 2$) of boundary $\partial\Omega$. 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, $\boldsymbol{\alpha}:\Omega\to\mathbb{R}^{d\times d}$ represents an anisotropic heterogeneous dispersion tensor, which is itself a function of the Darcy velocity field $\boldsymbol{\beta}:\Omega\to\mathbb{R}^d$, $\mu:\Omega\to\mathbb{R}^+$, the reaction coefficient, and $f:\Omega\to\mathbb{R}$, a forcing term. We then assume that the constitutive coefficients of (4.1) satisfy the following minimal regularity requirements :

- $\boldsymbol{\alpha}\in[L^\infty(\Omega)]^{d\times d}$ is a symmetric positive semidefinite matrix-valued function verifying that
$$\underline{\alpha}\|\boldsymbol{\zeta}\|^2\leq\boldsymbol{\zeta}^t\boldsymbol{\alpha}(x)\boldsymbol{\zeta}\leq\bar{\alpha}\|\boldsymbol{\zeta}\|^2,\quad\forall\boldsymbol{\zeta}\in\mathbb{R}^d,\quad\text{a.e.}\quad x\in\Omega. \tag{4.2}$$
where $\bar{\alpha}\geq\underline{\alpha}\geq 0$ denote respectively the largest and smallest eigenvalues of $\boldsymbol{\alpha}$, and $\|\cdot\|$ the Euclidean norm. Under the hypothesis (4.2), we shall assume the existence of a subdomain Ω^{ell} (resp. Ω^{hyp}) corresponding to the elliptic (resp. hyperbolic) region s.t.
$$\begin{array}{rcl}\Omega^{\text{ell}} & := & \{x\in\Omega\,:\,\boldsymbol{\zeta}^t\boldsymbol{\alpha}(x)\boldsymbol{\zeta}>0,\quad\forall\boldsymbol{\zeta}\in\mathbb{R}^d\},\\ \Omega^{\text{hyp}} & := & \{x\in\Omega\,:\,\boldsymbol{\zeta}^t\boldsymbol{\alpha}(x)\boldsymbol{\zeta}=0,\quad\forall\boldsymbol{\zeta}\in\mathbb{R}^d\},\end{array} \tag{4.3}$$
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. $\boldsymbol{\nabla}\cdot\boldsymbol{\beta}\in L^\infty(\Omega)$,
- $\mu\in L^\infty(\Omega)$ verifying that the following standard coercivity condition holds
$$\exists\,\mu_0\in\mathbb{R}\quad\text{s.t.}\quad\mu(x)+\frac{1}{2}\boldsymbol{\nabla}\cdot\boldsymbol{\beta}(x)\geq\mu_0>0,\quad\text{a.e.}\quad x\in\Omega. \tag{4.4}$$

Let us now introduce the following disjoint boundary sets as defined by Oleĭnik and Radkevič [Oleinik 1973] (see, e.g., [Di Pietro 2015b] and references therein) ;

$$\begin{array}{l}\Gamma^-:=\{x\in\partial\Omega\,:\,\boldsymbol{n}^t\boldsymbol{\alpha}(x)\boldsymbol{n}>0\quad\text{or}\quad\boldsymbol{\beta}\cdot\boldsymbol{n}<0\},\\ \Gamma^+:=\{x\in\partial\Omega\,:\,\boldsymbol{n}^t\boldsymbol{\alpha}(x)\boldsymbol{n}=0\quad\text{and}\quad\boldsymbol{\beta}\cdot\boldsymbol{n}\geq 0\},\end{array} \tag{4.5}$$

where $\boldsymbol{n}$ denotes the unit outward normal to $\partial\Omega$. Thus, the disjoint subsets $\Gamma^\pm$ will be referred to as the *inflow* and *outflow* 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 (4.1) with homogeneous Dirichlet boundary conditions, namely,

$$u = 0 \quad \text{on } \Gamma^-. \tag{4.6}$$

We point out 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 (4.6) adapt automatically :

- **Nondegenerate problems** : Here, $\Omega^{\text{hyp}} = \emptyset$ and $\boldsymbol{\alpha}$ is assumed to be a symmetric positive definite matrix-valued function on the whole domain Ω. Hence, 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 (4.1) must respect some regularity requirements i.e., $u \in H_0^1(\Omega)$, but it may present some sharp fronts along the characteristic direction $\boldsymbol{\beta}$, particularly for the advective-dominated regime.
- **Fully degenerate problems** : Here, $\Omega^{\text{ell}} = \emptyset$ and we recover the standard advection-reaction problem defined in Ω^{hyp} respecting the coercivity condition (4.4) and the usual definition of the inflow Γ^- and outflow Γ^+ parts of the boundary $\partial\Omega$, respectively. Following these assumptions, the problem is well-posed with no smoothing properties, i.e., $u \in V$ where $V := \{v \in L^2(\Omega) \mid \boldsymbol{\nabla} \cdot (\boldsymbol{\beta} v) \in L^2(\Omega)\}$, and discontinuities in the solution u induced by f (or eventually by Dirichlet boundary conditions) will propagate along the flow field $\boldsymbol{\beta}$, hence giving rise to internal layers.
- **Locally degenerate problems** : We assume that both Ω^{hyp} and Ω^{ell} are nonempty subsets. The model problem is then purely hyperbolic in Ω^{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 $\mathcal{I}^- := \{x \in \mathcal{I} \,:\, \boldsymbol{\beta}(x) \cdot \boldsymbol{n}_{\mathcal{I}} < 0\}$, where $\boldsymbol{n}_{\mathcal{I}}$ 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 subregion to the elliptic one. Thus, we set $\mathcal{I}^+ := \mathcal{I} \backslash \mathcal{I}^-$.

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

$$\begin{aligned}
\boldsymbol{\nabla} \cdot (-\boldsymbol{\alpha}\boldsymbol{\nabla} u + \boldsymbol{\beta} u) + \mu u &= f \quad \text{in} \quad \Omega^{\text{ell}} \cup \Omega^{\text{hyp}}, \\
[\![-\boldsymbol{\alpha}\boldsymbol{\nabla} u + \boldsymbol{\beta} u]\!] &= 0 \quad \text{on} \quad \mathcal{I}, \\
[\![u]\!] \cdot \boldsymbol{n}_{\mathcal{I}} &= 0 \quad \text{on} \quad \mathcal{I}^+, \\
u &= 0 \quad \text{on} \quad \Gamma^-,
\end{aligned} \tag{4.7}$$

where $[\![\cdot]\!]$ denotes the standard jump operator (in a DG sense) as defined in [Arnold 2002]. The well-posedness of the boundary value problem (4.7) has been analyzed by Oleĭnik and Radkevič in [Oleinik 1973]. 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 an inspired and unified H-IP formalism 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.

4.3 Discrete setting

4.3.1 Mesh notations

Let h be a positive parameter, and assume without loss of generality that $h \leq 1$. We denote by $\mathcal{T}_h^{\mathrm{ell}}$ (resp. $\mathcal{T}_h^{\mathrm{hyp}}$) a conformal partition of the subdomain Ω^{ell} (resp. Ω^{hyp}) satisfying $\mathcal{T}_h^{\mathrm{ell}} \cap \mathcal{T}_h^{\mathrm{hyp}} = \emptyset$. The set of all mesh elements is denoted by $\mathcal{T}_h$, i.e., $\mathcal{T}_h := \mathcal{T}_h^{\mathrm{ell}} \cup \mathcal{T}_h^{\mathrm{hyp}}$, where h stands for the largest diameter of all elements. We precisely state here that the sets of mesh elements $\mathcal{T}_h^{\mathrm{ell}}$ and $\mathcal{T}_h^{\mathrm{hyp}}$ 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 $\mathcal{F}_h^{\mathfrak{b}}$, i.e., $F \in \mathcal{F}_h^{\mathfrak{b}}$ if there exists E in $\mathcal{T}_h$ such that $F := \partial E \cap \partial\Omega$. We assume that the set $\mathcal{F}_h^{\mathfrak{b}}$ coincides with the disjoint boundary partition $\mathcal{F}_h^{\mathfrak{b}\pm}$ corresponding to the discrete counterpart of $\Gamma^{\pm}$. Likewise, we denote by $\mathcal{F}_h^{\mathfrak{i}}$ the set of interior interfaces, i.e., $F \in \mathcal{F}_h^{\mathfrak{i}}$ 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., $\mathcal{F}_h := \mathcal{F}_h^{\mathfrak{i}} \cup \mathcal{F}_h^{\mathfrak{b}}$, and we set $\mathcal{F}_h^{+} := \mathcal{F}_h^{\mathfrak{i}} \cup \mathcal{F}_h^{\mathfrak{b}+}$. In particular, we denote by $\mathcal{I}_h$ the subset of $\mathcal{F}_h^{\mathfrak{i}}$, which belongs to $\mathcal{I}$ ($\mathcal{I}_h \subset \mathcal{F}_h^{\mathfrak{i}}$), i.e., $\mathcal{I}_h$ (resp. $\mathcal{I}_h^{\pm}$) corresponds here to the discrete counterpart of $\mathcal{I}$ (resp. $\mathcal{I}^{\pm}$).

Assumption 4.1 (Compatible meshes)**.** The h-refined mesh sequences $(\mathcal{T}_h)_h$ must fulfill the following assumptions :

(A) Any element $E \in \mathcal{T}_h$ lies entirely in one of the disjoint subsets $\mathcal{T}_h^{\mathrm{ell}}$ or $\mathcal{T}_h^{\mathrm{hyp}}$;

(B) Any interface $F \in \mathcal{F}_h^{\mathfrak{b}}$ lies entirely in one of the disjoint subsets $\mathcal{F}_h^{\mathfrak{b}\pm}$;

(C) Any interface $F \in \mathcal{I}_h$ lies entirely in one of the disjoint subsets $\mathcal{I}_h^{\pm}$.

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 $\eta_E := \mathrm{card}(\mathcal{F}_E)$. We finally set $\partial\mathcal{T}_h := \{\cup\mathcal{F}_E \ : \ \forall E \in \mathcal{T}_h\}$ the collection of interfaces composing the boundary of all mesh elements, and we assume that $\partial\mathcal{T}_h := \partial\mathcal{T}_h^{\mathrm{ell}} \cup \partial\mathcal{T}_h^{\mathrm{hyp}}$. For any element X of $\mathcal{T}_h$ or $\mathcal{F}_h$, we then denote by $|X|$ and h_X its measure and diameter, respectively.

4.3.2 Discrete approximation spaces

For any polyhedral domain $\mathcal{D} \subset \mathbb{R}^d$ with $\partial\mathcal{D} \subset \mathbb{R}^{d-1}$, we denote by $(\cdot,\cdot)_{0,\mathcal{D}}$ (resp. $\langle\cdot,\cdot\rangle_{0,\partial\mathcal{D}}$) the L^2-inner product in $L^2(\mathcal{D})$ (resp. $L^2(\partial\mathcal{D})$) equipped with its natural norm $\|\cdot\|_{0,\mathcal{D}}$ (resp. $\|\cdot\|_{0,\partial\mathcal{D}}$). For any $E \in \mathcal{T}_h$, we set $\langle\cdot,\cdot\rangle_{0,\partial E} := \sum_{F\in\mathcal{F}_E}\langle\cdot,\cdot\rangle_{0,F}$, and we denote by $\|\cdot\|_{0,\partial E}$ its corresponding norm. Similarly, we denote by $H^s(\mathcal{D})$ the usual Hilbert space of index s on $\mathcal{D}$ equipped with its natural norm $\|\cdot\|_{s,\mathcal{D}}$ and seminorm $|\cdot|_{s,\mathcal{D}}$, respectively. In particular, if $s = 0$, then we set $H^0(\mathcal{D}) = L^2(\mathcal{D})$. We now denote by $H^s(\mathcal{T}_h)$ the usual broken Sobolev

space and by $\boldsymbol{\nabla}_h$ the broken gradient operator acting on $H^s(\mathcal{T}_h)$ with $s \geq 1$. Let us now introduce compact notations associated with the discrete L^2-inner scalar product :

$$(\cdot,\cdot)_{0,\mathcal{T}_h} := \sum_{E\in\mathcal{T}_h} (\cdot,\cdot)_{0,E}, \quad \langle\cdot,\cdot\rangle_{0,\partial\mathcal{T}_h} := \sum_{E\in\mathcal{T}_h} \langle\cdot,\cdot\rangle_{0,\partial E} \quad \text{and} \quad \langle\cdot,\cdot\rangle_{0,\mathcal{F}_h} := \sum_{F\in\mathcal{F}_h} \langle\cdot,\cdot\rangle_{0,F}, \quad (4.8)$$

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 DG methods, we consider broken polynomial spaces :

$$\mathcal{P}_k(\mathcal{T}_h) := \{v_h \in H^1(\mathcal{T}_h) \,:\, v_h|_E \in \mathcal{P}_k(E),\ \forall E \in \mathcal{T}_h\}, \quad (4.9a)$$

$$\mathcal{P}_k(\mathcal{F}_h) := \{\hat{v}_h \in L^2(\mathcal{F}_h) \,:\, \hat{v}_h|_F \in \mathcal{P}_k(F),\ \forall F \in \mathcal{F}_h\}, \quad (4.9b)$$

where $L^2(\mathcal{F}_h)$ corresponds to the space of functions that are single-valued on all elements of $\mathcal{F}_h$. Here, $\mathcal{P}_k(X)$ 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 (4.7). First, the discrete variable $u_h \in \mathcal{V}_h$ which is defined within each mesh element, and its trace $\hat{u}_h \in \hat{\mathcal{V}}_h^0$ defined on the mesh skeleton verifying the imposed Dirichlet boundary conditions at the boundary part Γ^-. Thus, we set :

$$\mathcal{V}_h := \mathcal{P}_k(\mathcal{T}_h) \quad \text{and} \quad \hat{\mathcal{V}}_h^0 := \{\hat{v}_h \in \mathcal{P}_k(\mathcal{F}_h) \,:\, \hat{v}_h|_F = 0,\ \forall F \in \mathcal{F}_h^{\mathfrak{b}-}\}. \quad (4.10)$$

For clarity, we then introduce the compact discrete variable $\boldsymbol{u}_h := (u_h, \hat{u}_h)$ belonging to the composite approximation space $\boldsymbol{\mathcal{V}}_h := \mathcal{V}_h \times \hat{\mathcal{V}}_h^0$, i.e., $\boldsymbol{u}_h \in \boldsymbol{\mathcal{V}}_h$. Finally, let us retain the discrete trace inequality. For all $v_h \in \mathcal{V}_h$, all $E \in \mathcal{T}_h$, and all $F \in \mathcal{F}_E$, then the following holds,

$$\|v_h\|_{0,F} \leq C_{\mathrm{tr}} h_E^{-1/2} \|v_h\|_{0,E}, \quad (4.11)$$

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

Remark 4.1 (Regularity of $\hat{u}$ vs. $\hat{u}_h$). Let $\mathcal{T}_h$ be a compatible partition as given in Assumption 4.1. As discussed in Section 4.2, the trace $\hat{u}$ of exact solution u of the problem (4.7) is globally single-valued on $\mathcal{F}_h$ except at the portion $\mathcal{I}_h^-$. Namely, for all $F := \partial E_1 \cap \partial E_2 \in \mathcal{I}_h^-$ with $E_1 \in \mathcal{T}_h^{\mathrm{hyp}}$ and $E_2 \in \mathcal{T}_h^{\mathrm{ell}}$ and by setting $\hat{u}_i := \hat{u}|_{E_i,F}$, we then infer that its jump $\hat{u}_1 - \hat{u}_2$ is non-null on F. However, the situation is somewhat different at the discrete level. Indeed, its discrete approximation $\hat{u}_h \in \hat{\mathcal{V}}_h^0 \subset L^2(\mathcal{F}_h)$ is assumed to be single-valued (continuous) at the set of all interfaces of the mesh skeleton, including $\mathcal{I}_h^-$ which seems to be inappropriate. However, the tightened regularity of $\hat{u}_h$ on $\mathcal{I}_h^-$ is rendered possible by the surprising fact that continuous interface-based DOFs can capture a discontinuous solution well upon an appropriate stabilization strategy. We refer the interested readers to the discussion given in Section 4.4.4.

4.3.3 Discrete trace operators

We first specify the functional setting concerning the discrete jump trace operator used in the rest of the paper. For this purpose, let $L^2(\partial\mathcal{T}_h)$ be the space of functions which are

double-valued on $\mathcal{F}_h^{\mathsf{i}}$ and single-valued on $\mathcal{F}_h^{\mathsf{b}}$. We point out here the main difference with the usual space $L^2(\mathcal{F}_h)$ which is a subset of $L^2(\partial\mathcal{T}_h)$, i.e., $L^2(\mathcal{F}_h) \subset L^2(\partial\mathcal{T}_h)$. We are now in a position to define the jump operator in a HDG sense.

Definition 4.1 (HDG-jump operator)**.** *Assume $s \geq 1$, we then define the HDG-jump operator as an application such that $[\![\cdot]\!] : H^s(\mathcal{T}_h) \times L^2(\partial\mathcal{T}_h) \to [L^2(\partial\mathcal{T}_h)]^d$. Concretely, for all $E \in \mathcal{T}_h$, all $F \in \mathcal{F}_E$ and all $\boldsymbol{v} := (v, \hat{v}) \in H^s(\mathcal{T}_h) \times L^2(\partial\mathcal{T}_h)$, its restriction to F is given by the formula :*

$$[\![\boldsymbol{v}]\!]|_{E,F} := (v|_{E,F} - \hat{v}|_{E,F})\boldsymbol{n}_{E,F}, \tag{4.12}$$

where $\boldsymbol{n}_{E,F}$ denotes the unit normal to F pointing out of E. For clarity, we then denote it in its compact form as follows : $[\![\boldsymbol{v}]\!] := (v - \hat{v})\boldsymbol{n}$ on $\partial\mathcal{T}_h$.

Following Definition 4.1, we point out that $[\![\boldsymbol{v}]\!] \in [L^2(\partial\mathcal{T}_h)]^d$ can be distinctively evaluated on the two branches of any interior interface $F \in \mathcal{F}_h^{\mathsf{i}}$. Finally, let us remind the well-known definitions of the weighted-average operator and its conjugate denoted by $\{\!\{\cdot\}\!\}_{\boldsymbol{\omega}}$ and $\{\!\{\cdot\}\!\}^{\boldsymbol{\omega}}$, respectively. For all $F \in \mathcal{F}_h^{\mathsf{i}}$ with $F := \partial E_1 \cap \partial E_2$ and $v \in H^s(\mathcal{T}_h)$ with $s \geq 1$, we set :

$$\{\!\{v\}\!\}_{\boldsymbol{\omega}} := \omega_1 v_1 + \omega_2 v_2 \quad \text{and} \quad \{\!\{v\}\!\}^{\boldsymbol{\omega}} := \omega_2 v_1 + \omega_1 v_2, \tag{4.13}$$

where $v_i = v|_{E_i,F}$ and $\boldsymbol{\omega} := (\omega_1, \omega_2)$ is a double-valued function verifying that weights satisfy $\omega_1 + \omega_2 = 1$. If $F \in \mathcal{F}_h^{\mathsf{b}}$, we then assume that $\{\!\{v\}\!\}_{\boldsymbol{\omega}} = \{\!\{v\}\!\}^{\boldsymbol{\omega}} := v$. If $\boldsymbol{\omega} = (1/2, 1/2)$, we then recover the classical average operator, and we will omit the subscript $\boldsymbol{\omega}$ in their definitions. Let us precise that (4.13) can be extended to the class of vector-valued functions, i.e., for all $\boldsymbol{\varphi} \in [H^s(\mathcal{T}_h)]^d$.

Remark 4.2 (Restricted-jump operator on $\boldsymbol{\mathcal{V}}_h$). Let us now precise the restriction of the jump operator as given in Definition 4.1 to the set of discrete functions $\boldsymbol{v}_h$ belonging to $\boldsymbol{\mathcal{V}}_h$, i.e., we now assume that $[\![\cdot]\!] : H^s(\mathcal{T}_h) \times L^2(\mathcal{F}_h) \to [L^2(\partial\mathcal{T}_h)]^d$. Thus, for all $E \in \mathcal{T}_h$, all $F \in \mathcal{F}_E$ and all $\boldsymbol{v}_h \in \boldsymbol{\mathcal{V}}_h$, its restriction to F is now explicitly given by :

$$[\![\boldsymbol{v}_h]\!]|_{E,F} := (v_h|_{E,F} - \hat{v}_h|_F)\boldsymbol{n}_{E,F},$$

since $\hat{v}_h$ is assumed to be single-valued on $\mathcal{F}_h$.

4.4 Hybridized interior penalty discontinuous Galerkin method

In this section, we describe the primal HDG method for solving the problem (4.7). First, we derive it intuitively, and we then propose a compact notation of all of these variants. The strong consistency and coercivity properties are then 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.

4.4.1 Intuitive derivation

The discrete formulation of the continuous model problem (4.7) can be intuitively derived with respect to the three following steps :

Step 1. **Global weak formulation** – Let $\boldsymbol{u}_h, \boldsymbol{v}_h \in \boldsymbol{\mathcal{V}}_h$. For all $E \in \mathcal{T}_h$, multiplying (4.1) by a test function v_h, and integrating by parts over E, we obtain a local equation :

$$-(\boldsymbol{\sigma}_h(u_h), \boldsymbol{\nabla}_h v_h)_{0,E} + \langle \hat{\boldsymbol{\sigma}}_h(\boldsymbol{u}_h) \cdot \boldsymbol{n}, v_h \rangle_{0,\partial E} + (\mu u_h, v_h)_{0,E} = (f, v_h)_{0,E}, \tag{4.14}$$

where $\boldsymbol{\sigma}_h(u_h) := -\boldsymbol{\alpha}\boldsymbol{\nabla}_h u_h + \boldsymbol{\beta} u_h$ corresponds to the approximation of the total flux on E, and $\hat{\boldsymbol{\sigma}}_h(\boldsymbol{u}_h)$ its trace approximation on ∂E that we will precisely define below. By summing (4.14) over all elements E of $\mathcal{T}_h$, we then obtain the global equation,

$$-(\boldsymbol{\sigma}_h(u_h), \boldsymbol{\nabla}_h v_h)_{0,\mathcal{T}_h} + \langle \hat{\boldsymbol{\sigma}}_h(\boldsymbol{u}_h) \cdot \boldsymbol{n}, v_h \rangle_{0,\partial\mathcal{T}_h} + (\mu u_h, v_h)_{0,\mathcal{T}_h} = (f, v_h)_{0,\mathcal{T}_h}. \tag{4.15}$$

We emphasize that $\hat{\boldsymbol{\sigma}}_h(\boldsymbol{u}_h)$ can be evaluated distinctively on both sides of a given interface. Consequently, an additional equation needs to be included in (4.15) to ensure the continuity requirements.

Step 2. **Transmission conditions** – By considering that $\hat{v}_h \in \hat{\mathcal{V}}_h^0$ vanishes on $\mathcal{F}_h^{\mathfrak{b}-}$, we impose the global transmission requirements respecting the outflow boundary condition on $\mathcal{F}_h^{\mathfrak{b}+}$:

$$\langle \hat{\boldsymbol{\sigma}}_h(\boldsymbol{u}_h) \cdot \boldsymbol{n}, \hat{v}_h \rangle_{0,\partial\mathcal{T}_h} - \langle (\boldsymbol{\beta} \cdot \boldsymbol{n}) \hat{u}_h, \hat{v}_h \rangle_{0,\mathcal{F}_h^{\mathfrak{b}+}} = 0, \tag{4.16}$$

which is the statement of weak continuity of the normal flux $\hat{\boldsymbol{\sigma}}_h(\boldsymbol{u}_h)$ across the mesh skeleton $\mathcal{F}_h^+$. Thus, subtracting (4.16) from (4.15) leads to

$$\begin{aligned} -(\boldsymbol{\sigma}_h(u_h), \boldsymbol{\nabla}_h v_h)_{0,\mathcal{T}_h} + (\mu u_h, v_h)_{0,\mathcal{T}_h} + \langle \hat{\boldsymbol{\sigma}}_h(\boldsymbol{u}_h), [\![\boldsymbol{v}_h]\!] \rangle_{0,\partial\mathcal{T}_h} \\ + \langle (\boldsymbol{\beta} \cdot \boldsymbol{n}) \hat{u}_h, \hat{v}_h \rangle_{0,\mathcal{F}_h^{\mathfrak{b}+}} = (f, v_h)_{0,\mathcal{T}_h}. \end{aligned} \tag{4.17}$$

Step 3. **Numerical Flux in the HDG sense** – The discrete problem (4.17) is then closed by assuming the following simple form of $\hat{\boldsymbol{\sigma}}_h(\boldsymbol{u}_h)$ on the mesh skeleton, i.e.,

$$\hat{\boldsymbol{\sigma}}_h(\boldsymbol{u}_h) := \boldsymbol{\sigma}_h(u_h) + \tau [\![\boldsymbol{u}_h]\!] \quad \text{on } \partial\mathcal{T}_h, \tag{4.18}$$

where τ denotes the stabilization penalty function that we describe precisely below.

Inserting (4.18) into (4.17) leads to the incomplete form of the hybridized interior penalty DG (H-IIP) method. Several variations of this methodology can then be derived by controlling the introduction of additional consistent terms to the discrete formulation (4.17) that we summarize below.

4.4.2 Compact discrete formulation

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

$$\mathcal{B}_h^{(\varepsilon)}(\boldsymbol{u}_h, \boldsymbol{v}_h) = \mathcal{L}_h(\boldsymbol{v}_h), \qquad \forall \boldsymbol{v}_h \in \boldsymbol{\mathcal{V}}_h, \tag{4.19}$$

where $\mathcal{L}_h(\boldsymbol{v}_h) := (f, v_h)_{0,\mathcal{T}_h}$ and the bilinear form $\mathcal{B}_h^{(\varepsilon)} : \boldsymbol{\mathcal{V}}_h \times \boldsymbol{\mathcal{V}}_h \to \mathbb{R}$ can be linearly decomposed following its diffusive, advective-reactive, and stability contributions, respectively :

$$\mathcal{B}_h^{(\varepsilon)}(\boldsymbol{u}_h, \boldsymbol{v}_h) := \mathcal{B}_{\boldsymbol{\alpha},h}^{(\varepsilon)}(\boldsymbol{u}_h, \boldsymbol{v}_h) + \mathcal{B}_{\boldsymbol{\beta},\mu,h}(\boldsymbol{u}_h, \boldsymbol{v}_h) + \mathcal{S}_h(\boldsymbol{u}_h, \boldsymbol{v}_h). \tag{4.20}$$

Here, $\mathcal{B}^{(\varepsilon)}_{\boldsymbol{\alpha},h}$ corresponds solely to the H-IP discretization of the diffusive part and is given by :

$$\mathcal{B}^{(\varepsilon)}_{\boldsymbol{\alpha},h}(\boldsymbol{u}_h,\boldsymbol{v}_h) := (\boldsymbol{\alpha}\boldsymbol{\nabla}_h u_h, \boldsymbol{\nabla}_h v_h)_{0,\mathcal{T}_h^{\mathrm{ell}}} - \langle \boldsymbol{\alpha}\boldsymbol{\nabla}_h u_h, [\![\boldsymbol{v}_h]\!]\rangle_{0,\partial\mathcal{T}_h^{\mathrm{ell}}} - \varepsilon\langle \boldsymbol{\alpha}\boldsymbol{\nabla}_h v_h, [\![\boldsymbol{u}_h]\!]\rangle_{0,\partial\mathcal{T}_h^{\mathrm{ell}}}, \quad (4.21)$$

where the parameter $\varepsilon \in \{0, \pm 1\}$ controls the introduction of the (consistent) symmetry term $\langle \boldsymbol{\alpha}\boldsymbol{\nabla}_h v_h, [\![\boldsymbol{u}_h]\!]\rangle_{0,\partial\mathcal{T}_h^{\mathrm{ell}}}$. Similar to the standard IPDG methods, $\varepsilon = 0$ corresponds to the Incomplete scheme denoted as the H-IIP method as described above, while $\varepsilon = +1$ (resp. $\varepsilon = -1$) denotes the Symmetric (resp. Non-Symmetric) scheme denoted as the H-SIP (resp. H-NIP) method. Let us point out that the diffusive contribution automatically vanishes for hyperbolic problems since $\Omega^{\mathrm{ell}} := \emptyset$. Thus, the advective-reactive part is discretized as follows :

$$\mathcal{B}_{\boldsymbol{\beta},\mu,h}(\boldsymbol{u}_h,\boldsymbol{v}_h) := -(\boldsymbol{\beta}u_h, \boldsymbol{\nabla}_h v_h)_{0,\mathcal{T}_h} + (\mu u_h, v_h)_{0,\mathcal{T}_h} + \langle \boldsymbol{\beta}u_h, [\![\boldsymbol{v}_h]\!]\rangle_{0,\partial\mathcal{T}_h} + \langle \beta\hat{u}_h, \hat{v}_h\rangle_{0,\mathcal{F}_h^{\mathfrak{b}+}}, \quad (4.22)$$

where $\beta := \boldsymbol{\beta}\cdot\boldsymbol{n}$. The last quantity $\mathcal{S}_h$ in (4.20) is called the discrete stability form based on jump-penalty terms as in Definition 4.1 and is given by :

$$\mathcal{S}_h(\boldsymbol{u}_h,\boldsymbol{v}_h) := \langle \tau[\![\boldsymbol{u}_h]\!], [\![\boldsymbol{v}_h]\!]\rangle_{0,\partial\mathcal{T}_h}. \quad (4.23)$$

Here $\tau \in L^2(\partial\mathcal{T}_h)$ denotes the stabilization function, which can also be distinctively evaluated on the mesh skeleton of the hyperbolic or elliptic subregions, respectively ; e.g.,

$$\tau := \begin{cases} \tau^{\mathrm{hyp}} & \text{on} \quad \partial\mathcal{T}_h^{\mathrm{hyp}}, \\ \tau^{\mathrm{ell}} & \text{on} \quad \partial\mathcal{T}_h^{\mathrm{ell}}. \end{cases} \quad (4.24)$$

Thus, a tailored design of τ will be investigated later in Section 4.4.4, taking into account the mathematical structure of the local cellwise problem alone. In practice, the definition of τ^{ell} is generally prescribed accordingly to the predominant mechanism, i.e., diffusion vs. advection, inside the elliptic region. Likewise, τ^{hyp} must account for only the advection mechanism inside the hyperbolic part. This adaptive strategy turns out to be particularly appropriate for the set of interfaces of $\mathcal{I}_h$ separating both subregions, as discussed in the remark 4.4. In the following, we will specify its dependence on the physical variables of the problem, namely, $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and then detail its assembly methods.

Thus, the rest of our analysis hinges on the following assumptions regarding τ.

Assumption 4.2 (Characterization of $\boldsymbol{\tau}$)**.** The stabilization function $\tau : \mathbb{R}\times\mathbb{R} \to \mathbb{R}_*^+$ (4.23) is a positive definite double-valued function chosen such that :

(A1) For all $s, t \in \mathbb{R}_*$, the following invariance properties hold :

$$|s| := \tau(s,0), \quad \text{and} \quad |t| := \tau(0,t). \quad (4.25)$$

The first and second assumptions in (4.25) allow for an encompassed treatment of both extremal configurations, i.e., the pure-diffusive and pure-advective regimes characterized by $\tau(\cdot,0)$ and $\tau(0,\cdot)$, respectively. We refer the interested reader to Remark 4.3 discussing dependencies of τ and their concrete physical meanings.

(A2) For all $s, t \in \mathbb{R}$, we shall assume the existence of a positive constant $\tau_{\min}$ such that,

$$\min \left[\tau(s,t) - |s| + \frac{\boldsymbol{\beta} \cdot \boldsymbol{n}}{2} \right] \Big|_{\partial \mathcal{T}_h} \geq \tau_{\min} > 0. \tag{4.26}$$

The second assumption is crucial to introduce an appropriate definition of the energy-norm which is regime dependent, and thereby ensure a discrete coercivity stability condition (see, e.g., Lemma 4.4.3). The formula (4.26) means that τ must be chosen on the mesh skeleton such that the following constraint holds : for all $s, t \in \mathbb{R}$, then

$$\tau(s,t) > |s| - \frac{\boldsymbol{\beta} \cdot \boldsymbol{n}}{2} \quad \text{on} \quad \partial \mathcal{T}_h. \tag{4.27}$$

Remark 4.3 (Dependencies of τ). Different formulations of the stabilization function τ have been derived and analyzed in the literature for the class of advection-diffusion-reaction problems. We assume that τ is a positive definite bi-valued function such that $\tau(\tau_{\boldsymbol{\alpha}}, \tau_{\boldsymbol{\beta}})$, where $\tau_{\boldsymbol{\alpha}}$ and $\tau_{\boldsymbol{\beta}}$ correspond to an appropriate and arbitrary definition of the penalty parameter in the context of a pure-diffusive and pure-advective regime, respectively. Under hypothesis (A1) given in Assumption 4.2, we infer that $\tau_{\boldsymbol{\alpha}} := \tau(\cdot, 0) \geq 0$ and $\tau_{\boldsymbol{\beta}} := \tau(0, \cdot) \geq 0$. We further assume that both parameters $\tau_{\boldsymbol{\alpha}}$ and $\tau_{\boldsymbol{\beta}}$ cannot be simultaneously null and that one of these distinctive mechanisms is still involved at least.

Remark 4.4 (Neutralization of the role of $\mathcal{I}_h^-$). Analyzing the discrete bilinear form $\mathcal{B}_h^{(\varepsilon)}$ in (4.20), we emphasize that no specific ad hoc techniques are required to capture the (possible) jump of the exact solution on $\mathcal{I}_h^-$. Indeed, all interfaces are addressed in the same fashion as standard $\partial \mathcal{T}_h$ elements, and the penalty terms are locally designed in an automated fashion according to the mathematical structure of PDEs inside the corresponding adjacent cell. Thus, under the mesh assumption 4.1, the subsets of separative interfaces $\mathcal{I}_h^{\pm}$ are not required to be clearly identified a priori, which is a significant numerical improvement. Conversely, we refer the interested readers to the work of [Gastaldi 1989], where they employed specific ad hoc techniques based on the removal of relevant terms at the portion $\mathcal{I}_h^-$.

Remark 4.5 (Static condensation). As mentioned in [Cockburn 2016a] (see, e.g., [Cockburn 2016b]), the technique of static condensation was introduced to reduce the size of the discrete matrix associated with the global problem (4.19). Let $\mathbf{U}_h := [\mathrm{U}_h, \hat{\mathrm{U}}_h]$ denote the vector of DOFs of the composite variable $\boldsymbol{u}_h \in \boldsymbol{\mathcal{V}}_h$, which is composed of interior- and interface-based DOFs denoted by U_h and $\hat{\mathrm{U}}_h$, respectively. The strategy consists of eliminating interior-based unknowns from the above equations by successively projecting (4.19) on $(v_h, 0)$ and $(0, \hat{v}_h)$. We thus obtain

$$\begin{array}{rcl} \mathcal{B}_h^{(\varepsilon)}(\boldsymbol{u}_h, (v_h, 0)) & : & \mathcal{B}_{uu}\mathrm{U}_h + \mathcal{B}_{u\hat{u}}\hat{\mathrm{U}}_h = \mathrm{F}, \\ \mathcal{B}_h^{(\varepsilon)}(\boldsymbol{u}_h, (0, \hat{v}_h)) & : & \mathcal{B}_{\hat{u}u}\mathrm{U}_h + \mathcal{B}_{\hat{u}\hat{u}}\hat{\mathrm{U}}_h = 0, \end{array} \tag{4.28}$$

where $\mathcal{B}_{\hat{u}u}$ corresponds to the interaction matrix between the discrete variables $\hat{u}_h$ and u_h, and similarly for $\mathcal{B}_{uu}$, $\mathcal{B}_{u\hat{u}}$ and $\mathcal{B}_{\hat{u}\hat{u}}$, respectively. Due to the discontinuous nature of $\mathcal{V}_h$, all

computations can be performed cellwise, leading to a block-diagonal matrix $\mathcal{B}_{uu}$, which can be easily inverted and locally eliminated. Thus we obtain the final linear system :

$$[\mathcal{B}_{\hat{u}\hat{u}} - \mathcal{B}_{\hat{u}u}\mathcal{B}_{uu}^{-1}\mathcal{B}_{u\hat{u}}]\hat{\mathrm{U}}_h = \mathrm{G}. \tag{4.29}$$

The matrix on the left-hand-side of (4.29) is called the Shur complement of $\mathcal{B}_{uu}$. We note that the original discrete problem (4.19) and its reduced version (4.29) are globally nonsymmetric since the continuous problem (4.7) is itself nonsymmetric. This is due to the presence of the advective term.

4.4.3 Convergence analysis

We now check some favorable properties such as the (strong) consistency and discrete stability, based on coercivity of all ε-variants of the H-IP method. This is to ensure the existence and uniqueness of the numerical solution via the Lax–Milgram Lemma.

4.4.3.1 Consistency

Lemma 4.4.1 (Strong consistency)**.** *Let u be the exact solution of the problem (4.7), $\hat{u} := u$ on $\partial\mathcal{T}_h$, and we denote $\boldsymbol{u} := (u, \hat{u})$ its compact notation. In addition to hypothesizing Assumption 4.1, the following holds :*

$$\mathcal{B}_h^{(\varepsilon)}(\boldsymbol{u}, \boldsymbol{v}_h) = \mathcal{L}_h(\boldsymbol{v}_h), \qquad \forall \boldsymbol{v}_h \in \boldsymbol{\mathcal{V}}_h, \tag{4.30}$$

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

Proof. Let $\mathcal{T}_h$ be a given compatible partition of Ω as given in Assumption 4.1 ; then, following the discussion given in Remark 4.1, the regularity of $\boldsymbol{u}$ implies that its trace $\hat{u}$ is :

(1) single-valued on $\mathcal{F}_h \backslash \mathcal{I}_h^-$ and,

(2) (possibly) double-valued only on the subset $\mathcal{I}_h^-$.

The strategy to prove that $\mathcal{B}_h^{(\varepsilon)}$ is consistent is to establish first that the jump of the composite variable $\boldsymbol{u}$ is null on the mesh skeleton. Toward this aim, we exploit the properties of the jump operator as given in Definition 4.1, and the regularity of $\boldsymbol{u}$ as discussed in Remark 4.1. By setting $\hat{u} := u$ on $\partial\mathcal{T}_h$, we infer that $[\![\boldsymbol{u}]\!] := (u - \hat{u})\boldsymbol{n}$ is always null on $\partial\mathcal{T}_h$ (including $\mathcal{I}_h^-$) simplifying the expression of (4.20). Thus, the rest of the proof follows by using classical arguments :

(1) Second, by setting $\boldsymbol{v}_h := (v_h, 0)$, integrating by parts on each element of the mesh, and finally observing that $[\![-\boldsymbol{\alpha}\boldsymbol{\nabla}_h u + \boldsymbol{\beta} u]\!] = 0$ on internal interfaces, we immediately infer

$$\mathcal{B}_h^{(\varepsilon)}(\boldsymbol{u}, (v_h, 0)) := \sum_{E \in \mathcal{T}_h} (\boldsymbol{\nabla}_h \cdot (-\boldsymbol{\alpha}\boldsymbol{\nabla}_h u + \boldsymbol{\beta} u) + \mu u, v_h)_{0,E} = \sum_{E \in \mathcal{T}_h} (f, v_h)_{0,E}. \tag{4.31}$$

(2) Finally, by setting $\boldsymbol{v}_h := (0, \hat{v}_h) \in \boldsymbol{\mathcal{V}}_h$, we then obtain

$$\mathcal{B}_h^{(\varepsilon)}(\boldsymbol{u}, (0, \hat{v}_h)) := -\sum_{E \in \mathcal{T}_h} \langle (-\boldsymbol{\alpha}\boldsymbol{\nabla}_h u + \boldsymbol{\beta} u) \cdot \boldsymbol{n}, \hat{v}_h \rangle_{0,\partial E} + \langle \beta u, \hat{v}_h \rangle_{0,\mathcal{F}_h^{\flat +}} = 0, \tag{4.32}$$

which corresponds to the (imposed) transmission conditions.

The proof is then completed by summing (4.31) and (4.32). ■

Remark 4.6. Let us highlight here the originality of the proof of Lemmata 4.4.1 using novel arguments based on HDG-jump operators. It fundamentally differs from the one proposed in [Di Pietro 2008] by privileging other arguments based on weighted average operators.

4.4.3.2 Discrete coercivity stability

We now introduce the self-adaptive energy-norm $\|\cdot\|_\dagger$ equipping $\boldsymbol{\mathcal{V}}_h$, which is regime-dependent as discussed in the following Remark 4.7. For all $\boldsymbol{v}_h \in \boldsymbol{\mathcal{V}}_h$, it is thus given by ;

$$\|\boldsymbol{v}_h\|_\dagger^2 := \|\boldsymbol{\alpha}^{1/2}\boldsymbol{\nabla}_h v_h\|_{0,\mathcal{T}_h^{\rm ell}}^2 + \|\mu_0^{1/2} v_h\|_{0,\mathcal{T}_h}^2 + \frac{1}{2}\|\beta^{1/2}\hat{v}_h\|_{0,\mathcal{F}_h^{\flat+}}^2 + |\boldsymbol{v}_h|_{\tau_{\min}}^2, \tag{4.33}$$

which clearly depends on constitutive coefficients $\boldsymbol{\alpha}$ defined inside the elliptic region $\Omega^{\rm ell}$, $\beta := \boldsymbol{\beta}\cdot\boldsymbol{n}$, defined on the outflow part of the boundary, and parameters μ_0 and $\tau_{\min}$ as given in (4.4) and (4.26), respectively. Here, $|\cdot|_\varrho$ corresponds to the (HDG-)jump seminorm, which is defined as follows ; i.e., for all $\boldsymbol{v}_h \in \boldsymbol{\mathcal{V}}_h$ then,

$$|\boldsymbol{v}_h|_\varrho^2 := \sum_{E\in\mathcal{T}_h} |\boldsymbol{v}_h|_{\varrho,\partial E}^2 \quad \text{with} \quad |\boldsymbol{v}_h|_{\varrho,\partial E}^2 := \sum_{F\in\mathcal{F}_E} \|\varrho_{E,F}^{1/2}[\![\boldsymbol{v}_h]\!]\|_{0,F}^2, \tag{4.34}$$

where $\varrho_{E,F} \geq 0$. The norm (4.33) is well-defined owing to assumptions (4.4) and (4.26) for all abovementioned regimes, and is used to prove the coercivity of (4.20) on $\boldsymbol{\mathcal{V}}_h$. For this purpose, we first need to introduce an appropriate definition of the diffusive penalty function $\tau_{\boldsymbol{\alpha}}$ on $\partial\mathcal{T}_h$ and hence establish an intermediate result, which is a minor adaptation to the $\boldsymbol{\alpha}$-tensor field of the proof provided by Wells in [Wells 2011].

Remark 4.7 (Self-adaptive energy-norm). Let us highlight that the energy-norm as defined in (4.33) allows for an encompassed treatment of all mentioned regimes and automatically adapts to the mathematical structure of the considered problem. Indeed, it appears that its first (resp. third) summand naturally vanishes for hyperbolic (resp. elliptic) problems since $\Omega^{\rm ell} := \emptyset$ (resp. $\Gamma^+ = \emptyset$).

Definition 4.2 (Diffusive stabilization function)**.** *For all $E \in \mathcal{T}_h$ and all $F \in \mathcal{F}_E$, the diffusive penalization function $\tau_{\boldsymbol{\alpha}}$ is defined such that its restriction on F is given by :*

$$\tau_{\boldsymbol{\alpha}}|_{E,F} := \begin{cases} \tau_0 C_{\rm tr}^2 \dfrac{\alpha_{E,F}}{h_E} & \text{if} \quad E \in \mathcal{T}_h^{ell}, \\ 0 & \text{if} \quad E \in \mathcal{T}_h^{hyp}, \end{cases} \tag{4.35}$$

where $\alpha_{E,F} := \boldsymbol{n}_{E,F}\boldsymbol{\alpha}_E\boldsymbol{n}_{E,F} > 0$ corresponds to the normal diffusivity which is positive by elliptic regularity, $\tau_0 > 0$, a given parameter determined by users, and $C_{\rm tr}$, the constant of the discrete trace inequality (4.11).

Lemma 4.4.2 (Bound of consistency term). *In addition to Definition 4.2 and for all* $\boldsymbol{v}_h \in \boldsymbol{\mathcal{V}}_h$, *the following holds :*

$$\left|\langle \boldsymbol{\alpha}\boldsymbol{\nabla}_h v_h, [\![\boldsymbol{v}_h]\!]\rangle_{0,\partial\mathcal{T}_h^{ell}}\right| \leq \frac{C_0}{2\xi}\|\boldsymbol{\alpha}^{1/2}\boldsymbol{\nabla}_h v_h\|^2_{0,\mathcal{T}_h^{ell}} + \frac{\xi}{2}|\boldsymbol{v}_h|^2_{\tau_{\boldsymbol{\alpha}}}, \tag{4.36}$$

where $C_0 := \eta_0/\tau_0$ *with* $\eta_0 := \max_{\forall E\in\mathcal{T}_h^{ell}}(\eta_E)$ *and* $\xi > 0$.

Proof. The decomposition of the consistency term yields

$$\left|\langle \boldsymbol{\alpha}\boldsymbol{\nabla}_h v_h, [\![\boldsymbol{v}_h]\!]\rangle_{0,\partial\mathcal{T}_h^{\mathrm{ell}}}\right| \leq \sum_{E\in\mathcal{T}_h^{\mathrm{ell}}}\left|\langle \boldsymbol{\alpha}^{1/2}\boldsymbol{\nabla}_h v_h, \boldsymbol{\alpha}^{1/2}[\![\boldsymbol{v}_h]\!]\rangle_{0,\partial E}\right|. \tag{4.37}$$

Successively applying the Cauchy–Schwarz inequality and the discrete trace inequality (4.11), and using the definition of $\tau_{\boldsymbol{\alpha}}$ given in (4.35), we thus infer that

$$\begin{aligned}\left|\langle \boldsymbol{\alpha}^{1/2}\boldsymbol{\nabla}_h v_h, \boldsymbol{\alpha}^{1/2}[\![\boldsymbol{v}_h]\!]\rangle_{0,\partial E}\right| &\leq \left[\frac{h_E}{\tau_0 C_{\mathrm{tr}}^2}\right]^{1/2}\|\boldsymbol{\alpha}^{1/2}\boldsymbol{\nabla}_h v_h\|_{0,\partial E}\|\tau_{\boldsymbol{\alpha}}^{1/2}[\![\boldsymbol{v}_h]\!]\|_{0,\partial E}, &(4.38)\\ &\leq \left[\frac{\eta_E}{\tau_0}\right]^{1/2}\|\boldsymbol{\alpha}^{1/2}\boldsymbol{\nabla}_h v_h\|_{0,E}|\boldsymbol{v}_h|_{\tau_{\boldsymbol{\alpha}},\partial E}, &(4.39)\end{aligned}$$

The proof is thus completed by applying Young's inequality with $\xi > 0$, summing over all mesh elements of $\mathcal{T}_h^{\mathrm{ell}}$ and observing that $\tau_{\boldsymbol{\alpha}}$ is null on $\partial\mathcal{T}_h^{\mathrm{hyp}}$ (see, Definition 4.2). ■

Lemma 4.4.3 (Discrete coercivity). *Besides the hypothesis of Assumption 4.2, then the discrete bilinear form* $\mathcal{B}_h^{(\varepsilon)}$ *in* (4.20) *is* $\boldsymbol{\mathcal{V}}_h$*-coercive with respect to the* $\|\cdot\|_\dagger$*-norm, namely,*

$$\forall \boldsymbol{v}_h \in \boldsymbol{\mathcal{V}}_h : \quad \mathcal{B}_h^{(\varepsilon)}(\boldsymbol{v}_h, \boldsymbol{v}_h) \geq C_{\mathrm{sta}}\|\boldsymbol{v}_h\|_\dagger^2, \tag{4.40}$$

where C_{sta} *is a regime-dependent positive constant, i.e.,*

$$C_{\mathrm{sta}} := \begin{cases} 1 & \text{if} \quad \Omega^{ell} = \emptyset, \\ 1 - C_0 & \text{else assuming} \quad \tau_0 > C_\varepsilon\eta_0, \end{cases}$$

with $C_\varepsilon := (1+\varepsilon)/2$.

Proof. Setting $\boldsymbol{u}_h = \boldsymbol{v}_h$ in discrete bilinear forms $\mathcal{B}_{\boldsymbol{\alpha},h}^{(\varepsilon)}$ (4.21), we directly obtain :

$$\mathcal{B}_{\boldsymbol{\alpha},h}^{(\varepsilon)}(\boldsymbol{v}_h, \boldsymbol{v}_h) := \|\boldsymbol{\alpha}^{1/2}\boldsymbol{\nabla}_h v_h\|^2_{0,\mathcal{T}_h^{\mathrm{ell}}} - (1+\varepsilon)\langle \boldsymbol{\alpha}\boldsymbol{\nabla}_h v_h, [\![\boldsymbol{v}_h]\!]\rangle_{0,\partial\mathcal{T}_h^{\mathrm{ell}}}. \tag{4.41}$$

Similarly, after integrating by parts, the advective-reactive bilinear form (4.22) yields that :

$$\mathcal{B}_{\boldsymbol{\beta},\mu,h}(\boldsymbol{v}_h, \boldsymbol{v}_h) := \|\mu_*^{1/2}v_h\|^2_{0,\mathcal{T}_h} + \|\beta^{1/2}\hat{v}_h\|^2_{0,\mathcal{F}_h^{\mathfrak{b}+}} + \underbrace{\langle \boldsymbol{\beta}v_h, [\![\boldsymbol{v}_h]\!]\rangle_{0,\partial\mathcal{T}_h} - \frac{1}{2}\langle \beta v_h, v_h\rangle_{0,\partial\mathcal{T}_h}}_{=\mathcal{T}_1}, \tag{4.42}$$

where $\mu_* := \mu + (\nabla\cdot\boldsymbol{\beta})/2 > 0$ by virtue of (4.4). Focusing now on the quantity $\mathcal{T}_1$ in (4.42) and rearranging all terms inside, one can quickly obtain :

$$\mathcal{T}_1 = \langle \beta[\![\boldsymbol{v}_h]\!], [\![\boldsymbol{v}_h]\!]\rangle_{0,\partial\mathcal{T}_h} + \underbrace{\langle \beta\hat{v}_h, v_h - \hat{v}_h\rangle_{0,\partial\mathcal{T}_h} - \frac{1}{2}\langle \beta v_h, v_h\rangle_{0,\partial\mathcal{T}_h}}_{=\mathcal{T}_2}. \tag{4.43}$$

Rearranging the two addends inside the term $\mathcal{T}_2$, we then deduce after some tedious algebraic manipulations that :

$$\mathcal{T}_2 = -\frac{1}{2}\Big[\langle\beta[\![\boldsymbol{v}_h]\!],[\![\boldsymbol{v}_h]\!]\rangle_{0,\partial\mathcal{T}_h} + \|\beta^{1/2}\hat{v}_h\|^2_{0,\mathcal{F}_h^{\mathfrak{b}+}}\Big]. \tag{4.44}$$

The equation (4.44) follows immediately by using the following identity relation :

$$\langle\beta\hat{v}_h,\hat{v}_h\rangle_{0,\partial\mathcal{T}_h} = \|\beta^{1/2}\hat{v}_h\|^2_{0,\mathcal{F}_h^{\mathfrak{b}+}}, \tag{4.45}$$

which is obtained by observing that $\hat{v}_h$ is single-valued on interfaces of the mesh skeleton $\mathcal{F}_h^+$ and is null on boundary interfaces $\mathcal{F}_h^{\mathfrak{b}-}$, and considering that $[\![\boldsymbol{\beta}]\!] = 0$ on $\mathcal{F}_h^{\mathfrak{i}}$. Thus, inserting (4.44) in (4.43), and hence assembling in (4.42), we obtain the following decomposition of the advective-reactive contribution :

$$\mathcal{B}_{\boldsymbol{\beta},\mu,h}(\boldsymbol{v}_h,\boldsymbol{v}_h) := \|\mu_*^{1/2}v_h\|^2_{0,\mathcal{T}_h} + \frac{1}{2}\|\beta^{1/2}\hat{v}_h\|^2_{0,\mathcal{F}_h^{\mathfrak{b}+}} + \frac{1}{2}\langle\beta[\![\boldsymbol{v}_h]\!],[\![\boldsymbol{v}_h]\!]\rangle_{0,\partial\mathcal{T}_h}, \tag{4.46}$$

Thus, collecting (4.41) and (4.46), inserting the contribution $\mathcal{S}_h(\boldsymbol{v}_h,\boldsymbol{v}_h)$, and finally considering Lemmata 4.4.2 and $\mu_* \geq \mu_0$ by virtue of (4.4), we immediately infer :

$$\mathcal{B}_h^{(\varepsilon)}(\boldsymbol{v}_h,\boldsymbol{v}_h) \geq C_\xi^{(\varepsilon)}\|\boldsymbol{\alpha}^{1/2}\boldsymbol{\nabla}_h v_h\|^2_{0,\mathcal{T}_h^{\mathrm{ell}}} + \|\mu_0^{1/2}v_h\|^2_{0,\mathcal{T}_h} + \frac{1}{2}\|\beta^{1/2}\hat{v}_h\|^2_{0,\mathcal{F}_h^{\mathfrak{b}+}} + \langle\tau_\xi^{(\varepsilon)}[\![\boldsymbol{v}_h]\!],[\![\boldsymbol{v}_h]\!]\rangle_{0,\partial\mathcal{T}_h}. \tag{4.47}$$

Here, $C_\xi^{(\varepsilon)}$ and $\tau_\xi^{(\varepsilon)}$ are both ε- and ξ-dependent parameters given by :

$$C_\xi^{(\varepsilon)} := 1 - C_0^{(\varepsilon)}\xi^{-1} \quad \text{and} \quad \tau_\xi^{(\varepsilon)} := \tau - \xi C_\varepsilon\tau_{\boldsymbol{\alpha}} + \beta/2, \tag{4.48}$$

where $C_\varepsilon := (1+\varepsilon)/2$ and $C_0^{(\varepsilon)} := C_0C_\varepsilon$ with C_0 as given in Lemmata 4.4.2. Considering now that $0 \leq C_\varepsilon \leq 1$ and choosing ξ such that $0 < \xi < 1$, we then infer that $\tau_\xi^{(\varepsilon)} > \tau - \tau_{\boldsymbol{\alpha}} + \beta/2 \geq \tau_{\min}$ by virtue of (4.26). We can thus distinguish two distinctive situations :

- **Fully-degenerate problems** : Here, $\Omega^{\mathrm{ell}} = \emptyset$ means that the physical process is purely advective-reactive inside $\Omega^{\mathrm{hyp}} := \Omega$. Thus, the formula (4.47) simply reduces to

$$\mathcal{B}_h^{(\varepsilon)}(\boldsymbol{v}_h,\boldsymbol{v}_h) \geq \|\mu_0^{1/2}v_h\|^2_{0,\mathcal{T}_h} + \frac{1}{2}\|\beta^{1/2}\hat{v}_h\|^2_{0,\mathcal{F}_h^{\mathfrak{b}+}} + |\boldsymbol{v}_h|^2_{\tau_{\min}} \geq \|\boldsymbol{v}_h\|^2_\dagger. \tag{4.49}$$

- **Non- or locally-degenerate problems** : Here, $\Omega^{\mathrm{ell}} \neq \emptyset$, and we emphasize that the penalty parameter τ_0 must be selected large enough, i.e., $\tau_0 > C_\varepsilon\eta_0$, to ensure that $C_\xi^{(\varepsilon)} > 0$. Hence, observing also that $C_\xi^{(\varepsilon)} \leq 1$, we thus infer

$$\begin{aligned}\mathcal{B}_h^{(\varepsilon)}(\boldsymbol{v}_h,\boldsymbol{v}_h) &\geq C_\xi^{(\varepsilon)}\Big[\|\boldsymbol{\alpha}^{1/2}\boldsymbol{\nabla}_h v_h\|^2_{0,\mathcal{T}_h^{\mathrm{ell}}} + \|\mu_0^{1/2}v_h\|^2_{0,\mathcal{T}_h} + \frac{1}{2}\|\beta^{1/2}\hat{v}_h\|^2_{0,\mathcal{F}_h^{\mathfrak{b}+}} + |\boldsymbol{v}_h|^2_{\tau_{\min}}\Big],\\ &\geq C_\dagger\|\boldsymbol{v}_h\|^2_\dagger,\end{aligned} \tag{4.50}$$

  where $C_\dagger$ corresponds to the infimum of $C_\xi^{(\varepsilon)}$; i.e.,

$$C_\dagger := \inf_{\substack{0<\xi<1\\ \varepsilon\in\{0,\pm1\}}} C_\xi^{(\varepsilon)} = 1 - C_0,$$

  with C_0 a positive constant dependent of the element shape only.

The proof is thus completed by collecting (4.49) and (4.50). ■

Thus, a straightforward consequence of the strong consistency (Lemma 4.4.1), and the discrete coercivity stability (Lemma 4.4.3) via the Lax–Milgram Theorem is the well-posedness of discrete problem (4.19) ; i.e., the existence and uniqueness of the discrete solution $\boldsymbol{u}_h \in \boldsymbol{\mathcal{V}}_h$ are ensured. Let us point out that our objective here is just to prove a stability result without deriving *a priori* error estimates. Later, the estimated convergence rates will be measured *a posteriori* based on our numerical experiments for all mentioned regimes.

Remark 4.8 (Inf-sup stability analysis). Let us also emphasize that the energy-norm (4.33) offers no bound on the advective derivative of the error, mainly when the diffusivity takes a tiny value in the elliptic region or vanishes locally in the hyperbolic one. This shortcoming can be remedied, and the convergence analysis can be improved by analyzing an Inf-Sup stability condition instead to prove a convergence result in a richer norm on $\boldsymbol{\mathcal{V}}_h$ denoted by $\|\cdot\|_{\ddagger}$ and defined as follows ; i.e., for all $\boldsymbol{v}_h \in \boldsymbol{\mathcal{V}}_h$,

$$\|\boldsymbol{v}_h\|_{\ddagger}^2 := \|\boldsymbol{v}_h\|_{\dagger}^2 + \sum_{E\in\mathcal{T}_h} h_E \|\boldsymbol{\beta}\cdot\boldsymbol{\nabla}_h v_h\|_{0,E}^2. \tag{4.51}$$

Tightened error estimates based on Inf-Sup stability analysis have been proposed by different authors for advection-dominated or pure-advective problems. We refer interested readers to the following references (see, e.g., [Ern 2009, Section 4, p.10] or [Wells 2011]).

4.4.4 Adaptive stabilization strategy

In practice, the choice of τ is quite delicate, as it strongly affects the accuracy of the H-IP method (4.19). Indeed, its definition directly impacts the numerical flux approximations $(\hat{u}_h, \hat{\boldsymbol{\sigma}}_h)$ on $\mathcal{F}_h^+$. To prove its relevance, let us apply a continuity argument, i.e., $[\![\hat{\boldsymbol{\sigma}}_h(\boldsymbol{u}_h)]\!] = 0$, on an interior interface $F \in \mathcal{F}_h^{\mathrm{i}}$. We refer the interested reader to Remark 4.10 for a discussion about degenerate outflow boundary interfaces. Thus, for any two adjacent cells $E_1, E_2 \in \mathcal{T}_h$, 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 $F := \partial E_1 \cap \partial E_2$,

$$\begin{cases} \hat{u}_h = \{\!\{u_h\}\!\}_{\boldsymbol{\omega}} + \kappa [\![\boldsymbol{\sigma}_h]\!], & \text{(4.52a)} \\ \hat{\boldsymbol{\sigma}}_h = \{\!\{\boldsymbol{\sigma}_h\}\!\}^{\boldsymbol{\omega}} + \eta [\![u_h]\!], & \text{(4.52b)} \end{cases}$$

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

$$\boldsymbol{\omega} := \left(\frac{\tau_1}{\tau_1+\tau_2}, \frac{\tau_2}{\tau_1+\tau_2}\right), \quad \kappa := \frac{1}{\tau_1+\tau_2}, \quad \text{and} \quad \eta := \frac{\tau_1\tau_2}{\tau_1+\tau_2}, \tag{4.53}$$

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

4.4.4.1 Hyperbolic regime

We assume here that $\tau_{\boldsymbol{\alpha}} := 0$, and hence $\tau^{\text{hyp}} := \tau_{\boldsymbol{\beta}}$ by virtue of (4.25). To ensure the minimal requirement (4.26) in the context of the hyperbolic regime, the advective penalty parameter is chosen such that $\tau^{\text{hyp}} > |\boldsymbol{\beta}\cdot\boldsymbol{n}|/2$ on the mesh skeleton. Furthermore, it is well known that an arbitrary choice of τ^{hyp} 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 that consists of adjusting the amount of artificial viscosity by including the normal advective effects. For this purpose, we now consider the following definition of τ^{hyp}.

Definition 4.3 (Penalty function on $\boldsymbol{\partial\mathcal{T}_h^{\text{hyp}}}$)**.** *The stabilization penalty function τ^{hyp} inside the hyperbolic region $\mathcal{T}_h^{hyp}$ is chosen to be of the following form,*

$$\tau^{\text{hyp}} := \begin{cases} \theta\,|\boldsymbol{\beta}\cdot\boldsymbol{n}| & \text{if} \quad \boldsymbol{\beta}\cdot\boldsymbol{n} \neq 0, \\ \theta\,\|\boldsymbol{\beta}\| & \text{else.} \end{cases} \tag{4.54}$$

where θ is a positive constant on the mesh skeleton $\partial\mathcal{T}_h^{hyp}$, which is determined by users and be chosen as follows :

$$\theta > \begin{cases} 1/2 & \text{if} \quad \boldsymbol{\beta}\cdot\boldsymbol{n} \neq 0, \\ 0 & \text{else,} \end{cases}$$

to fulfill the minimal requirement (A2) given in Assumption 4.2.

Remark 4.9 (A Lax–Friedrichs correction). The stabilization strategy inside the hyperbolic subregion as considered in Definition 4.3 offers attractive advantages since it does not require any assumptions concerning the orientation of the mesh interfaces. Indeed, to ensure that the hypothesis (A2) given in Assumption 4.2 is always fulfilled, a traditional upwinding-based strategy is preferred at non-aligned interfaces with $\boldsymbol{\beta}$, and another based strategy in the spirit of a Lax–Friedrichs scheme is favored at aligned interfaces. This correction is numerically performed in the experiment described in Section 4.5.2.

We now introduce the *signum function*, which is given as follows : for all $x \in \mathbb{R}$ then,

$$\text{sgn}(x) := \frac{x}{|x|} = \begin{cases} -1 & \text{if} \quad x < 0, \\ 0 & \text{if} \quad x = 0, \\ +1 & \text{if} \quad x > 0. \end{cases} \tag{4.55}$$

We then set another equivalent definition of the signum function - i.e., $\text{sgn}(x) \equiv {}^{|x|}/_{x}$.

Proposition 4.1 (Numerical fluxes)**.** *Let us now consider a non-aligned interface F with $\boldsymbol{\beta}$ such that $F := \partial E_1 \cap \partial E_2$, where E_1 and $E_2 \in \mathcal{T}_h^{hyp}$ are two adjacent cells, then the numerical fluxes $(\hat{u}_h^\theta, \hat{\boldsymbol{\sigma}}_h^\theta)$ on F are given by :*

$$\hat{u}_h^\theta := \{\!\{u_h\}\!\}_{\boldsymbol{\omega}_\theta} \quad \text{and} \quad \hat{\boldsymbol{\sigma}}_h^\theta := \boldsymbol{\beta}\{\!\{u_h\}\!\} + \frac{\tau_{\boldsymbol{\beta}}^\theta}{2}[\![u_h]\!], \tag{4.56}$$

where $\boldsymbol{\omega}_\theta := \left(\frac{1}{2} + \frac{\mathrm{sgn}(\boldsymbol{\beta}\cdot\boldsymbol{n}_1)}{2\theta}, \frac{1}{2} + \frac{\mathrm{sgn}(\boldsymbol{\beta}\cdot\boldsymbol{n}_2)}{2\theta}\right)$, $\tau_{\boldsymbol{\beta}}^{\theta} := \theta\,|\boldsymbol{\beta}\cdot\boldsymbol{n}_i|$, *and* $\boldsymbol{n}_i := \boldsymbol{n}_{E_i,F}$.

Proof. The proof is evident by substituting the definition (4.54) in (4.52) and observing that $\boldsymbol{\alpha}$ is null in $\mathcal{T}_h^{\mathrm{hyp}}$, that $[\![\boldsymbol{\beta}]\!] = 0$ on $\mathcal{F}_h^{\mathrm{i}}$, and that τ^{hyp} is single-valued on the mesh skeleton. ■

Example 4.1 (Traditional schemes). By appropriately selecting the value of θ in (4.54), we can establish bridges with some well-known stabilization schemes.

- **Upwind-scheme.** By setting $\theta = 1$, we then recover the usual definition of upwinding fluxes denoted by $(\hat{u}_h^{\mathrm{up}}, \hat{\boldsymbol{\sigma}}_h^{\mathrm{up}})$,

$$\hat{u}_h^{\mathrm{up}} := \begin{cases} u_i & \text{if} \quad \boldsymbol{\beta}\cdot\boldsymbol{n}_i > 0, \\ \{\!\{u_h\}\!\} & \text{if} \quad \boldsymbol{\beta}\cdot\boldsymbol{n}_i = 0, \end{cases} \quad \text{and} \quad \hat{\boldsymbol{\sigma}}_h^{\mathrm{up}} := \boldsymbol{\beta}\hat{u}_h^{\mathrm{up}}. \tag{4.57}$$

  The corresponding upwind penalty parameter is denoted $\tau^{\mathrm{up}} := |\boldsymbol{\beta}\cdot\boldsymbol{n}|$.

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

$$\hat{u}_h^{\mathrm{c}} := \lim_{\theta\to+\infty} \{\!\{u_h\}\!\}_{\boldsymbol{\omega}_\theta} = \{\!\{u_h\}\!\} \quad \text{and} \quad \hat{\boldsymbol{\sigma}}_h^{\mathrm{c}} := \boldsymbol{\beta}\{\!\{u_h\}\!\} + \frac{\tau_{\boldsymbol{\beta}}^{\infty}}{2}[\![u_h]\!]. \tag{4.58}$$

  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 $\tau_{\boldsymbol{\beta}}^{\infty} = \lim_{\theta\to+\infty}\tau_{\boldsymbol{\beta}}^{\theta} = +\infty$. 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 and/or significant artificial numerical diffusion near internal layers (see, e.g., the numerical experiments discussed in Section 4.5.2).

Remark 4.10 (Degenerate outflow boundaries). Let us precisely denote the transmission conditions at degenerative outflow boundaries that belong to the hyperbolic region. For all $F \in \mathcal{F}_h^{\mathfrak{b}+}$, we denote by $E \in \mathcal{T}_h^{\mathrm{hyp}}$ such that $F \in \mathcal{F}_E$, and we impose that $\hat{\boldsymbol{\sigma}}_h(\boldsymbol{u}_h)\cdot\boldsymbol{n} = (\boldsymbol{\beta}\cdot\boldsymbol{n})\hat{u}_h$, where $\hat{\boldsymbol{\sigma}}_h(\boldsymbol{u}_h) := \boldsymbol{\beta}u_h + \tau_{\boldsymbol{\beta}}[\![\boldsymbol{u}_h]\!]$. By combining these expressions, we observe that the role of $\tau_{\boldsymbol{\beta}}$ on outflow boundaries is clearly insignificant, since for any finite value of $\tau_{\boldsymbol{\beta}} > 0$, we then obtain $\hat{u}_h|_F = u_h|_{E,F}$.

4.4.4.2 Elliptic regime

We assume here that the diffusivity is not degenerate, i.e., $\underline{\alpha} > 0$. To ensure the minimal requirement (4.26) in the context of the elliptic regime, the stabilization penalty function is chosen such that $\tau^{\mathrm{ell}} \geq \tau_{\boldsymbol{\alpha}} + |\boldsymbol{\beta}\cdot\boldsymbol{n}|/2$ on $\partial\mathcal{T}_h^{\mathrm{ell}}$. In practice, its definition is generally prescribed accordingly to the predominant mechanism, i.e., diffusion vs. advection. Toward this aim, we now propose a Péclet-dependent definition of τ^{ell} to characterize the regime process inside the elliptic region.

Definition 4.4 (Penalty function on $\boldsymbol{\partial\mathcal{T}_h^{\mathrm{ell}}}$)**.** *The stabilization penalty function* τ^{ell} *inside the elliptic region* $\mathcal{T}_h^{ell}$ *is chosen of the following form,*

$$\tau^{\mathrm{ell}} := \tau_{\boldsymbol{\alpha}}\,\mathsf{A}(\mathrm{Pe}) \quad on \quad \partial\mathcal{T}_h^{ell}, \tag{4.59}$$

where $\mathsf{A}(\cdot)$ *is an appropriate single-valued function as described in Proposition 4.2. Here,* Pe *corresponds to the Péclet number and is given by :*

$$\mathrm{Pe} := \frac{\boldsymbol{\beta}\cdot\boldsymbol{n}}{\tau_{\boldsymbol{\alpha}}} \quad on \quad \partial\mathcal{T}_h^{ell}. \tag{4.60}$$

We emphasize that $\mathrm{Pe} \in L^2(\partial\mathcal{T}_h)$*, meaning that* $\mathrm{Pe}_{E_i,F}$ *can be distinctively evaluated on both sides of an interior interface* $F := \partial E_1 \cap \partial E_2$ *with* E_1 *and* E_2 *belonging to* $\mathcal{T}_h^{ell}$*.*

Remark 4.11 (Péclet-based stabilization strategy). The Péclet-based strategy is particularly appropriate or quantifying the preponderance of both diffusive vs. advective mechanisms in an automated fashion. Concretely, the regime is considered (*i*) diffusion-dominated if $|\mathrm{Pe}| \sim 1$, and (*ii*) advection-dominated if $|\mathrm{Pe}| \sim +\infty$. Furthermore, the oriented quantity (4.60) also indicates the windflow sense, i.e., for all $E \in \mathcal{T}_h^{\mathrm{ell}}$ and all $F \in \mathcal{F}_E$, then $\mathrm{Pe}_{E,F} > 0$ (resp. $\mathrm{Pe}_{E,F} < 0$) if F is an outflow- (resp. inflow-) interface.

We now specify some characteristics of the single-valued function A.

Proposition 4.2 (Characterization of A)**.** *The function* $\mathsf{A} : \mathbb{R} \to \mathbb{R}_*^+$ *must fulfill the following requirements :*

(P1) A *must be an even convex function, i.e.,* $\forall s \in \mathbb{R}$ *then* $\mathsf{A}(-s) = \mathsf{A}(s)$ *respecting* $\mathsf{A}(0) = 1$*. The latter condition allows recovery of the definition of the penalty term in the pure diffusive limit, i.e., if* $\mathrm{Pe} = 0$ *on* $\partial\mathcal{T}_h^{ell}$*, then* $\tau^{\mathrm{ell}} := \tau_{\boldsymbol{\alpha}}$*, which is in accordance with* (4.25) *as described in Assumption 4.2.*

(P2) The function A *must respect the following asymptotic behaviors : for all* $s \in \mathbb{R}$*, then*

$$\lim_{|s|\to+\infty} \frac{\mathsf{A}(s)}{s} = \mathrm{sgn}(s), \tag{4.61}$$

which is also in accordance with the parity requirement (P1) and (4.25) *in Assumption 4.2. The condition* (4.61) *allows recovery of the traditional upwind scheme as discussed in Remark 4.1 for hyperbolic problems, i.e.,*

$$\tau^{\mathrm{ell}} := \boldsymbol{\beta}\cdot\boldsymbol{n} \lim_{|\mathrm{Pe}|\to+\infty} \frac{\mathsf{A}(\mathrm{Pe})}{\mathrm{Pe}} = |\boldsymbol{\beta}\cdot\boldsymbol{n}| \equiv \tau^{\mathrm{up}}, \tag{4.62}$$

since $\mathrm{sgn}(\mathrm{Pe}) = \mathrm{sgn}(\boldsymbol{\beta}\cdot\boldsymbol{n})$*.*

(P3) For all $s \in \mathbb{R}$*, then the following holds :*

$$\mathsf{A}(s) \geq 1 + \frac{|s|}{2}. \tag{4.63}$$

This last constraint is an equivalent reformulation of hypothesis (A2) given in Assumption 4.2 to ensure that the energy-norm (4.33) *is well-defined and* $\mathcal{B}_h^{(\varepsilon)}$ *is* $\boldsymbol{\mathcal{V}}_h$*-coercive.*

Hence, we consider two well-known distinctive formulations of A introduced previously in the literature. The first example is undoubtedly the most intuitive one.

Example 4.2 (Standard-Upwind scheme). It is thus obtained by setting $\mathsf{A}_{\mathrm{su}}(s) := 1 + |s|$. Following this definition, we easily infer that

$$\tau^{\mathrm{ell}} := \tau_{\boldsymbol{\alpha}}(1 + |\mathrm{Pe}|) = \tau_{\boldsymbol{\alpha}} + \tau_{\boldsymbol{\beta}} \quad \text{on} \quad \partial\mathcal{T}_h^{\mathrm{ell}}, \tag{4.64}$$

where $\tau_{\boldsymbol{\beta}} := |\boldsymbol{\beta} \cdot \boldsymbol{n}|$ corresponds to the advective stability parameter chosen in accordance with the traditional upwind strategy. Here, we assume that τ^{ell} is simply equal to the sum of its distinctive diffusive and advective contributions. It is also evident that A_{su} respects all of the above criteria (P1-P3).

The second example was initially proposed by D.L. Scharfetter and H.K. Gummel in 1969 for the numerical approximation of the 1D drift-diffusion model [Scharfetter 1969].

Example 4.3 (Scharfetter–Gummel scheme). By now setting $\mathsf{A}_{\mathrm{sg}}(s) := \mathsf{B}_{\mathrm{e}}(-|s|)$ where $\mathsf{B}_{\mathrm{e}}(\cdot)$ denotes the Bernoulli function given by,

$$\mathsf{B}_{\mathrm{e}}(s) := \begin{cases} \dfrac{s}{e^s - 1} & \text{if} \quad s \neq 0, \\ 1 & \text{else.} \end{cases} \tag{4.65}$$

Proof of Example 4.3. We now verify that the function A_{sg} as introduced in Example 4.3 respects all requirements listed in Proposition 4.2.

- The proof of P1 is evident.
- The proof of P2 follows by observing that

$$\lim_{|s|\to+\infty} \frac{\mathsf{A}_{\mathrm{sg}}(s)}{s} = \lim_{|s|\to+\infty} \frac{|s|}{s}(1 - e^{-|s|})^{-1} = \mathrm{sgn}(s) \tag{4.66}$$

  since $\lim_{|s|\to+\infty} e^{-|s|} = 0$.
- The proof of P3 follows immediately by noticing that the following bound holds, i.e.,

$$1 + \frac{|s|}{2} \leq \mathsf{B}_{\mathrm{e}}(-|s|) \leq 1 + |s|\,, \quad \forall s \in \mathbb{R}, \tag{4.67}$$

  as depicted in Figure 4.1.

■

In practice, the Scharfetter–Gummel scheme is particularly interesting since it introduces less artificial diffusion than the standard-upwind scheme for all considered regimes. Indeed, for all $s \in \mathbb{R}$, we observe that $\mathsf{A}_{\mathrm{su}}(s/2) \leq \mathsf{A}_{\mathrm{sg}}(s) \leq \mathsf{A}_{\mathrm{su}}(s)$ by virtue of (4.67), as depicted below in Figure 4.1. Following this argument, we will always favor the Scharfetter–Gummel scheme in our upcoming numerical experiments described in Section 4.5.

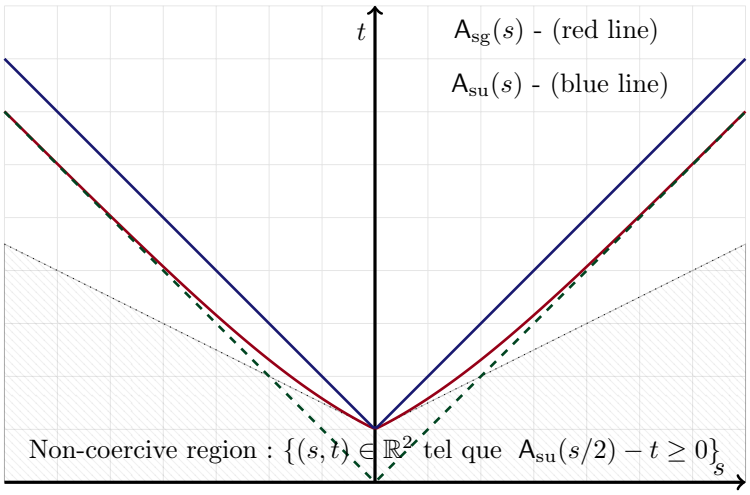


FIGURE 4.1 – An illustration of both stabilization functions $\mathsf{A}_{\mathrm{su}}(s)$ (blue line) and $\mathsf{A}_{\mathrm{sg}}(s)$ (red line). The dashed region corresponds to the set of non-admissible values of $\mathsf{A}(s)$ in the sense of P3 as given in (4.63). The dashed green line corresponds to the asymptotic limit in the pure-advective regime.

4.5 Numerical results

In this section, we provide some numerical experimental illustrating the robustness and accuracy of the proposed H-IP method for solving degenerate advection-diffusion-reaction model problems. To achieve this, we focus on three distinctive physical situations, namely, nondegenerate, fully degenerate and locally degenerate problems. All of our numerical experiments are performed using the high-performance finite element library called NGSOLVE containing a rich Python interface [Schöberl 2014]. All developed source codes (Python files) are available for free download from the following Github repository [3] and can be easily used by the interested reader to duplicate the numerical experiments presented below.

4.5.1 Test A - Nondegenerate problem

We assume here that both $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are nonnull for the whole domain $\Omega := [0,1]^2$. In this first example, 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 $\boldsymbol{\beta}$ and the dispersion matrix $\boldsymbol{\alpha}$ are assumed to be constant on the domain Ω – i.e., $\boldsymbol{\beta} := (\beta_1, \beta_2)$ with $\beta_{1,2} \in \mathbb{R}$ and $\boldsymbol{\alpha} := \alpha \boldsymbol{I}_2$, where $\boldsymbol{I}_d$ denotes the identity matrix in $\mathbb{R}^{d\times d}$ and $\alpha > 0$. The exact solution is given by

$$u(x,y) = \sin(\pi x)\sin(\pi y),$$

and the right-hand-side f is chosen such that the exact solution u is verified. Since the diffusive part is always nonnull inside Ω, we can distinguish three distinctive variants of the H-IP method. Hence, we focus on the H-IIP, H-NIP, and H-SIP schemes coupled with the Scharfetter-Gummel (SG) penalization strategy as described in Example 4.3. For our

3. https://github.com/GregoryETANGSALE/HDG-Degenerate-ADR-equation

numerical study, we set $\boldsymbol{\beta} := (2, 1)$, $\mu := 1$, and we select a wide range of $\alpha := \{1, 10^{-5}\}$ to control the ratio between the diffusive and advective contributions. Specifically, we analyze the influence of variations of the Péclet number over several orders of magnitude on the behavior of discrete solutions. Toward this aim, standard h- and k-refinement strategies are used to compute the discrete H^1- and L^2-errors and their estimated convergence rates (ECRs) for both diffusion- and advection-dominated regimes, respectively. The histories of convergence in the H^1- and L^2-norms are depicted in Figures 4.2 and 4.3 for different regimes $\alpha := \{1, 10^{-5}\}$ and polynomial degrees $k := \{1, \dots, 4\}$. A brief analysis first indicates that we can distinguish two distinctive situations.

- For the diffusion-dominated regime characterized by $\alpha = 1$ (see, e.g., Figure 4.2), we recover some well-know estimates indicating that the convergence in the discrete H^1-norm is optimal (with order k for all k) and for all schemes. However, the convergence in the L^2-norm indicates that both the H-IIP and H-NIP schemes behave differently from the H-SIP scheme. Indeed, their ECRs are strongly impacted by the polynomial parity of k. We observe that the convergence rate is suboptimal (with order k) 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. For this regime, the measured ECRs in the L^2-norm are mainly governed by the diffusive contribution. These observations agree with the theoretical results established by Shin *et al.* in [Shin 2015] (see e.g., [Etangsale 2022, Fabien 2020b] for the pure diffusive problem).
- For the advection-dominated regime characterized by $\alpha = 10^{-5}$ (see, e.g., Figure 4.3), we observe that the convergence is optimal in both the discrete H^1-norm with order k and in the L^2-norm with order $k + 1$ for all polynomial degrees $k \in \{1, \dots, 4\}$ and all schemes. For this regime, the measured ECRs in the L^2-norm are mainly governed by the advective contribution, and its diffusive part can be considered to be negligible. These observations also agree with the theoretical results of Shin *et al.* in [Shin 2015].

4.5.2 Test B - Fully degenerate problem

In the second example, we analyze the behavior of the H-IP method in the context of hyperbolic problems. Let us note that the discrete bilinear operator as defined in (4.19) is reduced to its advective-reactive part since $\boldsymbol{\alpha}$ is null throughout the whole domain $\Omega = [0, 1]^2$. Thus, we set $\boldsymbol{\beta} := (2, 1)$, $\mu := 1$, and $f := 0$. The following exact solution is obtained :

$$u(x, y) = \mathsf{H_e}(-x + 2y - 1) \quad \text{for all} \quad (x, y) \in \Omega, \tag{4.68}$$

where $\mathsf{H_e}(\cdot)$ denotes the well-known Heaviside function. Dirichlet boundary conditions are imposed (only) for the degenerate inflow part Γ^- – i.e., the left and bottom boundaries of Ω (see, e.g., Figure 4.4a). The exact solution displays a discontinuity along the characteristic direction $\boldsymbol{\beta}$ due to the jump in imposed boundary conditions at $x = 0$ and $y = 1/2$. We first assume that the actual location of internal layers is unknown, and an adaptive mesh refinement strategy based on the residual L^2–error calculation is investigated to capture them. We underline that no *a posteriori* error estimator is required here since the exact solution is known, simplifying the adaptive refinement strategy significantly as described

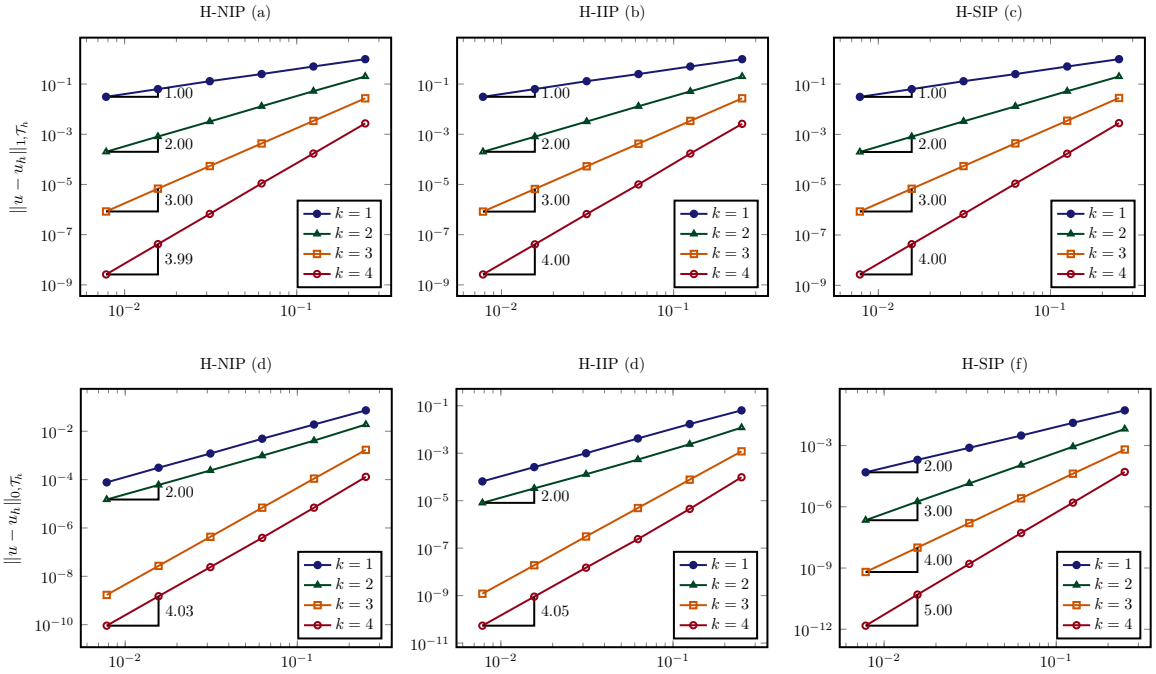


FIGURE 4.2 – Test A - History of convergence error in the discrete H^1- and L^2-norms (vs $1/h$) of the H-NIP (a,d), H-IIP (b,e) and H-SIP (c,f) methods for the diffusion-dominated regime ($\kappa = 1$).

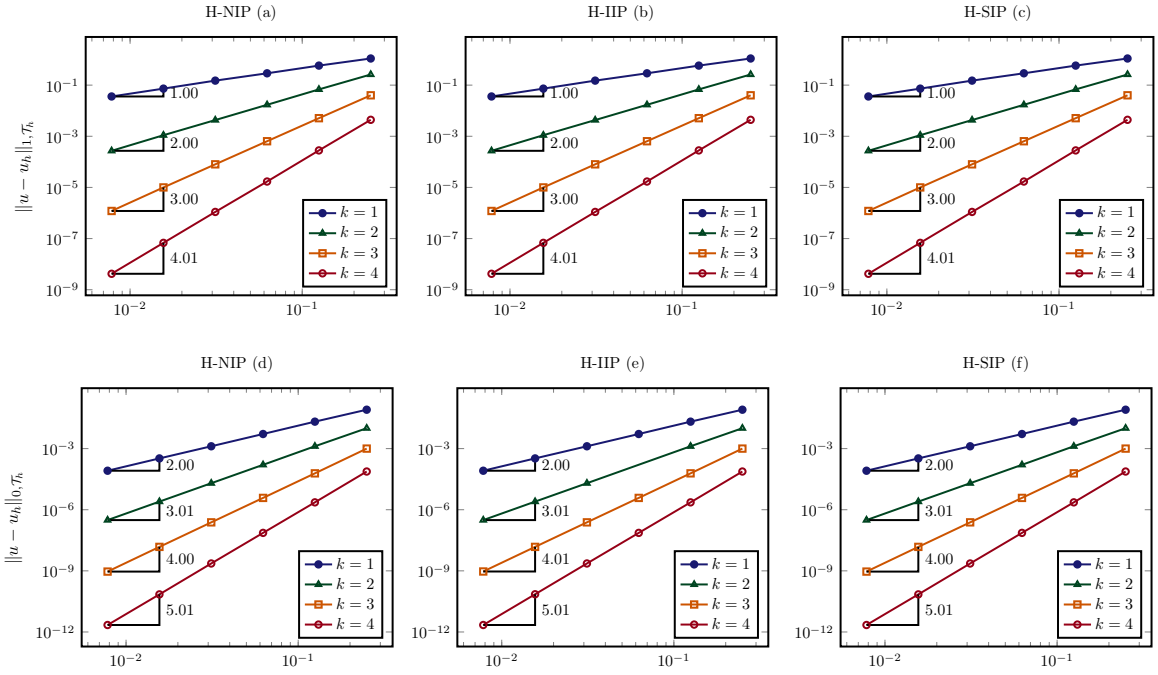
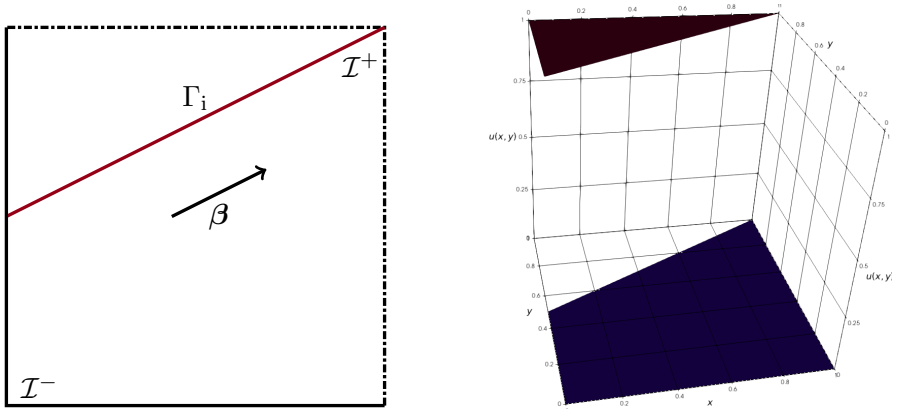


FIGURE 4.3 – Test A - History of convergence error in the discrete H^1- and L^2-norms (vs $1/h$) of the H-NIP (a,d), H-IIP (b,e) and H-SIP (c,f) methods for the advection-dominated regime ($\kappa = 10^{-5}$).



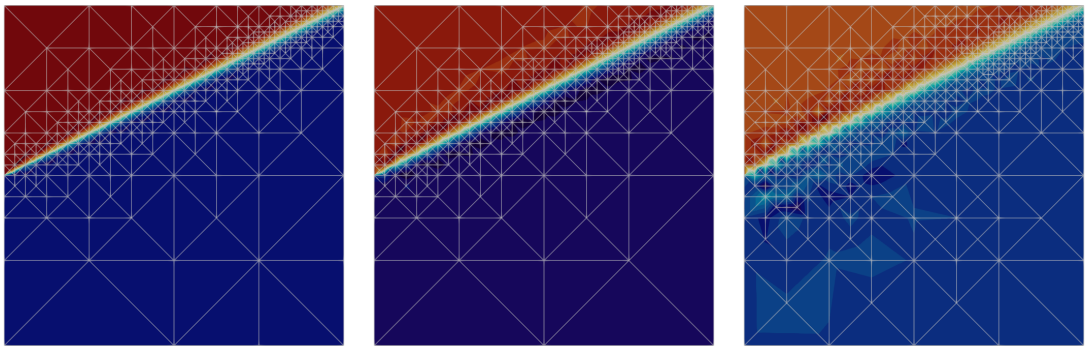
(a) (b)

FIGURE 4.4 – Test B : (a) Description of the fully degenerate case. The continuous black line denotes the inflow boundary Γ^- where Dirichlet conditions are imposed, the dashed line indicates the outflow boundary Γ^+, and the red line represents the internal layer Δ. (b) Illustration of the exact solution.

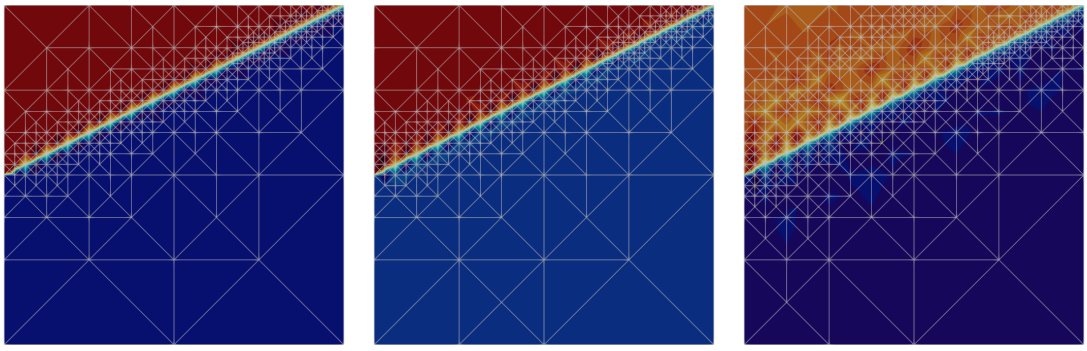
Algorithm 1 Adaptive mesh refinement

Require: $\mathcal{T}_h^0, u$ ▷ Initial coarse mesh \ Exact solution
  $k \leftarrow 0$
  Tol is a tolerance criteria
  ϱ is a threshold parameter given in percentage
  COMPUTE $\boldsymbol{u}_h^k := (u_h^k, \hat{u}_h^k)$ BY SOLVING (4.19) ON THE CURRENT MESH $\mathcal{T}_h^k$
  while $\|u - u_h^k\|_{0,\mathcal{T}_h^k} \leq Tol$ **do** ▷ Arbitrary stopping criteria
    for $E \in \mathcal{T}_h^k$ **do**
      $\varepsilon_E \leftarrow \|u - u_h^k\|_{0,E}$ ▷ Compute discrete L^2–error at mesh element
    end for
    $\varepsilon_0 \leftarrow \max_{\forall E \mathcal{T}_h^k} (\varepsilon_E)^{1/2}$
    for $E \in \mathcal{T}_h^k$ **do**
      if $\varepsilon_E \geq \varrho \varepsilon_0$ **then**
        MARK E FOR REFINEMENT
      end if
    end for
    APPLY CONFORMAL REFINEMENT STRATEGY TO GET $\mathcal{T}_h^{k+1}$
    COMPUTE $\boldsymbol{u}_h^{k+1} := (u_h^{k+1}, \hat{u}_h^{k+1})$ BY SOLVING (4.19) ON THE REFINED MESH $\mathcal{T}_h^{k+1}$
    Set $k \leftarrow k + 1$
  end while

below in Algorithm 1. We also refer the interested reader to the source code available from the repository and entitled FULLY_DEGENERATE.PY using another stopping criterion based on the cardinality of $\mathcal{T}_h$. We precise that the resulting mesh skeletons are (globally) not aligned with the characteristic direction $\boldsymbol{\beta}$. Here, we analyze here the role of the upwind-parameter θ (see, e.g., Definition 4.3) with respect to the accuracy of discrete solutions for different



(a) $k = 1$ and $\theta = 1$ (b) $k = 1$ and $\theta = 10$ (c) $k = 1$ and $\theta = 100$

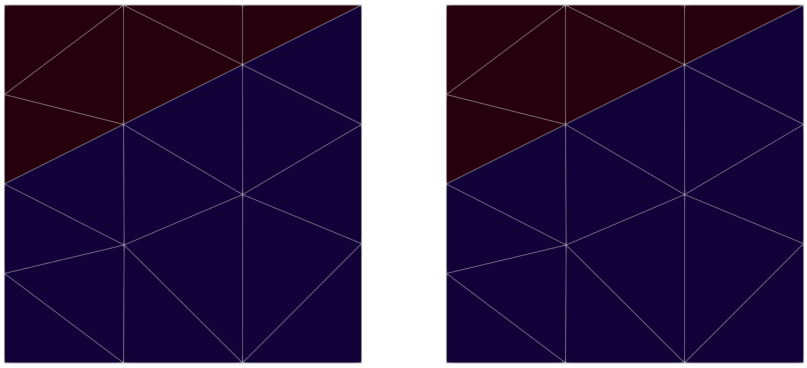


(d) $k = 2$ and $\theta = 1$ (e) $k = 2$ and $\theta = 10$ (f) $k = 2$ and $\theta = 100$

FIGURE 4.5 – Test B - Representation of linear (a)-(b)-(c) and quadratic (d)-(e)-(f) discrete solutions obtained on adaptive meshes for different values of the upwind-parameter $\theta \in \{1, 10, 100\}$, respectively.

polynomial degrees $k \in \{1, 2\}$. Since the exact solution (4.68) is only piecewise constant, increasing polynomial degrees can only yield better discrete approximations near internal layers. Following Figure 4.5, we first observe that the increase in θ significantly deteriorates the quality of the discrete solution near the internal layer for all polynomial degrees since there is more erratic behavior and/or significant additional artificial numerical diffusion. This observation agrees with Remark 4.1. We now assume that the location of the internal layer is clearly identified as in Figure 4.4a, and we consider a sequence of refined meshes aligned to the internal layer $\Delta : -x + 2y - 1 = 0$. Specifically, we use a Lax—Friedrichs correction as discussed in Remark 4.9 only at interfaces located along with Δ. Here, we analyze the ability of the H-IP scheme to well-capture piecewise constant exact solution and its jump discontinuity along Δ for different polynomial degree approximations $k := \{1, 2\}$. As depicted in Figure 4.6, a (piecewise) linear numerical solution is sufficient to deal with it. To conclude, let us also emphasize that the H-IP scheme handles outflow boundary conditions elegantly and naturally for all considered meshes (see, e.g., Figures 4.5 and 4.6). This observation agrees with Remark 4.10 indicating that $\hat{u}_h|_F = u_h|_{E,F}$, for all $F \in \mathcal{F}_h^{\mathfrak{b}+}$ and $F \in \mathcal{F}_E$. Accordingly, we will favor the traditional upwinding scheme obtained by selecting $\theta = 1$ in hyperbolic

regions in all future experiments.



(a) $k = 1$ (b) $k = 2$

FIGURE 4.6 – Representation of linear (a) and quadratic (b) discrete solutions obtained on a coarse mesh ($h = 1/3$) using a Lax–Friedrichs correction at interfaces along with the internal layer Δ.

4.5.3 Test C - 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 [Cockburn 2008] (see e.g., [Di Pietro 2015b]). The domain is now taken to be $\Omega := [-1,1]^2 \backslash \{x^2 + y^2 < 1/4\}$, which is divided into two disjoint subdomains Ω^{ell} and Ω^{hyp} corresponding to the elliptic and hyperbolic parts, respectively (see e.g., Figure 4.7a). Denoting by (r, ϑ) the polar coordinates (with azimuth ϑ measured in the anticlockwise sense starting from the positive x-axis) and by $\boldsymbol{e}_{\vartheta}$ the (unit) azimuthal vector, the problem coefficients are

$$\alpha := \begin{cases} \pi & \text{if } 0 < \vartheta \leq \pi, \\ 0 & \text{if } \pi < \vartheta \leq 2\pi, \end{cases} \qquad \boldsymbol{\beta} := \frac{\boldsymbol{e}_{\vartheta}}{r}, \qquad \mu := 10^{-6},$$

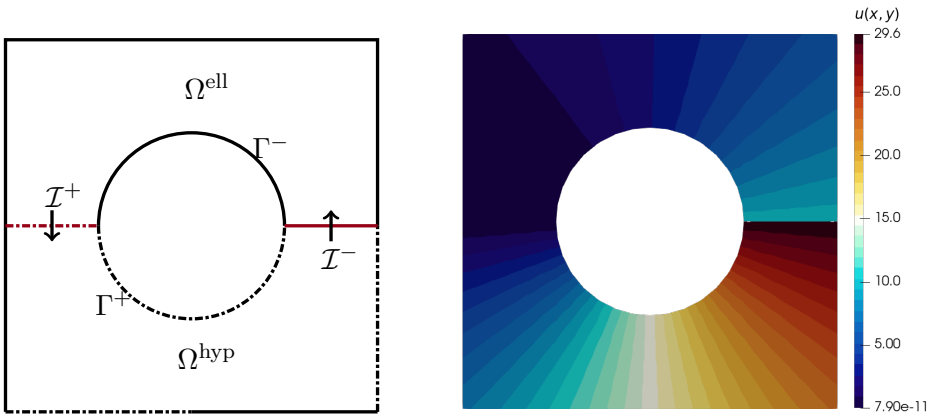
and $\boldsymbol{\alpha} := \alpha \boldsymbol{I}_2$. As illustrated in Figure 4.7b, the exact solution is given by,

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

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 strategy in the elliptic region (see, e.g., Example 4.3), and the H-IP method using the traditional upwind strategy in the hyperbolic one (see, e.g., Example 4.1). Standard h- and k-refinement strategies are used to compute the discrete L^2-errors and estimated convergence rates (ECRs). A history of convergence is illustrated in Figure 4.8a for different polynomial degrees $k \in \{1, \ldots, 4\}$. We observe that the ECR is optimal with order $k + 1$ for all k.

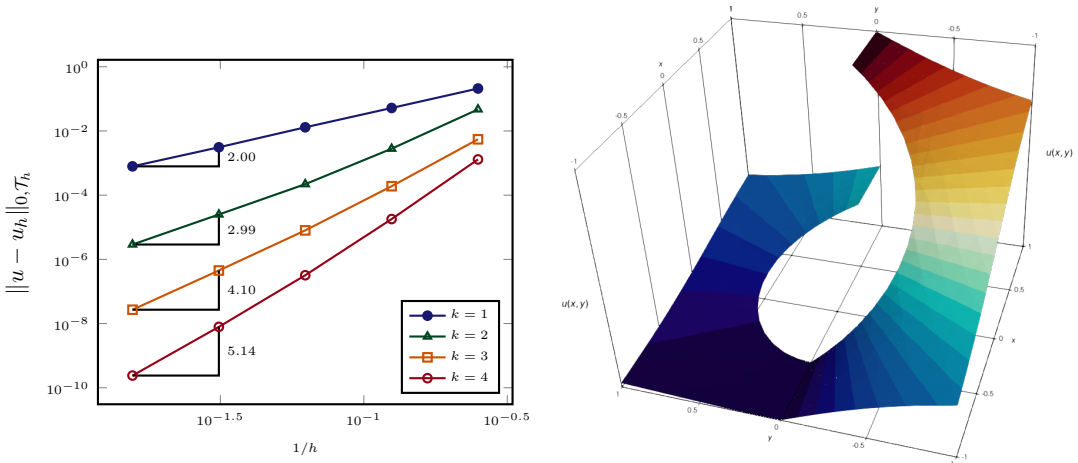
4.6 Conclusion

We have derived a compact hybridized interior penalty DG method for solving degenerate advection-diffusion-reaction problems. The proposed H-IP method can efficiently handle pure



(a) (b)

FIGURE 4.7 – Test C - (a) Description of the locally degenerate test case : the continuous black line denotes the nondegenerate inflow boundary Γ^- where Dirichlet conditions are imposed, the dashed black line indicates the degenerate no-flow/outflow boundary Γ^+, and the continuous red line represents the interior interface $\mathcal{I}^-$ where the exact solution is discontinuous. (b) Illustration of the exact solution (4.5.3).



(a) (b)

FIGURE 4.8 – Test C - (a) History of convergence error in the L_2-norm for different polynomial degrees $k \in \{1, \ldots, 4\}$. (b) Representation of u_h using piecewise linear approximations on a fine mesh ($h = 1/64$).

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. In the latter case, our discrete formulation does not require ad hoc techniques at the separative interface between elliptic/hyperbolic regions, which is a significant numerical improvement. An adaptive stabilization strategy is carried out, automatically accounting for the mathematical nature of (4.7) and the predominance of the diffusion vs. advection mechanisms. An upwinding-based scheme was favored for the hyperbolic region using a Lax–Friedrichs correction at aligned interfaces along with $\boldsymbol{\beta}$, and an inspired Scharfetter–Gummel

scheme was preferred for the elliptic region. 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 evidence.

Chapitre 5

Coupling of conforming and nonconforming Galerkin methods for fractured groundwater flow problems

---

Pre-printed article to be submitted in *Computer Physics & Communications*, under the title
On the coupling of conforming and nonconforming primal high-order Galerkin methods for fractured groundwater flow problems

Grégory Etangsale [1], Marwan Fahs [2], Vincent Fontaine[1], Hussein Hoteit [3]

Abstract

The present paper focuses on numerical modeling groundwater flows in fractured porous media using the codimensional model. In this description, fractures are defined explicitly as a $(d-1)$-dimensional geometric object immersed in a d-dimensional region and can act arbitrarily as a drain or a barrier. We investigate a novel numerical strategy combining distinct classes of conforming and nonconforming high-order Galerkin methods, which are both eligible for static condensation. This element-level procedure is particularly beneficial since it always leads to a smaller and sparser final system with degrees of freedom located solely on the mesh skeleton. Precisely, we combine an inspired hybridizable interior penalty discontinuous Galerkin (H-IP) formulation in the bulk region and a standard continuous Galerkin (CG) approximation in the fracture network. The distinctive discretization of corresponding PDEs is exposed, and the local matrix assemblies are detailed. Numerical experiments are also considered to prove the flexibility and efficiency of the discretization strategy even in realistic 2D and 3D test cases, and comparisons with the commercial software Comsol are also studied.

1. Department of Building and Environmental Sciences, University of La Réunion - South Campus, France
2. Institut Terre et Environnement de Strasbourg, Université de Strasbourg, CNRS, ENGEES, UMR 7063, Strasbourg, France
3. Physical Science and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia

Contents

5.1 Introduction

Modeling fluid flow in fractured porous media have received tremendous attention and development from a broad scientific community in recent decades [Flemisch 2016]. Novel theoretical and computational advancements directly impacted several engineering and environmental applications, such as groundwater resource management [Gläser 2019, Koohbor 2020b], petroleum extraction [Hoteit 2005, Khoei 2016], and nuclear waste storage, to name just a few. Fractures strongly affect fluid flow since they can operate as a conduit, i.e., a preferential paths with a high permeability material, or barriers with low permeability [Martin 2005, Angot 2009, Boon 2018]. Thus, modeling interactions between fractures and the bulk region is quite challenging.

Equivalent or dual porosity models, simulate flow in fractured porous media with equivalent parameters. They do not consider fractures explicitly and they cannot simulate preferential flows [Koohbor 2020b]. With Discrete Fracture Models the fractures are considered

explicitly. They are more representatitve of flow processes than equivalent models, but they are more consuming in computational requirements [Koohbor 2020b]. Recently, with the development on new computational techniques, Discrete Fracture Models are receiving increasing attention.

In practice, we can explicitly retrieve two distinctive approaches for representing fractures. Firstly, the *equidimensional* approach, that considers the bulk region and fractures with the same space dimensionality [Salinas 2018]. This description is straightforward, and does not need any coupling conditions to simulate the exchange between fractures and the bulk region. However, due to the small aperture of fractures, it requires locally refined meshes at the fracture layer. This drastically increases the number of mesh elements and hence the computational burden. Secondly, the *codimensional* model, which describes the fracture as objects of lower dimension by one order than the domain. This description is also referred to as the Discrete Fracture-Matrix (DFM) model [Martin 2005] (see, e.g., [Angot 2009, Fumagalli 2012, Tatomir 2012]. Thus, this description offers manifold advantages since (*i*) it leads to a simplified mesh geometry of the fracture network, (*ii*) it improves the mesh quality, and (*iii*) it drastically reduces the number of coupled degrees of freedom (DOFs) compared to the equidimensional approach.

The codimensional approach is an excellent alternative for modeling fluid flow in fractured domains, however it requires specific numerical schemes to couple $(d-1)$ and (d)–dimensional flow equations. The vast literature dedicated to the development of numerical schemes for the codimensional approach is very diversified. Among the various existing numerical approaches, we can retrieve those in which the mesh partition of the bulk is aligned and compliant with the fracture network. In this context, [Martin 2005] successfully derive a a (low-order) Mixed Finite Element (MFE) methods for modeling both the conductive and blocking fractures for flow in porous media (see e.g., [Boon 2018, Budiša 2021]). Angot *et al.* [Angot 2009]employed a cell-centered finite volume schemes, and a control-volume distributed multi-point flux approximation [Ahmed 2015] for the codimensional setting. The Mimetic Finite Differences method inside the bulk is coupled with the finite volume discretization inside the fractures [Formaggia 2018]. A standard Discontinuous Galerkin formalisms coupled with the Continuous Galerkin approach (DG/CG) was developed in [Antonietti 2019, Ma 2021, Zhao 2022]. Kadeethum and coworkers [Kadeethum 2020] have recently compared a novel combined Enriched Galerkin/DG method to the prior DG/CG scheme. A fully DG/DG strategy has been investigated for addressing flow in fractured porous media in Antonietti *et al.* [Antonietti 2020] and Mozolevski *et al.* [Mozolevski 2021]. Recently Younes *et al.* [Younes 2023] combined the MFE-DG methods for modeling variable density flow in fractured porous media. Several authors have also explored original numerical strategies for nonconforming grids in which the bulk region and fractures are meshed independently. Tunc *et al.* [Tunc 2012], and D'Angelo *et al.* [D'Angelo 2012] extended the cell-centered finite volume discretisation and MFE method for the non-matching grids. Several authors also considered the Embedded-DFM finite element (FEM) approach [Odsæter 2019, Vasilyeva 2019, Nikitin 2020]. The extended finite element method [Fumagalli 2012, Flemisch 2016] fulfills the idea of the unfitted FEM elements (see also the developments of the FEM with a local enrichment of the basis functions [Watanabe 2012, Manzoli 2021, Camargo 2022]). Mortar-type methods connect the flow along the matrix-fracture interface as a new variable [Nordbotten 2019], and the methods

based-on Lagrange multipliers functions follow the same idea [Köppel 2019a, Köppel 2019b, Zulian 2022]. A comparative study of different discretization approaches, based on a series of 2D and 3D benchmarks, can be found in Flemish *et al.* [Flemisch 2018], and Berre *et al.* [Berre 2021].

In the present work, we consider the approach of aligned grid. We focus on the famous class of discontinuous skeletal methods for modeling single-phase flow in fractured porous media. They belong to the popular category of DG methods eligible for static condensation. To that aim, an additional discrete variable is introduced on the boundary of the elements to approximate the state variable with Lagrange multipliers [Nguyen 2009, Egger 2010, Wells 2011, Cockburn 2012c, Fabien 2020b]. Thus, interior-based DOFs can be locally eliminated by static condensation in favor of those located on the mesh skeleton. This procedure is now well-established, and we refer the interested reader to these following works for a complete description [Nguyen 2009, Lehrenfeld 2010, Kirby 2012]. To the authors' knowledge, only [Chave 2019] have recently proposed an inspired hybrid high-order discontinuous skeletal method for groundwater flow in fractured porous media. They then extend their formalism to the mass transport problem [Chave 2018, Chave 2019]. Here, we combine a Hybridizable Interior Penalty (H-IP) discontinuous Galerkin method [Etangsale 2022] in the bulk-region with the standard Conforming Galerkin (CG) approach in the fracture. The H-IP method can be regarded as the Hybrid version of the famous Interior Penalty DG (IPDG) scheme, which have been analyzed recently by many authors [Arnold 1982, Rivière 2008, Di Pietro 2011]. The choice of the HDG framework is motivated by the discontinuous nature of the approximation of the state variable along with the matrix-fracture interface. Although, this class is particularly relevant for the codimensional description since their DOFs are located on the mesh skeleton, ensuring the natural contribution of the fractures between the mesh elements. We analyze the benefits of the HDG methods by reformulating the static condensation procedure in the codimensional description. The use of the reduced system might be of interest in the context of highly fractured porous media. We shall notice that the final matrix system is quasi-comparable to the one suggested in the standard HDG framework. Using an appropriate stabilization function inspired by [Etangsale 2022], the coupled H-IP/CG formalism can capture significant variations of the heterogeneity and anisotropy ratios of the bulk. The performances of the novel discretization strategy are supported by a series of numerical experiments. We also investigated the flexibility of the coupled H-IP/CG method to handle complex geometries, including intersecting and immersed fractures. Comparisons with the standard finite element solutions obtained with COMSOL and existing benchmarks aer performed to investigate the robustness of the combined H-IP/CG scheme in challenging and realistic physical situations.

The material is organized as follows. The mathematical model problem and the geometry are described in Section 5.2. We then derive, in Section 5.3, the discrete coupled H-IP/CG formalism and briefly discuss some implementation aspects relative to the static condensation procedure. In the last Section 5.4, extensive numerical experiments in two- and three-dimensional space are presented to corroborate our assertion regarding stability, accuracy, and efficiency. We end with some concluding remarks and perspectives.

5.2 Continuous model problem

In this section, we briefly review the codimensional model, the so-called Discrete-Fracture-Matrix (DFM) model in the literature, governing the groundwater flow in fractured porous media. In this framework, fractures are described as interfaces of codimension one, i.e., $(d-1)$-dimensional geometric objects immersed in a d-dimensional bulk domain. This *reasonable* assumption is justified by the fact that the thickness of the aperture is negligible compared to the size of the physical domain. Our presentation is strongly inspired by the seminal work of [Martin 2005]. Firstly, we will precise some notations related to the domain geometry characterization. Then, we will remind distinctive governing equations in the bulk-fracture regions and coupling conditions to close the physical problem. To alleviate the exposition, we will consider the simplest case, namely the *single* fractured porous media. However, the extension to a collection of single fractures is straightforward.

5.2.1 Geometry notation and assumptions

Let Ω be a bounded polygonal/polyhedral domain in $\mathbb{R}^d$, $d = 2, 3$, with a boundary $\partial\Omega \subset \mathbb{R}^{d-1}$. For simplicity, we assume that the porous medium is fulfilled by a compressible fluid and is crossed by a single fracture γ described) as an interface of codimension one, i.e., $\gamma \subset \mathbb{R}^{d-1}$, which separates Ω into exactly two disjoint subdomains Ω_i (see Figure 5.1-b). Let us now characterize the bulk region and its boundary by Ω_{B} and $\partial\Omega_{\mathrm{B}}$, respectively :

$$\Omega_{\mathrm{B}} := \Omega \backslash \gamma \quad \text{and} \quad \partial\Omega_{\mathrm{B}} := \partial\Omega \backslash \partial\gamma, \tag{5.1}$$

where $\partial\gamma := \gamma \cap \partial\Omega$ denotes the external frontier of the corresponding fracture. Without loss of generality, we further assume that (i) γ is an interface with no curvature, namely, a line segment if $d = 2$ and a planar surface if $d = 3$, and (ii) the fracture thickness is not necessary a single constant but it can be a scalar-valued function, i.e., $\ell_\gamma : \gamma \to \mathbb{R}$. We then introduce our primal unknowns of the groundwater flow problem, namely continuous pressure fields evaluated distinctively inside the bulk region $p_{\mathrm{B}} : \Omega_{\mathrm{B}} \to \mathbb{R}$ and the fracture $p_\gamma : \gamma \to \mathbb{R}$. Since they are both governed by distinctive equations, we further assume that these quantities can be discontinuous at bulk-fracture interface, i.e.,

$$p_{\mathrm{B}}(\boldsymbol{x}) - p_\gamma(\boldsymbol{x}) \neq 0, \quad \forall \boldsymbol{x} \in \gamma. \tag{5.2}$$

This assumption is crucial to account for fractures acting as drains or barriers. As we will see in the next section, the pressure-jump condition (5.2) directly impacts the flux mass exchange between the bulk region and fracture.

Remark 5.1 (Discontinuous pressure-jump at γ). Following our notations and as illustrated in Figure 5.1, let us emphasize that the pressure-jump given in relation (5.2) can be distinctively evaluated on the two branches of γ, i.e.,

$$p_{\mathrm{B}}(\boldsymbol{x}) - p_\gamma(\boldsymbol{x}) = \begin{cases} p_1(\boldsymbol{x}) - p_\gamma(\boldsymbol{x}) & \forall \boldsymbol{x} \in \partial\Omega_1 \cap \gamma, \\ p_2(\boldsymbol{x}) - p_\gamma(\boldsymbol{x}) & \forall \boldsymbol{x} \in \partial\Omega_2 \cap \gamma. \end{cases} \tag{5.3}$$

This characteristic results from the discontinuous nature of the bulk-pressure p_{B} on both sides of γ.

Finally, we introduce useful trace operators defined at the fracture interface γ. For a (regular enough) scalar-valued function φ, we then define the jump and average of φ at γ as follows :

$$[\![\varphi]\!]_\gamma := \varphi_1 \boldsymbol{n}_1 + \varphi_2 \boldsymbol{n}_2, \quad \text{and} \quad \{\!\{\varphi\}\!\}_\gamma := \frac{1}{2}(\varphi_1 + \varphi_2), \tag{5.4}$$

where φ_i denotes the restriction of φ to Ω_i and $\boldsymbol{n}_i$, the unit normal vector to the fracture γ pointing out of Ω_i. The generalization of these definitions to the set of (smooth enough) vector-valued functions $\boldsymbol{\varphi}$ is straightforward.

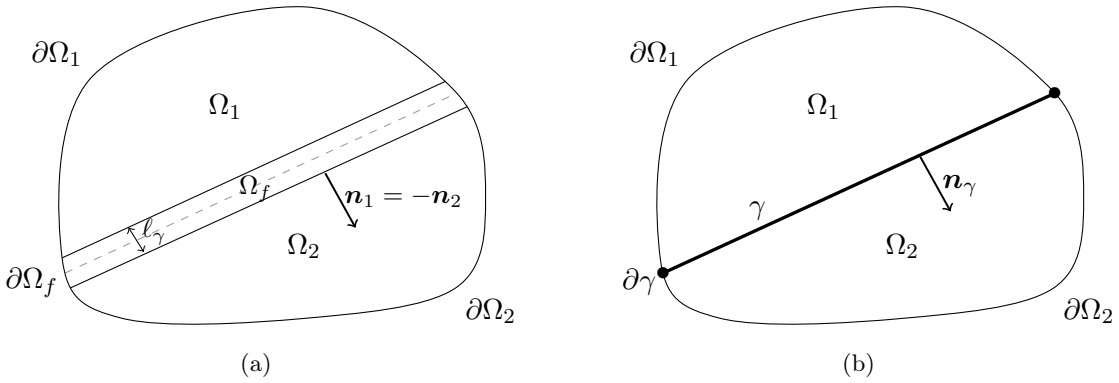


FIGURE 5.1 – Description of the equidimensional (a) and the codimensional description (b) for a single fractured porous media, with the boundaries $\partial\gamma$ and $\partial\Omega_j$, $j = \{1, 2, f\}$. The reduced interface $\gamma \in \mathbb{R}^{d-1}$ corresponds to the central axis of the fracture domain $\Omega_f = \{\boldsymbol{x} \in \mathbb{R}^d \mid \boldsymbol{x} = \boldsymbol{s} + r\boldsymbol{n}_\gamma,\ \boldsymbol{s} \in \gamma,\ r \in (-\ell_\gamma/2, \ell_\gamma/2)\}$.

5.2.2 The coupled bulk-fracture flow problem

Here, we describe the motion of a compressible fluid in bulk and the fracture, distinctively. Owing to the hypothesis of compressibility, we will assume the transient character of the global physical process and denote by $\mathrm{T} := [0, T]$ the time-interval with $T > 0$. For clarity, we then use the following superscript notation to characterize a space-time domain. Let $\mathcal{D}$ be an arbitrary bounded physical domain in $\mathbb{R}^p$ with $p \geq 1$, then we set $\mathcal{D}^{\mathrm{T}} := \mathcal{D} \times \mathrm{T}$. For a complete and rigorous derivation of the codimensional model, we refer the interested readers to [Martin 2005].

- **Bulk region :** For the simplicity of the exposition, we then assume that the flow is driven by the standard Darcy's law in the bulk region using homogeneous Dirichlet boundary conditions :

$$\boldsymbol{\sigma}_{\mathrm{B}} = -\boldsymbol{\alpha}_{\mathrm{B}} \boldsymbol{\nabla} p_{\mathrm{B}} \qquad \text{in } \Omega_{\mathrm{B}}^{\mathrm{T}}, \tag{5.5a}$$

$$\phi_{\mathrm{B}} \partial_t p_{\mathrm{B}} + \boldsymbol{\nabla} \cdot \boldsymbol{\sigma}_{\mathrm{B}} = f_{\mathrm{B}} \qquad \text{in } \Omega_{\mathrm{B}}^{\mathrm{T}}, \tag{5.5b}$$

$$p_{\mathrm{B}} = p_{\mathrm{B}}^0 \qquad \text{in } \Omega_{\mathrm{B}}^0, \tag{5.5c}$$

$$p_{\mathrm{B}} = 0 \qquad \text{on } \partial\Omega_{\mathrm{B}}^{\mathrm{T}}. \tag{5.5d}$$

Here, $\boldsymbol{\sigma}_{\mathrm{B}} : \Omega_{\mathrm{B}}^{\mathrm{T}} \to \mathbb{R}^d$ corresponds to the Darcy velocity of the fluid in the bulk domain, $\boldsymbol{\alpha}_{\mathrm{B}} : \Omega_{\mathrm{B}}^{\mathrm{T}} \to \mathbb{R}^{d \times d}$, an anisotropic and heterogeneous permeability tensor that we assume

to be (at least) symmetric and positive definite, $f_{\mathrm{B}} : \Omega_{\mathrm{B}}^{\mathrm{T}} \to \mathbb{R}$, a given source term, $p_{\mathrm{B}}^0 : \Omega_{\mathrm{B}} \to \mathbb{R}$, an initial condition, and $\phi_{\mathrm{B}} : \Omega_{\mathrm{B}}^{\mathrm{T}} \to \mathbb{R}$, a given storage coefficient. Thus, we shall assume that all these physical parameters are regular enough.

- **Fracture region :** The strategy for establishing the reduced flow model inside the fracture consists of using averaging techniques by integrating the standard d-dimensional Darcy's model along the normal direction of γ. Thus, let us first characterize the fracture permeability tensor $\boldsymbol{\alpha}_\gamma$ along its normal and tangential directions at the fracture level. We then assume that it has the following block-diagonal structure :

$$\boldsymbol{\alpha}_\gamma := \begin{bmatrix} \alpha_\gamma^{\boldsymbol{n}} & 0 \\ 0 & \boldsymbol{\alpha}_\gamma^{\boldsymbol{\tau}} \end{bmatrix}, \tag{5.6}$$

where $\alpha_\gamma^{\boldsymbol{n}} > 0$ corresponds to its normal component, and $\boldsymbol{\alpha}_\gamma^{\boldsymbol{\tau}} : \gamma^{\mathrm{T}} \to \mathbb{R}^{(d-1)\times(d-1)}$, its tangential one assuming here to be a symmetric and positive definite tensor-valued function satisfying the same regularity requirements as $\boldsymbol{\alpha}_{\mathrm{B}}$. We also point out here that $\boldsymbol{\alpha}_\gamma^{\boldsymbol{\tau}}$ simply reduces to a positive coefficient for $d = 2$. As a result, the flow is now governed by a (modified) tangential Darcy's law inside the fracture interface of the following form :

$$\begin{aligned}
\boldsymbol{\sigma}_\gamma &= -\ell_\gamma \boldsymbol{\alpha}_\gamma^{\boldsymbol{\tau}} \boldsymbol{\nabla}_{\boldsymbol{\tau}} p_\gamma && \text{in } \gamma^{\mathrm{T}}, && (5.7a)\\
\phi_\gamma \partial_t p_\gamma + \boldsymbol{\nabla}_{\boldsymbol{\tau}} \cdot \boldsymbol{\sigma}_\gamma &= f_\gamma + [\![\boldsymbol{\sigma}_{\mathrm{B}}]\!]_\gamma && \text{in } \gamma^{\mathrm{T}}, && (5.7b)\\
p_\gamma &= p_\gamma^0 && \text{in } \gamma^0, && (5.7c)\\
p_\gamma &= 0 && \text{on } \partial\gamma^{\mathrm{T}}, && (5.7d)
\end{aligned}$$

where $\boldsymbol{\sigma}_\gamma : \gamma^{\mathrm{T}} \to \mathbb{R}^{d-1}$ corresponds to the tangential Darcy velocity inside the fracture, $f_\gamma : \gamma^{\mathrm{T}} \to \mathbb{R}$, a given source term, $p_\gamma^0 : \gamma \to \mathbb{R}$, an initial condition, and $\phi_\gamma : \gamma^{\mathrm{T}} \to \mathbb{R}$, a given storage coefficient that we also assume to be regular enough. Here, $\boldsymbol{\nabla}_{\boldsymbol{\tau}}$ and $\boldsymbol{\nabla}_{\boldsymbol{\tau}}\cdot$ denote, respectively, the gradient and divergence operators in the tangential direction of the fracture as defined by [Fumagalli 2012]. The jump-term appearing in equation (5.7b) corresponds to the mass flux exchange from the bulk towards the fracture flow.

- **Compatible coupling condition :** Finally, we close the bulk-fracture problem by supplementing the system (5.5-5.7) with coupling conditions related to the normal flux contribution along the fracture interface γ. Different mathematical formulations of these conditions have been derived in the literature, mainly inspired by the constraints of DG formalism. Here, we privilege a specific formulation close to a Finite Volume decomposition which is naturally compatible with the HDG formalism, as we will see in the next section. The coupling condition depends on the pressure jump at the fracture and is given in a compact form as follows :

$$\boldsymbol{\sigma}_{\mathrm{B}} \cdot \boldsymbol{n}_{\mathrm{B},\gamma} = a_\gamma (p_{\mathrm{B}} - p_\gamma) \quad \text{on } \gamma, \tag{5.8}$$

where $a_\gamma := 2\alpha_\gamma^{\boldsymbol{n}}/\ell_\gamma$ represents the effective normal reduced-permeability, and $\boldsymbol{n}_{\mathrm{B},\gamma}$, the unit normal vector to γ pointing out of Ω_{B}. Following our notations (see, e.g., Figure 5.1), by setting $\boldsymbol{n}_{\mathrm{B},\gamma}|_{\Omega_i} := \boldsymbol{n}_i$ with $i = 1, 2$ and combining (5.8), we recover a standard

DG-formulation of the coupling condition based on the use of DG-trace operators as defined in (5.4) :

$$[\![\boldsymbol{\sigma}_{\mathrm{B}}]\!]_\gamma = 2a_\gamma \left(\{\!\{p_{\mathrm{B}}\}\!\}_\gamma - p_\gamma\right) \quad \text{and} \quad \{\!\{\boldsymbol{\sigma}_{\mathrm{B}}\}\!\}_\gamma = \frac{a_\gamma}{2}[\![p_{\mathrm{B}}]\!]_\gamma. \tag{5.9}$$

A vast amount of literature has suggested a variety of averaging techniques. For instance, we refer to [Martin 2005] for a detailed description of families of the codimensional model (see also [Angot 2009]).

Remark 5.2 (Others boundary conditions). For clarity and simplicity of our upcoming discussion about discretization strategy, global matrices assembly, and the static condensation, then only homogeneous Dirichlet boundary conditions are considered here at boundaries of the physical domain $\partial\Omega_{\mathrm{B}}$ and $\partial\gamma$, respectively. However, we underline that extending to non-homogeneous Dirichlet, Neumann or Robin boundary conditions is straightforward as proposed in Section 5.4 devoted to numerical experiments.

5.3 Numerical discretization of the coupled bulk-fracture problem

Here, we present a primal-primal nonconforming/conforming discretization methods for solving the coupled bulk-fracture problem (5.5), (5.7) and (5.8). Specifically, we use a Hybridizable Interior Penalty (H-IP) discontinuous Galerkin scheme in the bulk region and a standard continuous Galerkin approach in the fracture. The interest in such discretization methods is that they are both eligible for static condensation, leading to a smaller and sparser final matrix system where coupled degrees of freedom are solely located on the mesh skeleton. Before deriving the global discrete weak formulation and its statically condensed version, we first specify our discrete notations regarding the mesh and approximation spaces.

5.3.1 Discrete setting

We start by introducing some useful mesh notations. Let $\mathcal{T}_h$ be a shape-regular family of affine conformal triangulations of Ω. We then assume that $\mathcal{T}_h$ is compliant with the fracture γ, meaning that any element $E \in \mathcal{T}_h$ must belong to the bulk region Ω_{B} and cannot be cut by γ (except along its boundaries). Thus, a generic element $E \in \mathcal{T}_h$ is assumed to be a d-simplex, i.e., a triangle for $d = 2$ and a tetrahedron for $d = 3$, and we denote by h_E and $|E|$ its diameter and measure, respectively. Let us now precise the fundamental concept of *interface* associated with the mesh skeleton assuming to be a $(d-1)$-simplex, i.e., a line segment for $d = 2$ and a triangle for $d = 3$. The set of all interfaces denoted by $\mathcal{F}_h$ can be decomposed distinctively inside the bulk and fracture as follows :

$$\mathcal{F}_h := \mathcal{F}_h^{\mathrm{B}} \cup \mathcal{F}_h^{\gamma}, \tag{5.10}$$

with $\mathcal{F}_h^{\mathrm{B}} := \mathcal{F}_h^{\mathfrak{b}} \cup \mathcal{F}_h^{\mathfrak{i}}$. Here, $\mathcal{F}_h^{\mathfrak{b}}$ denotes the set of boundary interfaces, i.e., for any $F \in \mathcal{F}_h^{\mathfrak{b}}$, then there exists a unique element $E \in \mathcal{T}_h$ such that $F := \partial E \cap \partial\Omega$. Similarly, $\mathcal{F}_h^{\gamma}$ (resp. $\mathcal{F}_h^{\mathfrak{i}}$) denotes the set of fracture (resp. interior) interfaces that (resp. not) belong to γ, i.e., for any $F \in \mathcal{F}_h^{\gamma}$ (resp. $\mathcal{F}_h^{\mathfrak{i}}$), there exists distinct elements $E_1, E_2 \in \mathcal{T}_h$ such that $F := \partial E_1 \cap \partial E_2$ and $F \cap \gamma \neq \emptyset$ (resp. $F \cap \gamma = \emptyset$). Following the compliance hypothesis, we emphasize

that (i) the subset $\mathcal{F}_h^\gamma$ is automatically induced by the partition $\mathcal{T}_h$, and (ii) each fracture interface is always surrounded by only two adjacent elements of the bulk region. The latter constraint can be relaxed, leading to nonconforming grid meshes ; however, for the clarity of the exposition, we will not retain this hypothesis. For any $E \in \mathcal{T}_h$, we denote by $\mathcal{F}_E$ the set of interfaces composing its boundary ∂E, and we denote by $\boldsymbol{n}_{E,F}$ the unit normal vector to F pointing out of E. Likewise, the collection of interfaces of all mesh elements is provided by $\partial\mathcal{T}_h := \{\cup\mathcal{F}_E,\ \forall E \in \mathcal{T}_h\}$. Following the same arguments, we then set that $\partial\mathcal{T}_h := \partial\mathcal{T}_h^{\mathrm{B}} \cup \partial\mathcal{T}_h^\gamma$ where $\partial\mathcal{T}_h^{\mathrm{B}} := \{\cup F \text{ s.t. } F \in \mathcal{F}_E \text{ and } F \in \mathcal{F}_h^{\mathrm{B}},\ \forall E \in \mathcal{T}_h\}$ is the collections of interfaces belonging to the bulk domain. We then set $\partial\mathcal{T}_h^\gamma := \partial\mathcal{T}_h \backslash \partial\mathcal{T}_h^{\mathrm{B}}$.

For a bounded polyhedral domain $\mathcal{D} \subset \mathbb{R}^d$ with $\partial\mathcal{D} \subset \mathbb{R}^{d-1}$, we denote by $H^s(\mathcal{D})$ the usual Hilbert space of index s on $\mathcal{D}$ equipped with its natural norm $\|\cdot\|_{s,\mathcal{D}}$ and seminorm $|\cdot|_{s,\mathcal{D}}$, respectively. For $s = 0$, then we set $H^0(\mathcal{D}) = L^2(\mathcal{D})$, and we denote by $(\cdot,\cdot)_{0,\mathcal{D}}$ (resp. $\langle\cdot,\cdot\rangle_{0,\partial\mathcal{D}}$) the L^2-inner product in $L^2(\mathcal{D})$ (resp. $L^2(\partial\mathcal{D})$) equipped with its natural norm $\|\cdot\|_{0,\mathcal{D}}$ (resp. $\|\cdot\|_{0,\partial\mathcal{D}}$). For any $E \in \mathcal{T}_h$, we set $\langle\cdot,\cdot\rangle_{0,\partial E} := \sum_{F\in\mathcal{F}_E}\langle\cdot,\cdot\rangle_{0,F}$, and we denote by $\|\cdot\|_{0,\partial E}$ its corresponding norm. Assuming a given partition $\mathcal{T}_h$, we then denote by $H^s(\mathcal{T}_h)$ the usual broken Sobolev space equipped with its natural scalar-product $(\cdot,\cdot)_{s,\mathcal{T}_h}$ and its corresponding broken norm $\|\cdot\|_{s,\mathcal{T}_h}$. Hence, we introduce compact notations associated with the discrete L^2-inner scalar product :

$$(\cdot,\cdot)_{0,\mathcal{T}_h} := \sum_{E\in\mathcal{T}_h}(\cdot,\cdot)_{0,E}, \quad \langle\cdot,\cdot\rangle_{0,\partial\mathcal{T}_h} := \sum_{E\in\mathcal{T}_h}\langle\cdot,\cdot\rangle_{0,\partial E} \quad \text{and} \quad \langle\cdot,\cdot\rangle_{0,\mathcal{F}_h} := \sum_{F\in\mathcal{F}_h}\langle\cdot,\cdot\rangle_{0,F}, \tag{5.11}$$

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. Let us now introduce the following spaces of piecewise polynomial functions used to approximate our discrete primal variables, i.e.,

$$\mathcal{Q}_h^{\mathrm{B}} := \{q_h^{\mathrm{B}} \in H^1(\mathcal{T}_h) :\ q_h^{\mathrm{B}}|_E \in \mathcal{P}_k(E),\ \forall E \in \mathcal{T}_h\}, \tag{5.12a}$$

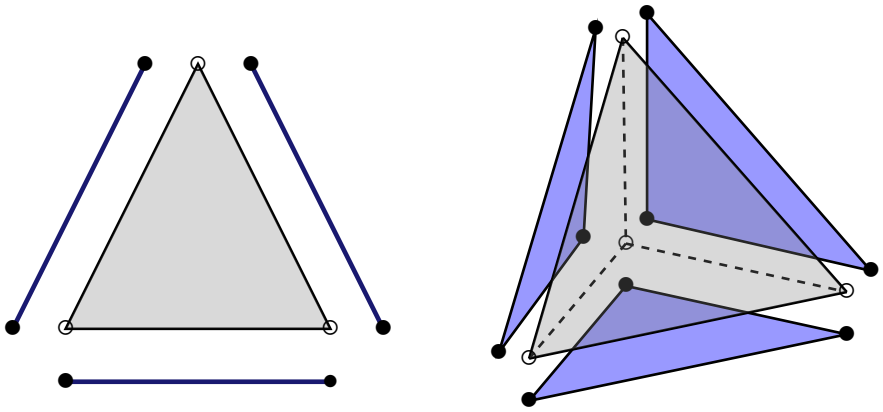
$$\hat{\mathcal{Q}}_h^{\mathrm{B}} := \{\hat{q}_h^{\mathrm{B}} \in L^2(\mathcal{F}_h^{\mathrm{B}}) :\ \hat{q}_h^{\mathrm{B}}|_F \in \mathcal{P}_k(F),\ \forall F \in \mathcal{F}_h^{\mathfrak{i}} \text{ and } \hat{q}_h^{\mathrm{B}}|_F = 0,\ \forall F \in \mathcal{F}_h^{\mathfrak{b}}\}, \tag{5.12b}$$

$$\bar{\mathcal{Q}}_h^\gamma := \{\bar{q}_h^\gamma \in L^2(\mathcal{F}_h^\gamma) \cap \mathcal{C}^0(\gamma) :\ \bar{q}_h^\gamma|_F \in \mathcal{P}_k(F),\ \forall F \in \mathcal{F}_h^\gamma\}, \tag{5.12c}$$

where $\mathcal{P}_k(X)$ denotes the space of polynomials of at least degree k on X, and X, a generic element of $\mathcal{T}_h$ or $\mathcal{F}_h$. Let us underline that the finite element space $\hat{\mathcal{Q}}_h^{\mathrm{B}}$ refers to the set of polynomial functions whose degrees of freedom are (exclusively) located on $\mathcal{F}_h^{\mathrm{B}}$, respecting homogeneous Dirichlet boundary conditions imposed on $\mathcal{F}_h^{\mathfrak{b}}$ and vanishing on $\mathcal{F}_h^\gamma$. For brevity, let us now introduce the novel compact discrete variable $\boldsymbol{q}_h := (q_h^{\mathrm{B}}, \hat{q}_h^{\mathrm{B}}, \bar{q}_h^\gamma)$ belonging to the composite approximation space $\boldsymbol{\mathcal{Q}}_h := \mathcal{Q}_h^{\mathrm{B}} \times \hat{\mathcal{Q}}_h^{\mathrm{B}} \times \bar{\mathcal{Q}}_h^\gamma$, i.e., $\boldsymbol{q}_h \in \boldsymbol{\mathcal{Q}}_h$. For convenience, let us indicate that we will use the same polynomial degree approximation k for each component of the composite variable $\boldsymbol{q}_h \in \boldsymbol{\mathcal{Q}}_h$. We are now in position to define the jump operator $[\![\cdot]\!] : \boldsymbol{\mathcal{Q}}_h \to [L^2(\partial\mathcal{T}_h)]^d$. Thus, for all $E \in \mathcal{T}_h$, all $F \in \mathcal{F}_E$, and all $\boldsymbol{q}_h \in \boldsymbol{\mathcal{Q}}_h$, then its restriction to F is explicitly given by :

$$[\![\boldsymbol{q}_h]\!]|_{E,F} := \begin{cases} (q_h^{\mathrm{B}}|_{E,F} - \hat{q}_h^{\mathrm{B}}|_F)\boldsymbol{n}_{E,F} & \text{if } F \in \partial\mathcal{T}_h^{\mathrm{B}}, \\ (q_h^{\mathrm{B}}|_{E,F} - \bar{q}_h^\gamma|_F)\boldsymbol{n}_{E,F} & \text{if } F \in \partial\mathcal{T}_h^\gamma. \end{cases} \tag{5.13}$$

Here, $L^2(\partial\mathcal{T}_h)$ denotes the space of functions that are double-valued on $\mathcal{F}_h^{\mathfrak{i}} \cup \mathcal{F}_h^\gamma$ and single-valued on $\mathcal{F}_h^{\mathfrak{b}}$. Let us point out the main difference with the usual space $L^2(\mathcal{F}_h)$, which is a subset of $L^2(\partial\mathcal{T}_h)$, i.e., $L^2(\mathcal{F}_h) \subset L^2(\partial\mathcal{T}_h)$.



(a) (b)

FIGURE 5.2 – Illustration of the DOFs for the H-IP method, with a d-simplex element, i.e., a triangle if $d = 2$ (a) and tetrahedron if $d = 3$ (b) using piecewise linear approximation. The black circle represents the trace-based unknowns and the circle the element-based unknowns.

5.3.2 The compact coupled discrete formulation

In this section, we combine primal conforming/nonconforming discretization methods for solving the coupled problem (5.5), (5.7) and (5.8). We then denote by $\boldsymbol{p}_h := (p_h^{\mathrm{B}}, \hat{p}_h^{\mathrm{B}}, \bar{p}_h^{\gamma}) \in \boldsymbol{\mathcal{Q}}_h$ the discrete approximation of the exact solution $\boldsymbol{p} := (p_{\mathrm{B}}, p_{\mathrm{B}}, p_\gamma)$. First, we derive distinctive discrete formulations following our physical intuition as much as possible and then assemble the global discrete weak formulation.

- **H-IP discretization** : Let $\boldsymbol{q}_h := (q_h^{\mathrm{B}}, \hat{q}_h^{\mathrm{B}}, \bar{q}_h^{\gamma}) \in \boldsymbol{\mathcal{Q}}_h$ be an arbitrary composite trial function. Multiplying (5.5b) by $q_h^{\mathrm{B}} \in \mathcal{Q}_h^{\mathrm{B}}$, integrating by parts over $E \in \mathcal{T}_h$, and summing over all mesh elements of $\mathcal{T}_h$, we then infer the global weak formulation :

$$(\phi_{\mathrm{B}} \partial_t p_h^{\mathrm{B}}, q_h^{\mathrm{B}})_{0,\mathcal{T}_h} - (\boldsymbol{\sigma}_h^{\mathrm{B}}(\boldsymbol{p}_h), \boldsymbol{\nabla}_h q_h^{\mathrm{B}})_{0,\mathcal{T}_h} + \langle \hat{\boldsymbol{\sigma}}_h^{\mathrm{B}}(\boldsymbol{p}_h), q_h^{\mathrm{B}} \boldsymbol{n} \rangle_{0,\partial\mathcal{T}_h} = (f_{\mathrm{B}}, q_h^{\mathrm{B}})_{0,\mathcal{T}_h}, \tag{5.14}$$

where $\boldsymbol{\sigma}_h^{\mathrm{B}}(\boldsymbol{p}_h) := -\boldsymbol{\alpha}_{\mathrm{B}} \boldsymbol{\nabla}_h p_h^{\mathrm{B}}$ corresponds to the discrete approximation of $\boldsymbol{\sigma}_{\mathrm{B}}$ on $\mathcal{T}_h$, and $\hat{\boldsymbol{\sigma}}_h^{\mathrm{B}}(\boldsymbol{p}_h)$ its trace approximation on $\partial\mathcal{T}_h$ that we will precisely define below. Let us underline that $\hat{\boldsymbol{\sigma}}_h^{\mathrm{B}}(\boldsymbol{p}_h)$ can be evaluated distinctively on both sides of any given interface. Hence, an additional equation needs to be included in (5.14) to ensure the continuity requirements. To do so, considering that $\hat{q}_h^{\mathrm{B}} \in \hat{\mathcal{Q}}_h^{\mathrm{B}}$ is single-valued on $\mathcal{F}_h^{\mathsf{i}}$ and vanishes on $\mathcal{F}_h^{\mathsf{b}}$, we can then impose the transmission condition on $\mathcal{F}_h^{\mathrm{B}}$, i.e.,

$$\langle \hat{\boldsymbol{\sigma}}_h^{\mathrm{B}}(\boldsymbol{p}_h), \hat{q}_h^{\mathrm{B}} \boldsymbol{n} \rangle_{0,\partial\mathcal{T}_h^{\mathrm{B}}} = 0. \tag{5.15}$$

Inserting (5.15) in (5.14) and supplementing with additional consistent term leads to the following compact form of the H-IP method :

$$\begin{aligned}(\phi_{\mathrm{B}} \partial_t p_h^{\mathrm{B}}, q_h^{\mathrm{B}})_{0,\mathcal{T}_h} - (\boldsymbol{\sigma}_h^{\mathrm{B}}(\boldsymbol{p}_h), \boldsymbol{\nabla}_h q_h^{\mathrm{B}})_{0,\mathcal{T}_h} + \langle \hat{\boldsymbol{\sigma}}_h^{\mathrm{B}}(\boldsymbol{p}_h), [\![\boldsymbol{q}_h]\!] \rangle_{0,\partial\mathcal{T}_h^{\mathrm{B}}} + \langle \hat{\boldsymbol{\sigma}}_h^{\mathrm{B}}(\boldsymbol{p}_h), q_h^{\mathrm{B}} \boldsymbol{n} \rangle_{0,\partial\mathcal{T}_h^{\gamma}} \\ + \varepsilon \langle \hat{\boldsymbol{\sigma}}_h^{\mathrm{B}}(\boldsymbol{q}_h), [\![\boldsymbol{p}_h]\!] \rangle_{0,\partial\mathcal{T}_h^{\mathrm{B}}} = (f, q_h^{\mathrm{B}})_{0,\mathcal{T}_h},\end{aligned} \tag{5.16}$$

where ε controls the introduction of the last (symmetry) term in (5.16) obtained by permuting the role of $\boldsymbol{p}_h$ and $\boldsymbol{q}_h$ in its third contribution. This procedure is well-established in the IP framework to restore the symmetry of discretized problem. Mainly, $\varepsilon = 0$ corresponds to the Incomplete scheme denoted as the H-IIP method, while $\varepsilon = +1$ (resp. $\varepsilon = -1$) corresponds to the Symmetric (resp. Non-symmetric) one named as the H-SIP (resp. H-NIP) method.

- **CG discretization** : The CG discretization is straightforward. Thus, multiplying (5.7b) by a H^1-conform test function $\bar{q}_h^\gamma \in \bar{\mathcal{Q}}_h^\gamma$, integrating by parts over $F \in \mathcal{F}_h^\gamma$, and summing overall element interfaces of $\mathcal{F}_h^\gamma$, we then obtain the following discrete formulation :

$$\langle \phi_\gamma \partial_t \bar{p}_h^\gamma, \bar{q}_h^\gamma \rangle_{0,\mathcal{F}_h^\gamma} - \langle \boldsymbol{\sigma}_h^\gamma(\boldsymbol{p}_h), \boldsymbol{\nabla}_{\boldsymbol{\tau},h} \bar{q}_h^\gamma \rangle_{0,\mathcal{F}_h^\gamma} - \langle \hat{\boldsymbol{\sigma}}_h^{\mathrm{B}}(\boldsymbol{p}_h), \bar{q}_h^\gamma \boldsymbol{n} \rangle_{0,\partial\mathcal{T}_h^\gamma} = \langle f_\gamma, \bar{q}_h^\gamma \rangle_{0,\mathcal{F}_h^\gamma}, \quad (5.17)$$

where $\boldsymbol{\sigma}_h^\gamma(\boldsymbol{p}_h) := -\ell_\gamma \boldsymbol{\alpha}_\gamma^{\boldsymbol{\tau}} \boldsymbol{\nabla}_{\boldsymbol{\tau},h} \bar{p}_h^\gamma$. The last term in the left-hand-side of (5.17) incorporates the fact that $\hat{\boldsymbol{\sigma}}_h^{\mathrm{B}}$ is distinctively evaluated on both side of the fracture interface.

- **Numerical flux $\hat{\boldsymbol{\sigma}}_h^{\mathrm{B}}$ on $\partial\mathcal{T}_h$** : The discrete problems (5.16) and (5.17) are then closed by assuming the following simple form of $\hat{\boldsymbol{\sigma}}_h^{\mathrm{B}}(\boldsymbol{p}_h)$ on the mesh skeleton,

$$\hat{\boldsymbol{\sigma}}_h^{\mathrm{B}}(\boldsymbol{p}_h) := \begin{cases} \boldsymbol{\sigma}_h^{\mathrm{B}}(\boldsymbol{p}_h) + \tau_{\mathrm{B}} [\![\boldsymbol{p}_h]\!] & \text{on } \partial\mathcal{T}_h^{\mathrm{B}}, \\ a_\gamma [\![\boldsymbol{p}_h]\!] & \text{on } \partial\mathcal{T}_h^\gamma. \end{cases} \quad (5.18)$$

where a_γ is defined as in (5.8) and τ_{B} denotes the stabilization penalty parameter on the mesh skeleton $\partial\mathcal{T}_h^{\mathrm{B}}$ of the bulk region. For all $E \in \mathcal{T}_h$ and all $F \in \mathcal{F}_E \cap \mathcal{F}_h^\gamma$, then τ_{B} is defined such that its restriction on F is given by :

$$\tau_{\mathrm{B}}|_{E,F} := C_{k,d} \alpha_{E,F} h_E^{-1}, \quad (5.19)$$

where $\alpha_{E,F} := \boldsymbol{n}_{E,F} \boldsymbol{\alpha}_E \boldsymbol{n}_{E,F} > 0$ corresponds to the normal permeability coefficient, and $C_{k,d} := \tau_0 (k+1)(k+d)/d$, a multiplicative constant depending on the space-dimension d, the polynomial degree k, and a user-dependent constant $\tau_0 > 0$. Let us emphasize the resemblances between the parameters a_γ and τ_{B}, namely the ratio of the normal permeability coefficient vs. a local length-scale.

Thus, inserting (5.18) in (5.17) and (5.14), we infer the compact formulation of the coupled H-IP/CG methods which consists of seeking $\boldsymbol{p}_h \in \boldsymbol{\mathcal{Q}}_h$ such that,

$$\mathcal{A}_h^{(\varepsilon)}(\boldsymbol{p}_h, \boldsymbol{q}_h) = (f_{\mathrm{B}}, q_h^{\mathrm{B}})_{0,\mathcal{T}_h} + \langle f_\gamma, \bar{q}_h^\gamma \rangle_{0,\mathcal{F}_h^\gamma}, \quad \forall \boldsymbol{q}_h \in \boldsymbol{\mathcal{Q}}_h \quad (5.20)$$

where the bilinear form $\mathcal{A}_h^{(\varepsilon)} : \boldsymbol{\mathcal{Q}}_h \times \boldsymbol{\mathcal{Q}}_h \to \mathbb{R}$ can be decomposed as follows :

$$\mathcal{A}_h^{(\varepsilon)}(\boldsymbol{p}_h, \boldsymbol{q}_h) := \mathcal{A}_h^{\mathrm{B}}(\boldsymbol{p}_h, \boldsymbol{q}_h) + \mathcal{A}_h^\gamma(\boldsymbol{p}_h, \boldsymbol{q}_h) + \mathcal{I}_h(\boldsymbol{p}_h, \boldsymbol{q}_h).$$

Here, $\mathcal{A}_h^{\mathrm{B}}$ and $\mathcal{A}_h^\gamma$ correspond respectively to the H-IP and CG discretization methods in the bulk region and the fracture, and $\mathcal{I}_h$ represents a coupling term between both sub-problems :

$$\begin{aligned} \mathcal{A}_h^{\mathrm{B}}(\boldsymbol{p}_h, \boldsymbol{q}_h) :=& (\phi_{\mathrm{B}} \partial_t p_h^{\mathrm{B}}, q_h^{\mathrm{B}})_{0,\mathcal{T}_h} + (\boldsymbol{\alpha}_{\mathrm{B}} \boldsymbol{\nabla}_h p_h^{\mathrm{B}}, \boldsymbol{\nabla}_h q_h^{\mathrm{B}})_{0,\mathcal{T}_h} + \langle \tau_{\mathrm{B}} [\![\boldsymbol{p}_h]\!], [\![\boldsymbol{q}_h]\!] \rangle_{0,\partial\mathcal{T}_h^{\mathrm{B}}} \\ & - \langle \boldsymbol{\alpha}_{\mathrm{B}} \boldsymbol{\nabla}_h p_h^{\mathrm{B}}, [\![\boldsymbol{q}_h]\!] \rangle_{0,\partial\mathcal{T}_h^{\mathrm{B}}} - \varepsilon \langle \boldsymbol{\alpha}_{\mathrm{B}} \boldsymbol{\nabla}_h q_h^{\mathrm{B}}, [\![\boldsymbol{p}_h]\!] \rangle_{0,\partial\mathcal{T}_h^{\mathrm{B}}}, \quad (5.21a) \\ \mathcal{A}_h^\gamma(\boldsymbol{p}_h, \boldsymbol{q}_h) :=& \langle \phi_\gamma \partial_t \bar{p}_h^\gamma, \bar{q}_h^\gamma \rangle_{0,\mathcal{F}_h^\gamma} + \langle \ell_\gamma \boldsymbol{\alpha}_\gamma^{\boldsymbol{\tau}} \boldsymbol{\nabla}_{\boldsymbol{\tau},h} \bar{p}_h^\gamma, \boldsymbol{\nabla}_{\boldsymbol{\tau},h} \bar{q}_h^\gamma \rangle_{0,\mathcal{F}_h^\gamma}, \quad (5.21b) \\ \mathcal{I}_h(\boldsymbol{p}_h, \boldsymbol{q}_h) :=& \langle a_\gamma [\![\boldsymbol{p}_h]\!], [\![\boldsymbol{q}_h]\!] \rangle_{0,\partial\mathcal{T}_h^\gamma}. \quad (5.21c) \end{aligned}$$

Remark 5.3 (Alternative discretization inside the fracture). Inside the fracture γ, the choice of the conforming approach is motivated by its fast implementation and eligibility for static condensation. However, we emphasize that other numerical strategies are also available such as a primal HHO method or, eventually, a classical mixed hybrid finite element (MHFE) method. The numerical approach remains globally unchanged, and the main difference relies on the mathematical construction of the bilinear form $\mathcal{A}_h^\gamma$.

5.3.3 Static condensation procedure

Typically, the coupled H-IP/CG discretization problem (5.20) for solving flow in fractured porous media yields a large-size system of differential/algebraic equations, even for a lower polynomial approximation. Indeed, the global dimension of the composite variable $\boldsymbol{p}_h \in \boldsymbol{\mathcal{Q}}_h$ is directly associated with each local component, namely, $p_h^{\mathrm{B}}, \hat{p}_h^{\mathrm{B}}$ and $\bar{p}_h^\gamma$. Different optimization techniques have been explored in the literature to decouple the set of degrees of freedom and drastically reduce the corresponding linear system's size. The static condensation procedure is undoubtedly one of these most famous techniques for which H-IP and CG methods are both eligible. The statically condensed linear system now only couples fewer degrees of freedom localized on the mesh skeleton. Here, we describe this procedure for the combined H-IP/CG discretization method (5.20) employed for solving the coupled flow model in fractured porous media. To do so, we first need to unify our notations and reorganize the set $\bar{\mathcal{Q}}_h^\gamma := \mathcal{Q}_h^\gamma \times \hat{\mathcal{Q}}_h^\gamma$. Here, $\mathcal{Q}_h^\gamma$ and $\hat{\mathcal{Q}}_h^\gamma$ represent the subset of functions of $\bar{\mathcal{Q}}_h^\gamma$ that have their support on the (strict) interior of all interfaces of $\mathcal{F}_h^\gamma$ and on its boundaries, respectively. Consequently, a generic composite element $\boldsymbol{q}_h \in \boldsymbol{\mathcal{Q}}_h$ can be equivalently reorganized as follows :

$$\boldsymbol{q}_h := (q_h^{\mathrm{B}}, q_h^\gamma, \hat{q}_h^{\mathrm{B}}, \hat{q}_h^\gamma) \in \mathcal{Q}_h^{\mathrm{B}} \times \mathcal{Q}_h^\gamma \times \hat{\mathcal{Q}}_h^{\mathrm{B}} \times \hat{\mathcal{Q}}_h^\gamma,$$

where the two-first and -last components correspond to the set of interior- and skeleton-based DOFs, respectively. Let us now consider that the discrete composite variable $\boldsymbol{p}_h(\cdot, t) \in \boldsymbol{\mathcal{Q}}_h$ can be decomposed on the set of trial functions as follows :

$$p_h^{\mathrm{B}} := \sum_i^{n_{\mathrm{B}}^{\mathfrak{i}}} p_i^{\mathrm{B}}(t) q_i^{\mathrm{B}}, \quad p_h^\gamma := \sum_i^{n_\gamma^{\mathfrak{i}}} p_i^\gamma(t) q_i^\gamma, \quad \hat{p}_h^{\mathrm{B}} := \sum_i^{n_{\mathrm{B}}^{\mathfrak{b}}} \hat{p}_i^{\mathrm{B}}(t) \hat{q}_i^{\mathrm{B}}, \quad \text{and} \quad \hat{p}_h^\gamma := \sum_i^{n_\gamma^{\mathfrak{b}}} \hat{p}_i^\gamma(t) \hat{q}_i^\gamma. \tag{5.22}$$

Here, $n_{\mathrm{B}}^{\mathfrak{i}} := \dim(\mathcal{Q}_h^{\mathrm{B}})$, $n_{\mathrm{B}}^{\mathfrak{b}} := \dim(\hat{\mathcal{Q}}_h^{\mathrm{B}})$, $n_\gamma^{\mathfrak{i}} := \dim(\mathcal{Q}_h^\gamma)$ and $n_\gamma^{\mathfrak{b}} := \dim(\hat{\mathcal{Q}}_h^\gamma)$. The rest of the statically condensed strategy consists of successively projecting (5.20) on (i) the set of interior-based DOFs, namely $\boldsymbol{q}_h := (q_h^{\mathrm{B}}, 0, 0, 0)$ and $\boldsymbol{q}_h := (0, q_h^\gamma, 0, 0)$, and hence on (ii) the set of skeleton-based DOFs, namely $\boldsymbol{q}_h := (0, 0, \hat{q}_h^{\mathrm{B}}, 0)$ and $\boldsymbol{q}_h := (0, 0, 0, \hat{q}_h^\gamma)$. Thus, the global discrete problem (5.20) can be recast in the following compact matrix-form :

$$\begin{bmatrix} \mathbb{A} & \mathbf{0} \\ \mathbf{0} & \mathbb{E} \end{bmatrix} \begin{bmatrix} \dot{\mathrm{P}}_h \\ \dot{\hat{\mathrm{P}}}_h \end{bmatrix} + \begin{bmatrix} \mathbb{B}^{(\varepsilon)} & \mathbb{C} \\ \mathbb{C}^{(\varepsilon)} & \mathbb{D} \end{bmatrix} \begin{bmatrix} \mathrm{P}_h \\ \hat{\mathrm{P}}_h \end{bmatrix} = \begin{bmatrix} \mathrm{F}_h \\ \hat{\mathrm{F}}_h \end{bmatrix}, \tag{5.23}$$

where $\mathrm{P}_h := [\mathrm{P}_h^{\mathrm{B}}, \mathrm{P}_h^\gamma]$ and $\hat{\mathrm{P}}_h := [\hat{\mathrm{P}}_h^{\mathrm{B}}, \hat{\mathrm{P}}_h^\gamma]$ corresponds to the vector of interior- and skeleton-based degrees of freedom associated with the discrete variables $(p_h^{\mathrm{B}}, p_h^\gamma)$ and $(\hat{p}_h^{\mathrm{B}}, \hat{p}_h^\gamma)$, respectively. Here, the *dot* symbol denotes the partial time-derivative operator. Let us point out

that each block-matrix of (5.23) can be decomposed as follows :

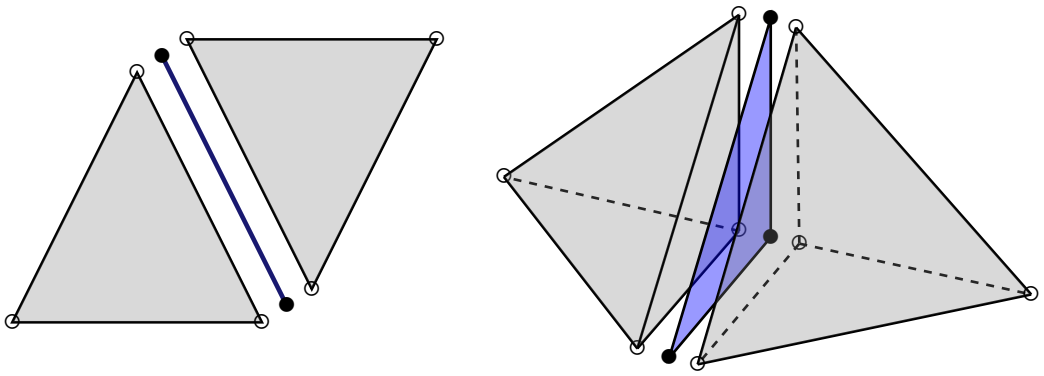
$$\left[\begin{array}{cc|cc} \mathbb{A}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbb{A}_2 & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbb{E}_2 \end{array}\right] \left[\begin{array}{c} \dot{\mathrm{P}}_h^{\mathrm{B}} \\ \dot{\mathrm{P}}_h^{\gamma} \\ \hline \dot{\hat{\mathrm{P}}}_h^{\mathrm{B}} \\ \dot{\hat{\mathrm{P}}}_h^{\gamma} \end{array}\right] + \left[\begin{array}{cc|cc} \mathbb{B}_1^{(\varepsilon)} & \mathbb{I}_1 & \mathbb{C}_1 & \mathbb{I}_2 \\ \mathbb{I}_1^t & \mathbb{B}_2 & \mathbf{0} & \mathbb{I}_3 \\ \hline \mathbb{C}_2^{(\varepsilon)} & \mathbf{0} & \mathbb{D}_1 & \mathbf{0} \\ \mathbb{I}_2^t & \mathbb{I}_3^t & \mathbf{0} & \mathbb{D}_2 \end{array}\right] \left[\begin{array}{c} \mathrm{P}_h^{\mathrm{B}} \\ \mathrm{P}_h^{\gamma} \\ \hline \hat{\mathrm{P}}_h^{\mathrm{B}} \\ \hat{\mathrm{P}}_h^{\gamma} \end{array}\right] = \left[\begin{array}{c} \mathrm{F}_h^{\mathrm{B}} \\ \mathrm{F}_h^{\gamma} \\ \hline \mathbf{0} \\ \hat{\mathrm{F}}_h^{\gamma} \end{array}\right]. \tag{5.24}$$

For conciseness of the presentation, the mathematical expressions of each subblock matrices given in (5.24) are reported in Appendix A1 (section 5.6.1). We also emphasize that all matrix assemblies can be performed at the element level due to the discontinuous nature of the discrete approximation spaces $\mathcal{Q}_h^{\mathrm{B}}$ and $\mathcal{Q}_h^{\gamma}$. In practice, the time derivative operator is substituted using an arbitrary time discretization scheme. The resulting final system to solve is of the following form : at each time step $\Delta t := t^n - t^{n-1}$,

$$\mathbb{K}\hat{\mathrm{P}}_h^n = \mathbb{F}^{n-1}, \tag{5.25}$$

where $\mathbb{K} \in \mathbb{R}^{n_{\mathfrak{b}} \times n_{\mathfrak{b}}}$, with $n_{\mathfrak{b}} = n_{\mathrm{B}}^{\mathfrak{b}} + n_{\gamma}^{\mathfrak{b}}$, corresponds to the final mass matrix, $\hat{\mathrm{P}}_h^n \in \mathbb{R}^{n_{\mathfrak{b}}}$, the vector of skeleton-based unknowns at the time-level t^n, and $\mathbb{F}^{n-1} \in \mathbb{R}^{n_{\mathfrak{b}}}$, a given right-hand-side. For instance, we refer the interested reader to Appendix A2 (section 5.6.2) describing the combination of an Euler (first-order) time derivative and the static condensation procedure.

Remark 5.4 (Influence of the static condensation). For the coupled H-IP/CG approach, the static condensation technique affects considerably the size of the matrix system. By applying this procedure inside the bulk region, it was shown that the number of unknowns and the CPU time can be reduced by half compared to the traditional DG schemes (see [Etangsale 2021b, Section 3.3]). Inside the fracture domain, the static condensation of the CG method is relevant since the interface-based unknowns are located on the nodes ($d = 2$) or the edges ($d = 3$) of the fracture elements (see Figure 5.3 for a geometric representation of the DOFs.



(a) (b)

FIGURE 5.3 – Exemple of a $(d-1)$-fracture element (in blue) aligned with two d-dimension neighbouring elements (in gray) in 2D (c) and 3D (d). The black dots represents the interface-based unknowns and the circle the interior-based unknowns.

5.4 Numerical experiments

In this section, numerical experiments are conducted to illustrate the performance of the newly developped H-IP/CG formalism for modeling flow in fractured porous media. The numerical experiments are designed to progressively increase the complexity of the fracture network. To validate the new formalism, we evaluate its accuracy and ability to handle complex geometries, providing some comparisons with finite element solutions obtained with COMSOL Multiphysics. We investigate the robustness of H-IP/CG approach for treating discontinuous and anisotropic diffusion coefficients inside the bulk region. The flexibility and the efficiency of this approach is studied for 2D and 3D groundwater flow at large scale. The H-IP/CG method is implementated using the high-performance finite element library, called NGSolve. For the test cases presented below, the results of the H-IP/CG approach are compared to a finite element (FE) solution obtained with COMSOL using quadratic continuous Lagrange functions. The COMSOL model is developed by combining the built-in modules *Subsurface Flow* for fluid flow inside the bulk region, *Fracture Flow* to deal with the presence of conductive fractures, and *Thin Barriers* to treat blocking interfaces. Finally, we further assume that both the COMSOL and the H-IP/CG solutions employ the same grid for the resolution.

5.4.1 Validation

To validate the suitability of the H-IP/CG sheme, we realize a series of numerical experiments including (*i*) a single fracture, (*ii*) a quarter five-spot pattern with a regular fracture network, and (*iii*) a production example with crossing barriers.

5.4.1.1 Test A - Single fracture

The first test case is designed to evaluate the estimated convergence rates (ECRs) of the H-IP/CG method in the simplest setting, i.e., a single fractured porous media. The configuration of the problem is inspired by the one proposed in [Antonietti 2020], selecting a particular choice of the average procedure of the coupling conditions (see also [Chave 2018, Antonietti 2019]). Let us consider the unit square domain, $\Omega = [0,1]^2$, which is separated into two sub-regions by the fracture $\gamma = \{(x,y) \in \Omega : x = 0.5\}$. Inside each sub-region Ω_i, $i = 1, 2$, and the fracture γ the exact solutions are chosen as follows :

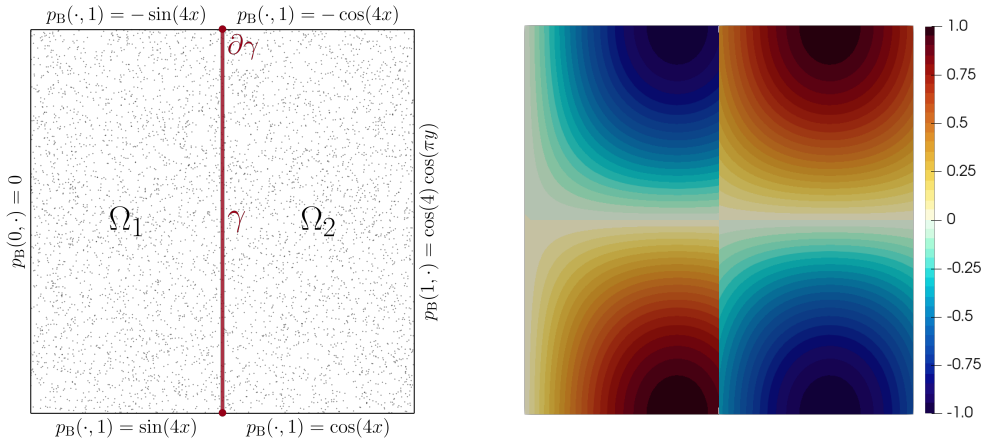
$$p_{\mathrm{B}}(x,y) = \begin{cases} \sin(4x)\cos(\pi y) & \text{if } x < 0.5, \\ \cos(4x)\cos(\pi y) & \text{if } x > 0.5, \end{cases} \quad \text{and} \quad p_\gamma(y) = [\cos(2) + \sin(2)]\cos(\pi y).$$

We note that the coupling conditions (5.8) are satisfied by setting $\boldsymbol{\alpha}_{\mathrm{B}} = \boldsymbol{I}_2$ (identity matrix) and $\alpha_\gamma^{\boldsymbol{n}}/\ell_\gamma = 2$. The source term verifies the exact solution and the coupling conditions as,

$$f_{\mathrm{B}}(x,y) = p_{\mathrm{B}}(x,y)(16+\pi^2), \quad \text{and} \quad f_\gamma(y) = \cos(\pi y)\,[\cos(2)+\sin(2)]\left(\frac{2\alpha_\gamma^{\boldsymbol{n}}}{\ell_\gamma} + \pi^2 \ell_\gamma \boldsymbol{\alpha}_\gamma^{\boldsymbol{\tau}}\right),$$

and we prescribe Dirichlet conditions on the whole boundaries. We take the tangential and normal components of the fracture permeability $\boldsymbol{\alpha}_\gamma^{\boldsymbol{\tau}} = 10^2$ and $\alpha_\gamma^{\boldsymbol{n}} = 4 \cdot 10^{-2}$, respectively.

The fracture thickness is fixed to $\ell_\gamma = 2 \cdot 10^{-2}$ m. In Figure 5.4, we give an illustration of the geometry and the discrete solution calculated with high-order polynomial degree $k = 4$. We clearly observe the discontinuity of the pressure along the interface γ.



(a) (b)

FIGURE 5.4 – Test A - Representation of the geometry and boundary conditions used for the convergence case (a), and pressure solution inside the bulk from the H-IIP scheme with $h = 1/64$ and $k = 4$ (b).

Now, we compute the L^2-norm error estimate ($\|\cdot\|_{0,\mathcal{T}_h,}$) using h- and k-refinement strategies. We use a sequence of structured rectangular meshes obtained by partitioning the unit square into $L \times L$ uniform quadrilaterals, with $L = \{4, 8, 16, 32, 64\}$. In Figure 5.5, we display the history of convergence of the error for all variations of the scheme, namely the Non-Symmetric ($\varepsilon = -1$), the Incomplete ($\varepsilon = 0$), and the Symmetric ($\varepsilon = 1$). We observe that the formalism inherits the behavior of the standard IP scheme inside the bulk region. Indeed, we recover some well-known estimates for both non-symmetric variants. Choosing $\varepsilon = \{-1, 0\}$, the H-IP/CG approach converge optimally with order $k+1$ (*resp.* sub-optimally with order k) for any odd (*resp.* even) polynomial degree k. Conversely, the method converge always optimally for any degree $k \geq 1$ and $\varepsilon = 1$. Inside the fracture γ, the sub-optimal convergence is due to the lack of symmetry of the formalism.

5.4.1.2 Test B - Network of fractures : local injection and production

Next, we assess the flexibility of the H-IP/CG method with more sophisticated situations. For these cases, the H-IP/CG solutions are compared to the finite element solutions obtained with COMSOL. In the rest of the presentation, we will favor the fully Symmetric variant (H-SIP/CG) for the resolution.

Test B.1 - Quarter five-spot problem with regular fracture network

We start with an injection/production process based-on the quarter five-spot setup. The five-spot pattern is widely used in underground energy extraction, where one injection well is located in the middle and four producers are placed at each corners of the square domain.

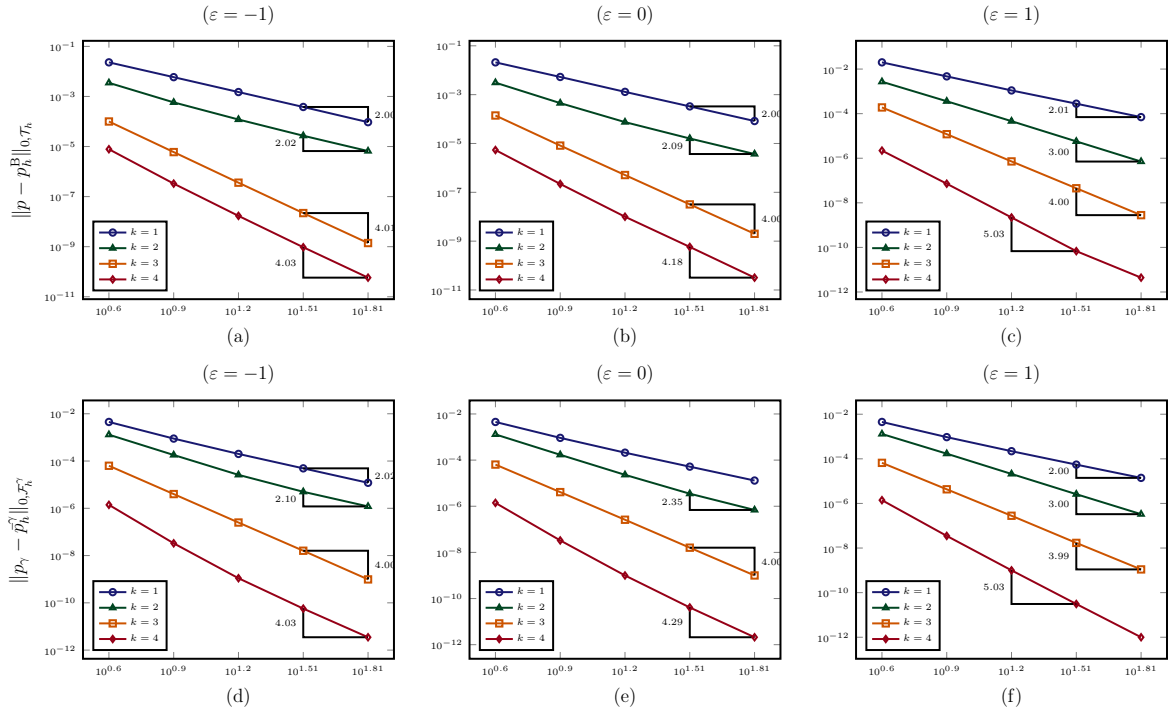


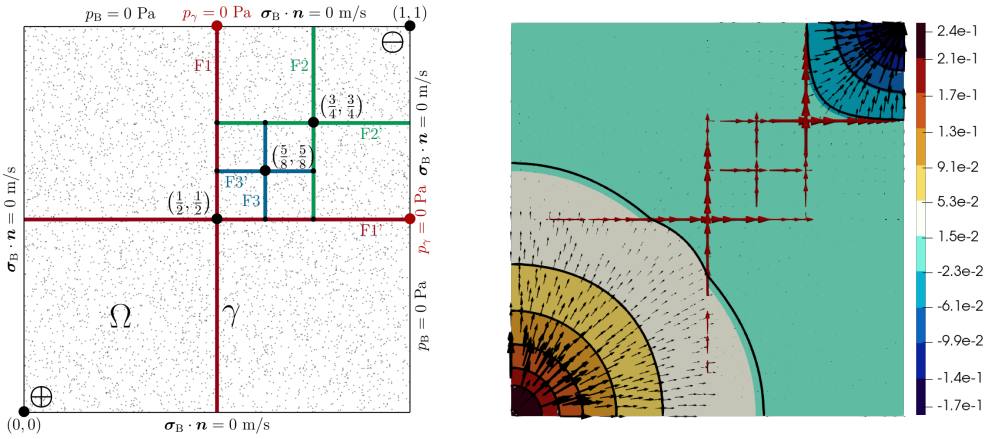
FIGURE 5.5 – Test A - From the left to the right : History of convergence in the $\|\cdot\|_{0,\mathcal{T}_h}$-norm (vs. $1/h$) on the top and in the $\|\cdot\|_{0,\mathcal{F}_h^\gamma}$-norm (vs. $1/h_\gamma$) on the bottom of the H-IP/CG method for the (a-d) Non-Symmetric, (b-e) Incomplete and (c-f) Symmetric variants.

Because the geometry is symmetric, we only consider a quarter of the domain ($\Omega = [0,1]^2$). The injection well is located at $\boldsymbol{x} = (0,0)$ and the single producer at $\boldsymbol{x} = (1,1)$. We increase the complexity by considering a network of intersecting fractures in the domain. The fracture network is inspired by Geiger, where the coordinates are specified in the referenced benchmark [Flemisch 2018]. — *The results of the H-IP/CG method for this benchmark are reported in Appendix B1 (section 5.7.1), and we observe that the data agreed with the existing solutions of the literature. Particularly, the scheme capture as well both the conductive and the blocking fractures.* — In this application, the thickness of the fractures are constant and take $\ell_\gamma = 10^{-2}$ m. The bulk permeability is chosen as the identity matrix, whereas the tangential and normal permeability of the fractures are given by $\boldsymbol{\alpha}_\gamma^{\boldsymbol{\tau}} = 10^2$ m^2 and $\alpha_\gamma^{\boldsymbol{n}} = 1$ m^2, respectively. The presence of the injection and production wells is modeled with the following source term inside the bulk,

$$f_{\mathrm{B}}(x,y) = \frac{1.01}{100}\left[\tanh\left(200\left(R - \sqrt{x^2+y^2}\right)\right) + \tanh\left(200\left(R - \sqrt{(x-1)^2+(y-1)^2}\right)\right)\right],$$

where $R = 0.2$ m is the radius of the wells. The source term is null inside the fractures. In Figure 5.6a, we illustrate the geometry and the boundary conditions used for the simulation.

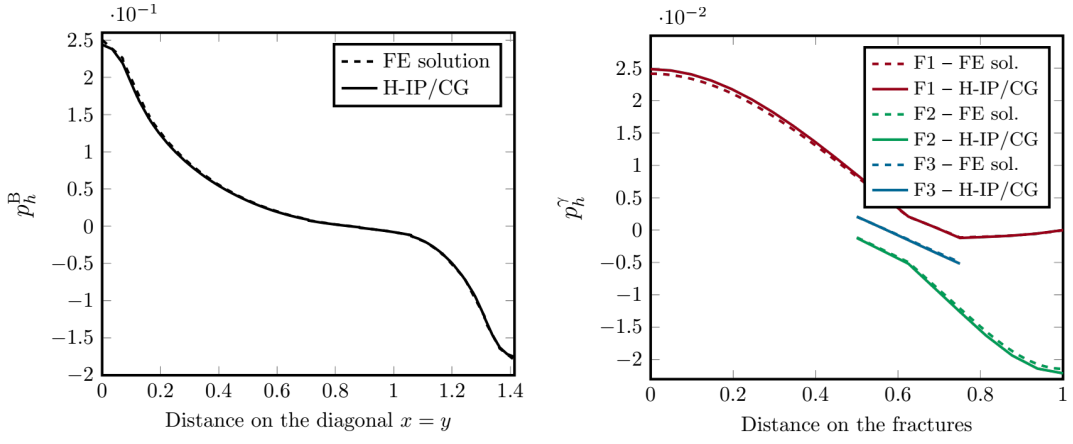
In Figure 5.6b, we present the pressure distribution obtained from the H-IP/CG method that we overlay on the FE solution with continuous black lines. Since the fractures are highly



(a) (b)

FIGURE 5.6 – Test B.1 - Representation of the geometry and boundary conditions used for the quarter five-spot problem (a), and pressure solution inside the bulk from the H-IP/CG scheme with linear functions (b). The FE solution is display with continuous black lines and quadratic functions. The scaled velocity field are presented in red for the fractures ($\times 2e3$) and black for the bulk ($\times 1e2$).

conductive, the velocity field is faster in the fracture network than in the bulk. Then, we suggest in Figure 5.7 a graphical comparison of the pressure profiles from the FE and approximate solutions. Figure 5.7a depicts the pressure profile of the bulk along $(x = y)$–line. The fracture pressure is given in Figure 5.7b. The symmetry of the problem enables us to only consider the fractures F1, F2 and F3. For each representations, the pressure profile decreases from the injection well towards the production well. We close by affirming that the results of the H-IP/CG scheme fits with the FE solution for both the bulk and the fracture network.



(a) (b)

FIGURE 5.7 – Test B.1 - Pressure distribution of the bulk along the line $x = y$ (a), and discrete pressure in the fractures F1, F2 and F3 (b). The results of the finite element solution are depicted with dashed lines.

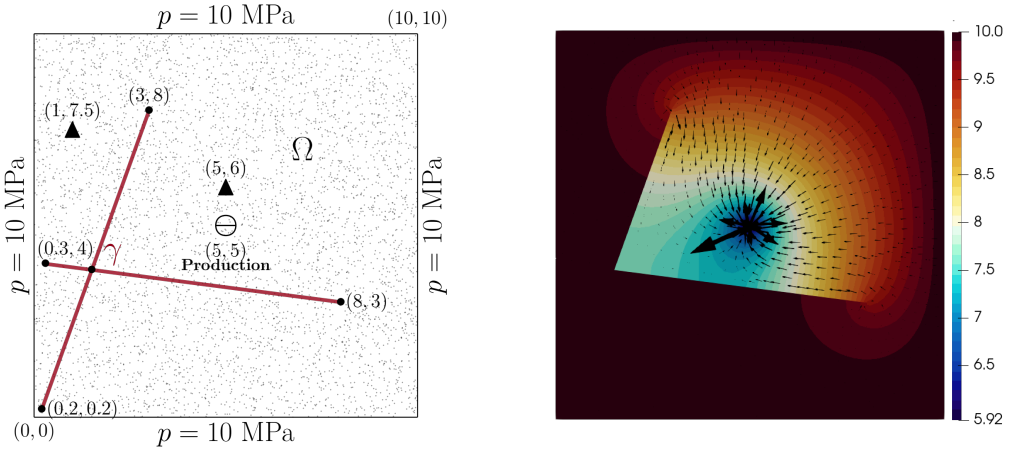
Test B.2 - Production with crossing barriers

Now, we investigate the ability of the methodology to treat immersed tips. Since the width of the fractures is small compared to the dimensions of the host medium, we can neglect the fluid exchange at the intersection of the bulk and the fracture boundaries. This assumption is translated into a *no-flow* boundary condition ($-\ell_\gamma \boldsymbol{\alpha}_\gamma^{\boldsymbol{\tau}} \boldsymbol{\nabla}_{\boldsymbol{\tau}} p_\gamma \cdot \boldsymbol{\tau} = 0$) at the immersed tips.

For this numerical experiment, we perform the simulation of a time-dependent water pumping through a confined aquifer. The square domain ($\Omega = [0, 10]^2$) is crossed by two immersed and intersecting geological barriers (see Figure 5.8a). The production well is located at the center of the domain, and its presence is embedded by the following source term :

$$f_{\mathrm{B}}(x, y) = \begin{cases} -4 \cdot 10^{-4} & \text{if } \left(R - \sqrt{(x-5)^2 + (y-5)^2}\right) > 0 \\ 0 & \text{else} \end{cases}$$

where $R = 0.2$ m is the radius of the well. The bulk permeability is set to $\boldsymbol{\alpha}_{\mathrm{B}} = 10^{-11} \boldsymbol{I}_2$ m^2. The tangential and normal permeability of the fracture network take value $\boldsymbol{\alpha}_\gamma^{\boldsymbol{\tau}} = \alpha_\gamma^{\boldsymbol{n}} = 2.5 \cdot 10^{-15}$ m^2. We fix the thickness of the barriers to $\ell_\gamma = 10^{-1}$ m. Then, we enforce Drichlet boundary conditions, $p_{\mathrm{B}} = 10$ MPa, on the entire frontier. At the initial time step ($t = 0$ s) the initial pressure value is set 10 MPa, and the storage coefficient $\phi_{\mathrm{B}} = \phi_\Gamma = 10^{-8}$. The domain setup, including the position of two observations points are reported in Figure 5.8a.

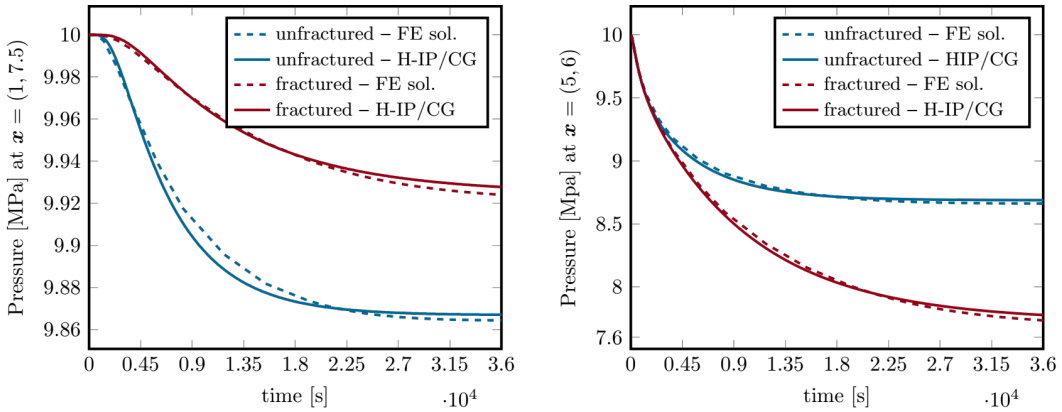


(a) (b)

FIGURE 5.8 – Test B.2 - Representation of the geometry and boundary conditions used for the crossing barriers problem (a), and pressure solution inside the bulk from the H-IP/CG scheme with unstructured triangular mesh ($h = 1/40$) and linear functions (b). The scaled velocity field is presented in black for the bulk, and the black triangles represents the observation points.

The simulation is performed until the solution reaches the steady state, using the BDF scheme to discretized the time derivative and the uniform time-step $\triangle t = 100$ s. We illustrate in Figure 5.8b the pressure of the bulk with the velocity field at the steady state. As expected, we clearly observe the jump of the pressure along each barriers. In the vicinity of the production well, the velocity is larger than in the rest of the domain leading to a drop of the pressure. Afterwards, we plot in Figure 5.9 the FE and H-IP/CG pressure distributions with

and without the barriers at the observation points, namely $\boldsymbol{x} = (1, 7.5)$ and $\boldsymbol{x} = (5, 6)$, for the time-domain T $= [0, 36000]$s. At least, we observe again a decrease of the pressure over time. All results of the H-IP/CG method are in agreement with the one of the FE solution.



(a) (b)

FIGURE 5.9 – Test B.2 - Evolution of the pressure at the observation points $\boldsymbol{x} = (1, 7.5)$ (a), and $\boldsymbol{x} = (5, 6)$ (b) vs. time $[s]$ with and without barriers. The results of the reference solution are depicted with dashed lines.

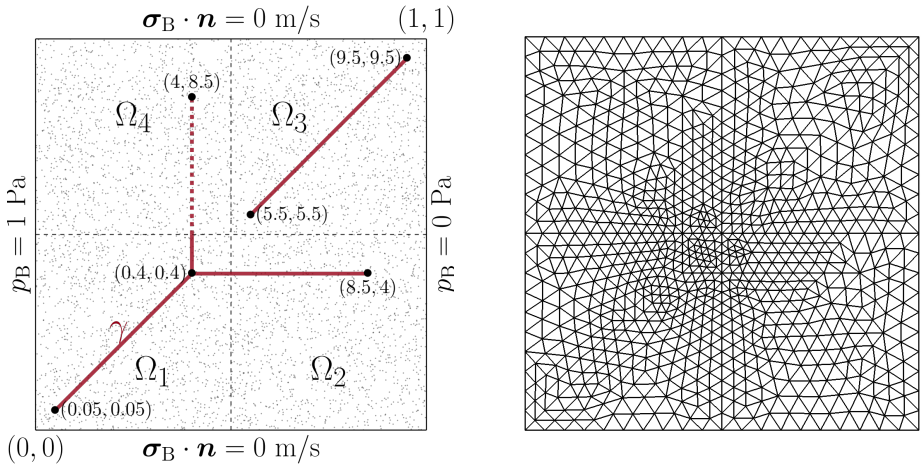
5.4.2 Test C - Heterogeneous and anisotropic matrix

All previous test cases deal with isotropic domains. Now, we evaluate the capability of the H-IP/CG method to deal simultaneously with the anisotropies of the bulk and the presence of fractures. We assume that the unit domain is split into four sub-regions : $\Omega_1 = [0, 1/2]^2$, $\Omega_2 = [1/2, 1] \times [0, 1/2]$, $\Omega_3 = [1/2, 1]^2$, and $\Omega_4 = [0, 1/2] \times [1/2, 1]$ such that, $\Omega := \cup_{i=1}^4 \Omega_i$. For each sub-domains, we define separately their respective permeability tensor-valued as follows,

$$\boldsymbol{\alpha}_{\mathrm{B}} = \begin{bmatrix} 1 & 0 \\ 0 & \alpha_{\mathrm{B}} \end{bmatrix} \text{ for } (x, y) \in \Omega_1, \Omega_3, \qquad \text{and} \qquad \boldsymbol{\alpha}_{\mathrm{B}} = \begin{bmatrix} \alpha_{\mathrm{B}}^{-1} & 0 \\ 0 & 1 \end{bmatrix} \text{ for } (x, y) \in \Omega_2, \Omega_4,$$

where $\alpha_{\mathrm{B}} = 10^{-4}$ stands for the strongest anisotropic ratio. We fix a constant tangential and normal permeability inside each fractures as $\boldsymbol{\alpha}_\gamma^{\boldsymbol{\tau}} = \alpha_\gamma^{\boldsymbol{n}} = 10$ m^2, and the aperture is taken as $\ell_\gamma = 10^{-3}$ m. The geometry of the problem is shown in Figure 5.10a, in which the fracture network is highlighted in red. According to the previous notations, the fractures acts as a preferential path to the flow in Ω_j, $j = \{1, 2, 3\}$, whereas we have a barrier in Ω_4. For the simulation, the flow path is driven by the boundary conditions, where we enforce a unit pressure gradient from the left to the right frontier. On the remaining boundaries, we impose no-flow boundary condition.

The pressure fields from the FE and H-IP/CG solution are depicted in Figure 5.11, with a coarser and finer meshes (see Figure 5.10b for an illustration of the coarse grid). We observe that the given results are relatively similar for each configurations. However, we remark that the FE solution exhibits spurious oscillations along the discontinuity of the bulk violating the minimum and maximum principle (see also Figure 5.24 in Appendix B2 (section 5.7.2),



(a) (b)

FIGURE 5.10 – Test C - Representation of the geometry and boundary conditions used for the heterogeneous and anisotropic problem with the four sub-domains (a). The fracture network is represented in red with dashed and solid lines for the impermeable and permeable fractures, respectively. Example of the coarse mesh with 1712 triangular elements and 48 fracture elements (b).

where the values out of the bounds are presented). On the contrary, the H-IP/CG method preserves the upper and lower limits independently of the mesh setup.

5.4.3 Test D - Normal fluxes evaluation

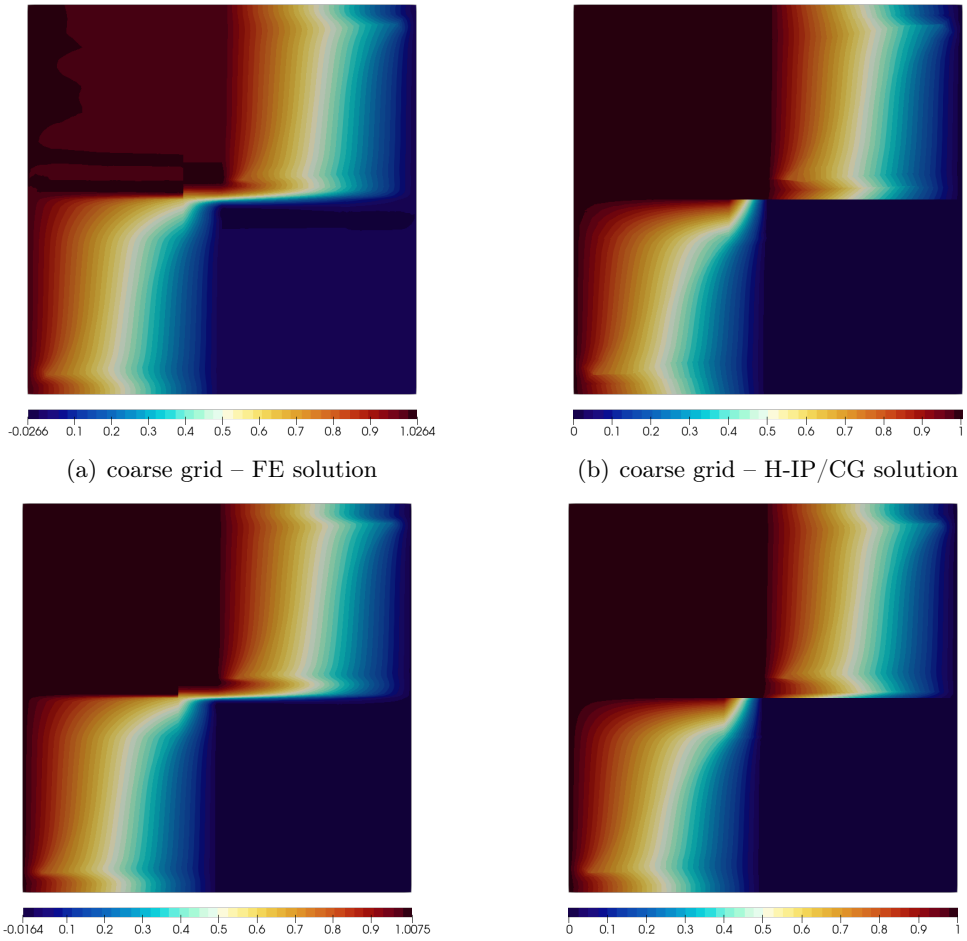
In environmental resources management, normal fluxes quantification in fractured porous media is of primordial interest. In this application, we measure the efficiency of the formalism in the computation of the internal fluxes. We assume that the rectangular domain, $\Omega = [0, 2] \times [0, 1]$, is composed of several layers and faults with aperture $\ell_\gamma = 10^{-3}$ m. We enforce Dirichlet boundary condition $p_{\mathrm{B}} = 5$ on the left, and $p_{\mathrm{B}} = 0$ on the right. On the remaining parts of the frontier we have a no-flow boundary condition. Next, we consider two cases :

- **Case 1** - Single fracture : The fault separates the bulk into two disjoints sub-domains Ω_l and Ω_r. In this context, we consider two different scenarios (i) the bulk is homogeneous with $\boldsymbol{\alpha}_{\mathrm{B}} = 10^{-4}\boldsymbol{I}_2$ m^2 ; and (ii) the permeability of the bulk is define separately in each layers such that $\boldsymbol{\alpha}_{\mathrm{B}|\Omega_l} = 10^{-2}\boldsymbol{I}_2$ m^2 and $\boldsymbol{\alpha}_{\mathrm{B}|\Omega_r} = 10^{-4}\boldsymbol{I}_2$ m^2.
- **Case 2** - Multiple fractures : The fracture domain consist of 2- or 4-parallel and equally spaced fractures. The permeability of the bulk is homogeneous $\boldsymbol{\alpha}_{\mathrm{B}} = 10^{-4}\boldsymbol{I}_2$ m^2.

For all configurations, the tangential and normal permeability of the fractures are set to $\boldsymbol{\alpha}_\gamma^{\boldsymbol{\tau}} = \alpha_\gamma^{\boldsymbol{n}} = 10^{-6}$ m^2. The geometry description of the three domains are given in Figure 5.12a for the Case 1 and Figure 5.13a-b for the Case 2. We listed in Table 5.2 the coordinates of the three fracture networks.

Here, we investigate the flux at the outlet frontier, i.e., the right boundary face Γ_{out}, by computing the flux with the following expression :

$$\mathrm{Flux}_{\mathrm{out}} = -\langle \boldsymbol{\alpha}_{\mathrm{B}} \boldsymbol{\nabla}_h p_h^{\mathrm{B}}, \boldsymbol{n}_{\mathrm{out}} \rangle_{0,\Gamma_{\mathrm{out}}} \quad \text{on } \Gamma_{\mathrm{out}},$$



(a) coarse grid – FE solution

(b) coarse grid – H-IP/CG solution

(c) fine grid – FE solution

(d) fine grid – H-IP/CG solution

FIGURE 5.11 – Test C - Illustration of the reference solution (a-c) and the discrete solution (b-d) with unstructured triangular meshes. The coarse grid (a-b) 1712 triangular elements and 48 fracture elements. The fine grid (c-d) uses 6848 triangular elements and 97 fracture elements.

where $\boldsymbol{n}_{\text{out}}$ refers to unit normal on the outlet Γ_{out}. In Figures 5.12b-c and 5.13c-d, we illustrate the pressure distribution for each configurations. As expected, the pressure decreases from the left to the right faces, and we retrieve the pressure drop across the fracture γ for the impermeable fracture setup. For the permeable fracture domain, the pressure field is quasi-continuous inside the bulk region. In Figures 5.14 and 5.15, we present the values of the flux at the outlet boundary with various mesh sizes for both cases. Globally, the normal flux evaluation is getting more accurate and precise when the mesh size reduces. The FE and the H-IP/CG solutions produce the same approximation with the finer mesh. However, the H-IP/CG method outperforms the FE solution providing a better estimation of the flux with coarser grids. An exception still persists for the heterogeneous media of Case 1, where both solutions behave similarly for the impermeable fracture. Even if the results are not presented, we observed a similar behavior for multiple and impermeable fractures.

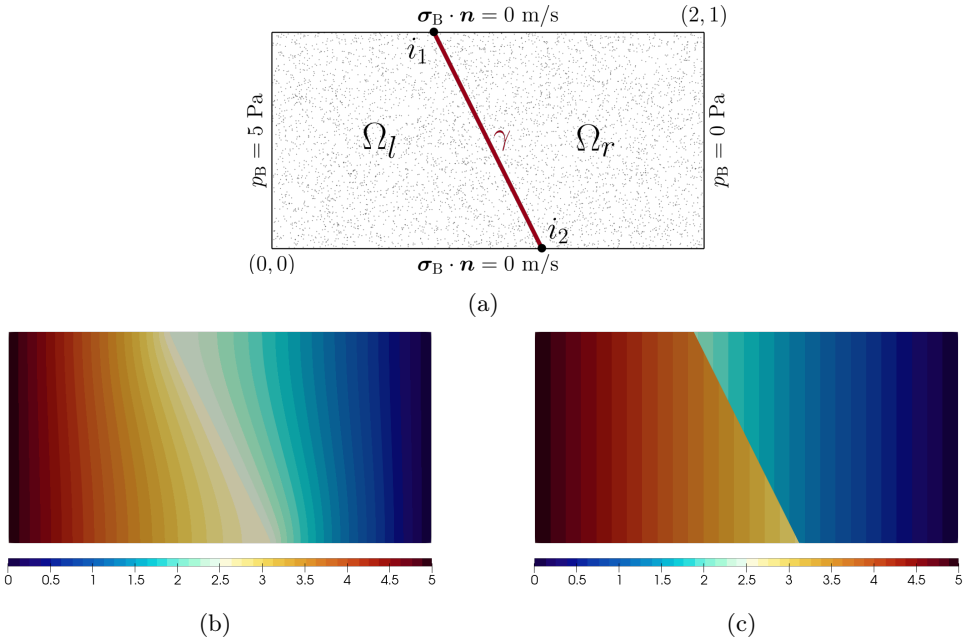


FIGURE 5.12 – Test D - Representation of the geometry and boundary conditions used for the flux evaluation problem for the Case 1 (a). Illustration of the pressure solution for the permeable (b) and the impermeable fracture (c) inside the bulk, with unstructured triangular mesh ($h = 1/32$) and linear functions.

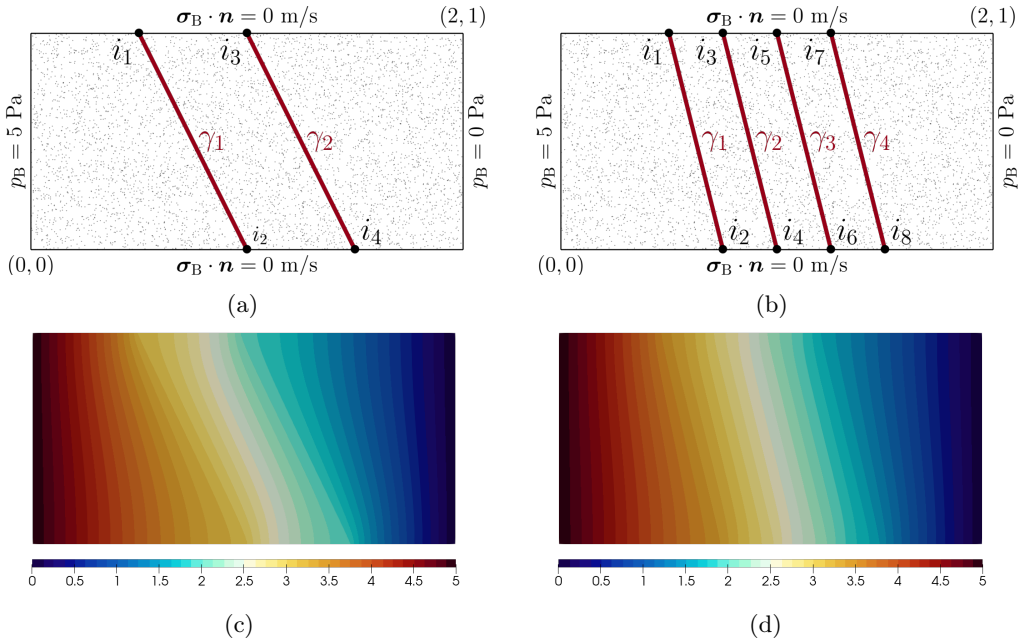
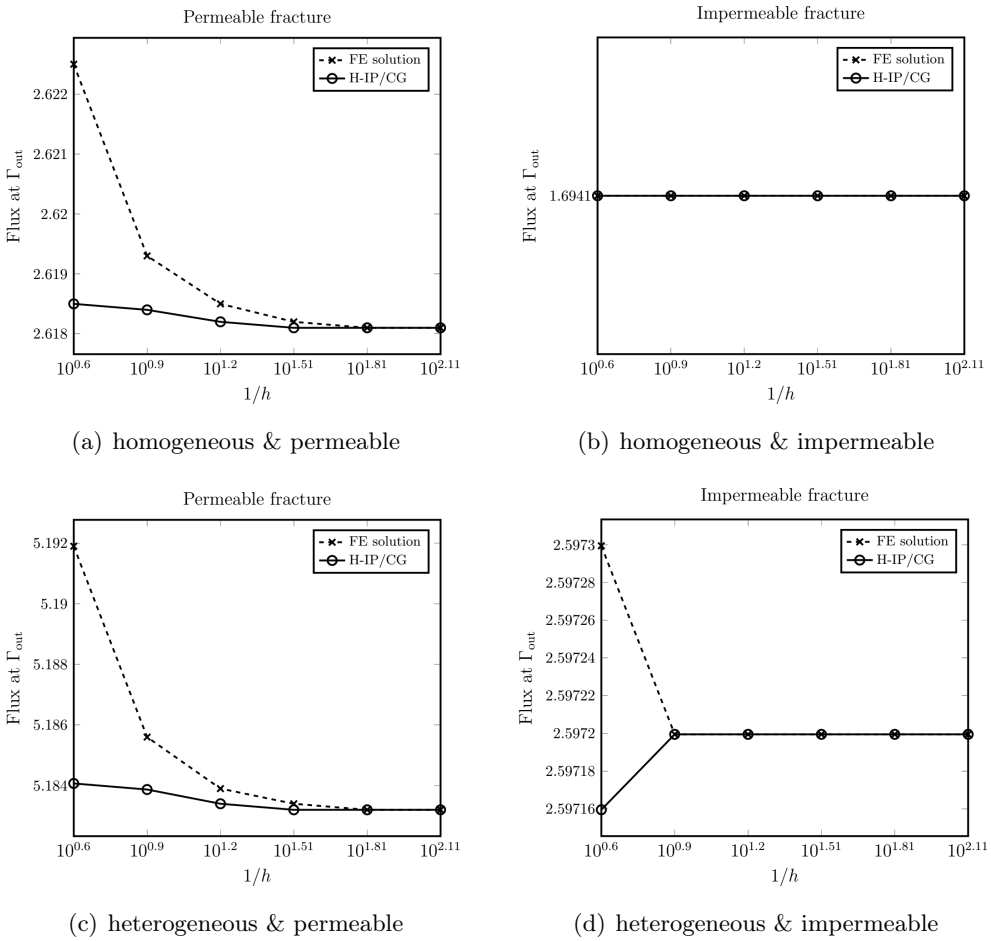


FIGURE 5.13 – Test D - Representation of the geometry and boundary conditions used for the flux evaluation problem for the Case 2 with 2- (a) and 4-fractures (b). Illustration of the pressure solution with 2- (c) and 4-fractures (d) inside the bulk using unstructured triangular mesh ($h = 1/32$) and linear functions.



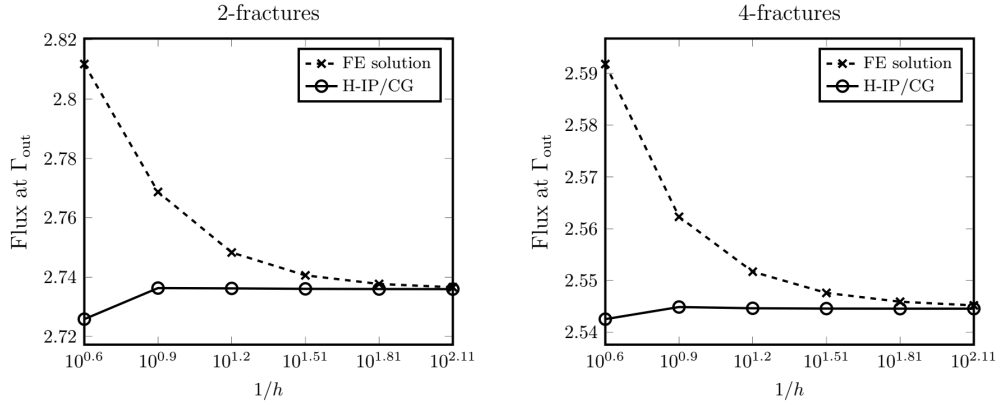
(a) homogeneous & permeable

(b) homogeneous & impermeable

(c) heterogeneous & permeable

(d) heterogeneous & impermeable

FIGURE 5.14 – Test D - Evaluation of the flux at the outlet boundary for the Case 1.



(a)

(b)

FIGURE 5.15 – Test D - Evaluation of the flux at the outlet boundary for the Case 2.

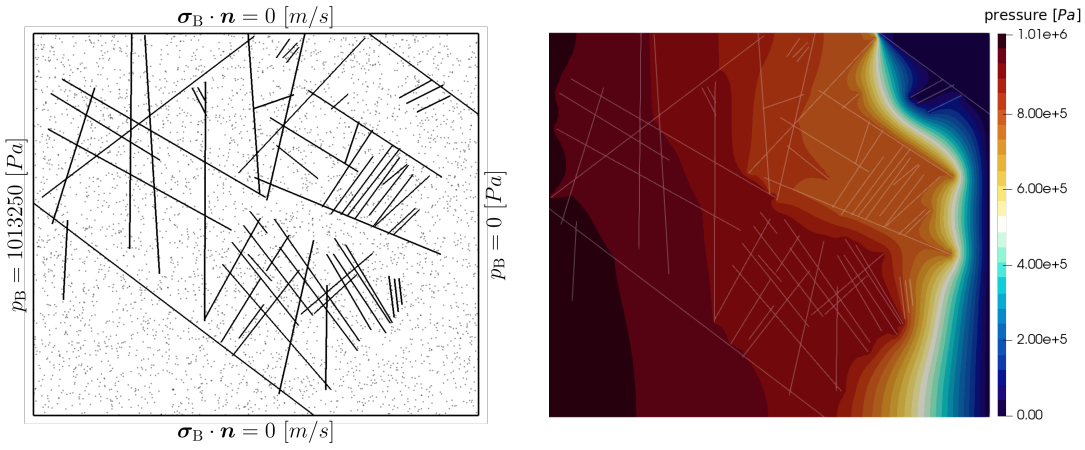
5.4.4 Test E - Realistic fracture networks

For this section, we evaluate the flexibility of the H-IP/CG formalism to handle more realistic fracture networks in 2D and 3D.

5.4.4.1 Test E.1 - 2D example : Sotra island benchmark

This example is used to analyze the methodology on a time-dependent realistic fracture network problem. The geometry represents a set of 64 intersecting fractures from an interpreted outcrop in Sotra island, near Bergen in Norway, of the benchmark study [Flemisch 2018]. The bulk has a size of 700×600 m, with permeability $\boldsymbol{\alpha}_{\mathrm{B}} = 10^{-14}\boldsymbol{I}_2$ m^2. The tangential and normal permeability of the fractures are fix to $\boldsymbol{\alpha}_{\gamma}^{\boldsymbol{\tau}} = 10^{-14}$ m^2 and $\alpha_{\gamma}^{\boldsymbol{n}} = 10^{-8}$ m^2, respectively. The storage coefficient is constant $\phi_{\mathrm{B}} = \phi_{\Gamma} = 10^{-8}$. The temporal domain is given as $\mathrm{T} = [0, 1094]$ years, where an uniform time step of length $\triangle t = 0.1$ years is used. At the initial time ($t = 0$ s) the pressure is null inside the bulk. Next, we consider two different scenarios depending on the fracture aperture :

- **Case 1** - Constant : the thickness of each fractures is set to $\ell_{\gamma} = 10^{-2}$ m ;
- **Case 2** - Variable : we categorize the fractures into three distintive clusters, namely FN1, FN2 and FN3. FNi refers to the i-th fracture network, $i = \{1, 2, 3\}$, which consists of the large faults (FN1), the medium size fractures (FN2), and the small fractures (FN3). The thickness of FN1, FN2 and FN3 are set to 20^{-2} m, 10^{-2} m and $5 \cdot 10^{-3}$ m, respectively. Here, we investigate the influence of the three fracture networks on the pressure distribution by removing FN3 and FN2 successively. To the best of our knowledge, this is the first attempt that a variable aperture is applied for this benchmark.



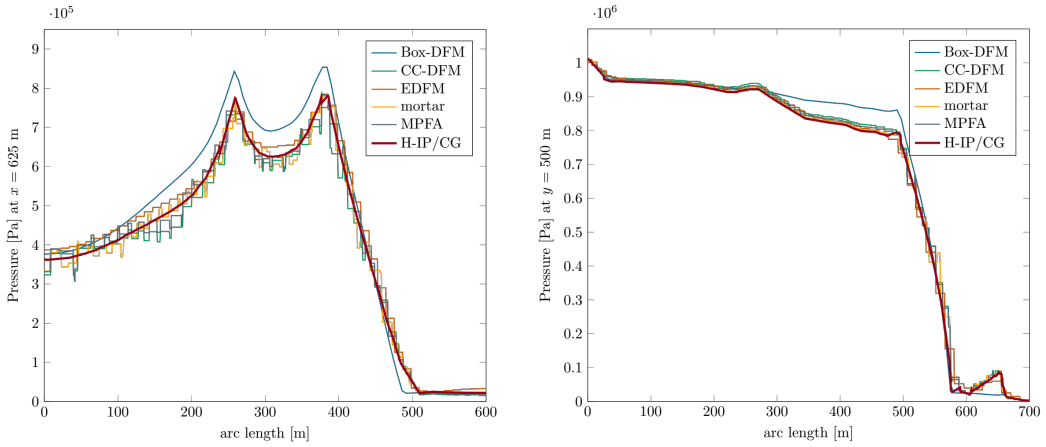
(a) (b)

FIGURE 5.16 – Test E.1 - Representation of the geometry and boundary conditions used for the realistic benchmark (a), and discrete pressure p_h^{B} (at the steady state) inside the bulk for the Case 1 with unstructured triangular mesh (2534 elements) and linear functions (b).

The geometry description and the boundary conditions are shown in Figure 5.16a. We depict, in Figure 5.18a, the fracture network FN1 in red, FN2 in blue and FN3 in green, where we

also include the location of four observations points with black triangles. Notice that at the final time, $T = 1094$ years, all solutions reach the steady-state.

In Figure 5.16b, we illustrate the pressure distribution for the Case 1 at the final time-level. Then, we compare the pressure distribution with the existing solutions of the literature. Figure 5.17 displays the approximate pressure along the lines $y = 500$ m and $x = 625$ m. We clearly observe that the pressure profiles of the H-IP/CG method are in agreement with the other schemes of the benchmark. Moreover, the number of DOFs and the conditioning of the methods, listed in Table 5.1, corroborate the conclusion established previously.



(a) (b)

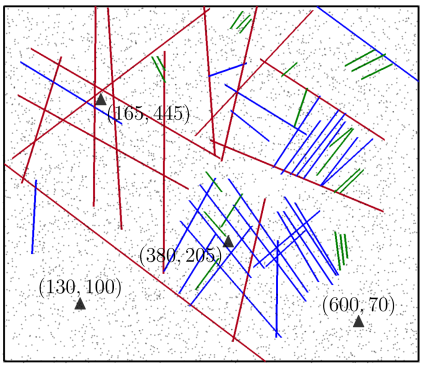
FIGURE 5.17 – Test E.1 - Comparison of the discrete pressure (at the steady state) with referenced solution of the benchmark study [Flemisch 2018] along the lines $x = 625$ (a) and $y = 500$ (b).

In Figure 5.18b-d, we present the discrete pressure for the Case 2 at the steady-state, and we evaluate the behavior of the H-IP/CG solution at the four observation points during the period T (see Figure 5.19). By removing successively the fracture networks FN3 or FN2, we remark that the pressure increase almost identically over time. An exception still remains, when the fracture network comprises only the large fractures. In this context, the absence of fractures in the bottom-right region dramatically reduces the pressure values.

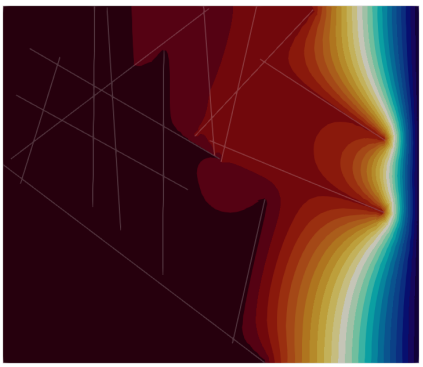
This numerical experiment highlights the influence of the large and medium fractures on the flow path, which cannot be afford. As a contrast, the small fractures can be neglected under certain conditions and hypothesis.

5.4.4.2 Test E.2 - 3D example : Algerøyna outcrop benchmark

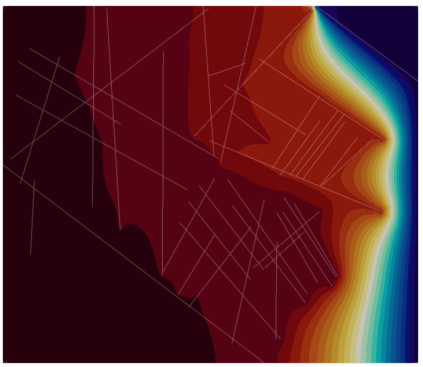
The last experiment is a 3D benchmark case proposed in [Berre 2021]. The geometry represents a set of 52 fractures from an interpreted outcrop of the island of Algerøyna, outside Bergen island, in Norway. The bulk domain is the box $\Omega = [-500\text{m}, 350\text{m}] \times [100\text{m}, 1500\text{m}] \times [-100\text{m}, 500\text{m}]$ (see Figure 5.20). The bulk permeability tensor-valued is set to $\boldsymbol{\alpha}_{\text{B}} = \boldsymbol{I}_3$ m^2. The tangential and normal permeability of the fractures take values $\boldsymbol{\alpha}_\gamma^{\boldsymbol{\tau}} = \alpha_\gamma \boldsymbol{I}_3$ m^2 and $\alpha_\gamma^{\boldsymbol{n}} = \alpha_\gamma$ m^2, respectively, with the parameter $\alpha_\gamma = 10^4$. The fracture aperture is $\ell_\gamma = 10^{-2}$



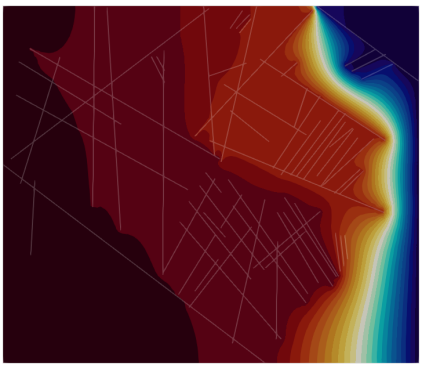
(a) fracture networks



(b) pressure with FN1



(c) pressure with FN1, FN2



(d) pressure with FN1, FN2, FN3

FIGURE 5.18 – Test E.1 - Representation of the three fracture networks for the Case 2, with FN1 in red, FN2 in blue and FN3 in green (a). The location of the observation points are represented with black triangles. Illustration of the discrete pressure (at the steady state) inside the bulk by removing FN2 and FN3, with unstructured triangular mesh (2534 elements) and linear functions (b).

Method	**DOFs**	**matrix el.**	**fracture el.**	**cond.**
Box-DFM	5563	10807 trigs	1386	$9.3 \cdot 10^5$
CC-DFM	8481	7614 trigs	867	$5.3 \cdot 10^6$
EDFM	3599	2491 quads	1108	$4.7 \cdot 10^6$
mortar	25258	8319 trigs	1317	$2.2 \cdot 10^{17}$
H-IP/CG	8984	2534 trigs	419	$6.07 \cdot 10^4$

TABLE 5.1 – Test E.1 - realistic : Performances of the H-IP/CG method compared to the existing schemes of the literature in terms of DOFs, number of elements, and condition number (cond.).

m. We enforce homogeneous Dirichlet boundary conditions ($p_{\mathrm{B}} = 0$) on the oulet frontier :

$$\partial\Omega_{\mathrm{out}} := (\{-500\} \times [100, 400] \times \times [-100, 100]) \cup (\{350\} \times [100, 400] \times [-100, 100]),$$

and an uniform unit inflow $\boldsymbol{\sigma}_{\mathrm{B}} \cdot \boldsymbol{n}_{\mathrm{in}} = 1$ m/s is imposed on the inlet boundary :

$$\partial\Omega_{\mathrm{in}} := ([-500, -200] \times \{1500\} \times [300, 500]) \cup (\{-500\} \times [1200, 1500] \times [300, 500]).$$

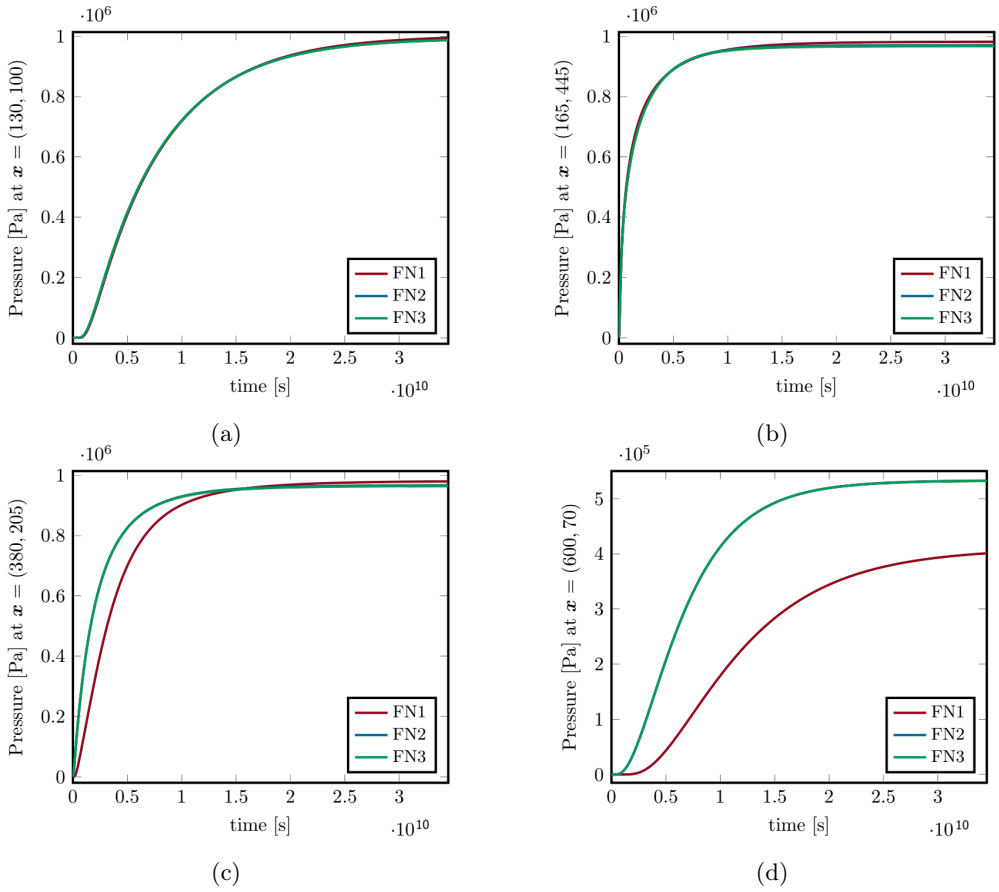


FIGURE 5.19 – Test E.1 - Evolution of the pressure p_h^{B} at the four observation points vs. time $[s]$ (Case 2).

On the remaining frontiers, i.e., $\partial\Omega_{\mathrm{N}} := \partial\Omega \backslash \{\partial\Omega_{\mathrm{in}} \cup \partial\Omega_{\mathrm{out}}\}$, we enforce a no-flow boundary condition. Notice that the test case description is accompanied by data of the geometry description, the existing simulation results, plotting scripts, available in the Git repository [4].

Because of the complexity of the geometry, the grid is generated using the mesh generator gmsh [Geuzaine 2009], which leads to 899682 globally coupled unknowns with linear polynomial functions (354496 bulk element-based, 534363 bulk trace-based DOFs, and 10523 interior-based DOFs). The simulation is performed using a tetrahedral mesh with 88624 bulk elements and 18234 fracture elements. An element-wise BDDC [5] preconditioner is favor for the resolution, which produces 169 iteration using a tolerence of 10^{-14}.

We compare in Figure 5.21 the pressure distribution of the H-IP/CG method with the existing solutions of of the benchmark, along the two diagonal lines $(-500, 100, 100)$–$(350, 1500, 500)$ and $(350, 100, -100)$–$(-500, 1500, 500)$. Finally, we can conclude that the method still permformed as well, since the results perfectly match with the published spread area.

4. https://git.iws.uni-stuttgart.de/benchmarks/fracture-flow-3d.git
5. BDDC : Balancing Domain Decomposition preconditioner with Constraints

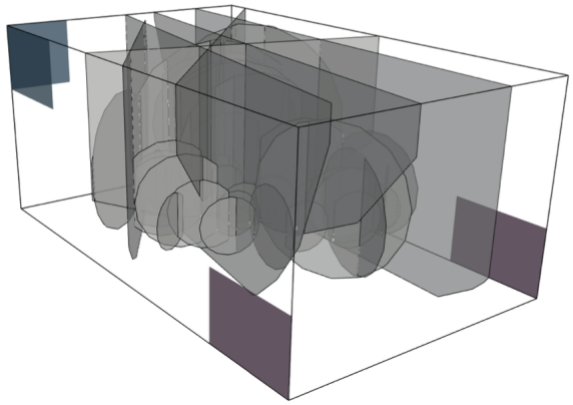
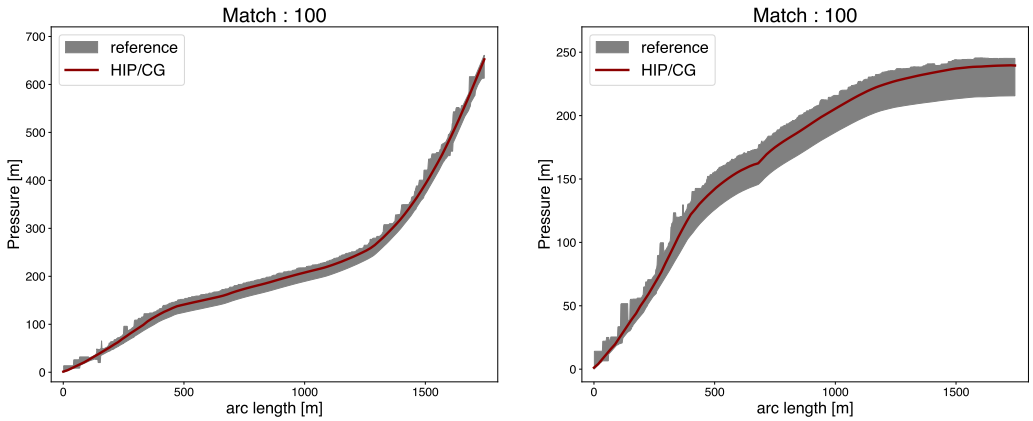


FIGURE 5.20 – Test E.2 - Representation of the bulk geometry and the fracture network given by the benchmark study [Berre 2021], where the inlet boundaries are display in blue and the outlet in purple.



(a) (b)

FIGURE 5.21 – Test E.2 - Representation of the pressure p_h^{B} (using linear function) with the referenced solution of the benchmark study [Flemisch 2018] along the lines $(350, 100, -100)$–$(-500, 1500, 500)$ (a) and $(-500, 100, 100)$–$(350, 1500, 500)$ (b). The shaded area displays the standard deviation of the methods of the benchmark, and the red line shows the results of the H-IP/CG scheme.

5.5 Conclusion

In the present paper, we derived a high-order nonconforming/conforming Galerkin discretization strategy for modeling fractured groundwater flow problems in the codimensional setting. In this context, the fracture is represented as a $(d-1)$–dimensional geometric object embedded in the bulk region.

The novelty of this work consists of the coupling of two formalisms eligible for static condensation, namely the Hybridizable Interior Penalty (H-IP) method in the bulk and the standard continuous Galerkin (CG) approach in the fractures. This result in a compact matrix system, which depends only on the skeleton-based DOFs and directly impacts the CPU time. Then, we measure the performances and efficiency of the scheme using a series of

extensive numerical experiments, even in the context of conductive and blocking fractures. Firstly, we retrieve the usual convergence properties of the IPDG framework for each values of $\varepsilon = \{0, \pm 1\}$. Second, we applied sucessfully the flexibility of the H-IP/CG method to more complex fracture networks, including realistic benchmarks in 2D and 3D, and time-dependent groundwater flow cases. Especially, the comparisons with Comsol helped in the validation of the methodology. We proved that the formalism covers as well the anisotropies of the bulk, the presence of fractures, and the use of various fracture aperture simultaneously.

In the forthcoming works, we will extend the concept of DG or HDG method to the $(d-1)$-dimensional setting, and apply the resulting discretization scheme to the coupled flow-transport problem in fractured porous media.

5.6 Appendix A - Implementation aspects

In this appendix, we provide complementary details for the implementation aspects of the static condensation procedure.

5.6.1 A1 - Notations for the block-matrices

In this part, we define the subblock-matrices associated with (5.24) by the following expressions :

- **Block-Matrices $\mathbb{A}$ and $\mathbb{E}$:**

$$\mathbb{A}_1|_{ij} := (\phi_{\mathrm{B}} q_i^{\mathrm{B}}, q_j^{\mathrm{B}})_{0,\mathcal{T}_h}, \tag{5.26a}$$

$$\mathbb{A}_2|_{ij} := \langle \phi_\gamma q_i^\gamma, q_j^\gamma \rangle_{0,\mathcal{F}_h^\gamma}, \tag{5.26b}$$

$$\mathbb{E}_2|_{ij} := \langle \phi_\gamma \hat{q}_i^\gamma, \hat{q}_j^\gamma \rangle_{0,\mathcal{F}_h^\gamma}. \tag{5.26c}$$

  We point out that matrices $\mathbb{A}_1$, $\mathbb{A}_2$ and $\mathbb{A}_3$ have a block-diagonal structure rendering their inversion so quite.

- **Block-Matrix $\mathbb{B}^{(\varepsilon)}$:**

$$\begin{aligned}\mathbb{B}_1^{(\varepsilon)}|_{ij} :=& (\boldsymbol{\alpha}_{\mathrm{B}} \boldsymbol{\nabla}_h q_i^{\mathrm{B}}, \boldsymbol{\nabla}_h q_j^{\mathrm{B}})_{0,\mathcal{T}_h} - \langle \boldsymbol{\alpha}_{\mathrm{B}} \boldsymbol{\nabla}_h q_i^{\mathrm{B}}, q_j^{\mathrm{B}} \rangle_{0,\partial\mathcal{T}_h^{\mathrm{B}}} - \varepsilon \langle \boldsymbol{\alpha}_{\mathrm{B}} \boldsymbol{\nabla}_h q_j^{\mathrm{B}}, q_i^{\mathrm{B}} \rangle_{0,\partial\mathcal{T}_h^{\mathrm{B}}} + \\ & \langle \tau_{\mathrm{B}} q_i^{\mathrm{B}}, q_j^{\mathrm{B}} \rangle_{0,\partial\mathcal{T}_h^{\mathrm{B}}} + \langle a_\gamma q_i^{\mathrm{B}}, q_j^{\mathrm{B}} \rangle_{0,\partial\mathcal{T}_h^\gamma},\end{aligned} \tag{5.27a}$$

$$\mathbb{B}_2|_{ij} := \langle \ell_\Gamma \boldsymbol{\alpha}_\gamma^{\boldsymbol{\tau}} \boldsymbol{\nabla}_{\boldsymbol{\tau}} q_j^\gamma, \boldsymbol{\nabla}_{\boldsymbol{\tau}} q_i^\gamma \rangle_{0,\mathcal{F}_h^\gamma} + \langle a_\gamma q_j^\gamma, q_i^\gamma \rangle_{0,\partial\mathcal{T}_h^\gamma}, \tag{5.27b}$$

$$\mathbb{I}_1|_{ij} := -\langle a_\gamma q_j^\gamma, q_i^{\mathrm{B}} \rangle_{0,\partial\mathcal{T}_h^\gamma}. \tag{5.27c}$$

  Let us emphasize that the global matrix $\mathbb{B}^{(\varepsilon)}$ is symmetric if and only if its local one $\mathbb{B}_1^{(\varepsilon)}$ is itself symmetric too, i.e., if $\varepsilon = 1$. We also underline the block-diagonal structure of matrices $\mathbb{B}_1^{(\varepsilon)}$ and $\mathbb{B}_2$ rendering their inversion immediate at the element level. Here, $\mathbb{I}_1$ denotes an interaction matrix between the interior-based DOFs q_h^{B} and q_h^γ, respectively.

- **Block-Matrices $\mathbb{C}$ and $\mathbb{C}^{(\boldsymbol{\varepsilon})}$:**

$$\mathbb{C}_1|_{ij} := \langle \boldsymbol{\alpha}_{\mathrm{B}} \boldsymbol{\nabla}_h q_i^{\mathrm{B}} \cdot \boldsymbol{n} - \tau_{\mathrm{B}} q_i^{\mathrm{B}}, \hat{q}_j^{\mathrm{B}} \rangle_{0,\partial\mathcal{T}_h^{\mathrm{B}}}, \tag{5.28a}$$

$$\mathbb{C}_2^{(\varepsilon)}|_{ij} := \langle \varepsilon \boldsymbol{\alpha}_{\mathrm{B}} \boldsymbol{\nabla}_h q_j^{\mathrm{B}} \cdot \boldsymbol{n} - \tau_{\mathrm{B}} q_j^{\mathrm{B}}, \hat{q}_i^{\mathrm{B}} \rangle_{0,\partial\mathcal{T}_h^{\mathrm{B}}}, \tag{5.28b}$$

$$\mathbb{I}_2|_{ij} := -\langle a_\gamma \hat{q}_j^\gamma, q_i^{\mathrm{B}} \rangle_{0,\partial\mathcal{T}_h^\gamma}, \tag{5.28c}$$

$$\mathbb{I}_3|_{ij} := \langle \ell_\gamma \boldsymbol{\alpha}_\gamma^{\boldsymbol{\tau}} \boldsymbol{\nabla}_{\boldsymbol{\tau}} \hat{q}_j^\gamma, \boldsymbol{\nabla}_{\boldsymbol{\tau}} q_i^\gamma \rangle_{0,\mathcal{F}_h^\gamma} + \langle a_\gamma \hat{q}_j^\gamma, q_i^\gamma \rangle_{0,\partial\mathcal{T}_h^\gamma}. \tag{5.28d}$$

We underline the symmetric property $\mathbb{C}_1^t = \mathbb{C}_2^{(\varepsilon)}$ if $\varepsilon = 1$. Here, $\mathbb{I}_2$ (*resp.* $\mathbb{I}_3$) denotes an interaction matrix between the interior-based DOFs q_h^{B} (*resp.* q_h^γ) and the skeleton-based DOFs $\hat{q}_h^\gamma$, respectively.

- **Block-Matrix** $\mathbb{D}$:

$$\mathbb{D}_1|_{ij} := \langle \tau_{\mathrm{B}} \hat{q}_j^{\mathrm{B}}, \hat{q}_i^{\mathrm{B}} \rangle_{0,\partial\mathcal{T}_h^{\mathrm{B}}}, \tag{5.29a}$$

$$\mathbb{D}_2|_{ij} := \langle \ell_\gamma \boldsymbol{\alpha}_\gamma^{\boldsymbol{\tau}} \boldsymbol{\nabla}_{\boldsymbol{\tau}} \hat{q}_j^\gamma, \boldsymbol{\nabla}_{\boldsymbol{\tau}} \hat{q}_i^\gamma \rangle_{0,\mathcal{F}_h^\gamma} + \langle a_\gamma \hat{q}_j^\gamma, \hat{q}_i^\gamma \rangle_{0,\partial\mathcal{T}_h^\gamma}, \tag{5.29b}$$

where the subblock-matrices $\mathbb{D}_1$ (*resp.* $\mathbb{D}_2$) denotes the interaction of the interior-based DOFs $\hat{q}_h^\Gamma$ (*resp.* q_h^Γ) with itself.

5.6.2 A2 - Time-Discretization

Here, we apply the BDF scheme for the discretization of the time derivative. Let us consider the time domain $\mathrm{T} = [0, T]$, which is partionned into N sub-intervals $\mathrm{T}^n = [t^{n-1}, t^n]$ such that, $0 = t^0 < \cdots < t^n < \cdots < t^N = T$, and $\triangle t^n = t^n - t^{n-1}$ refers to the length of each sub-intervals. At the time-step t^n, the linear system (5.23) can be expressed into the following equivalent block-matrix system :

$$\begin{bmatrix} \mathbb{A} + \mathbb{B}^{(\varepsilon)} & \mathbb{C} \\ \mathbb{C}^{(\varepsilon)} & \mathbb{E} + \mathbb{D} \end{bmatrix} \begin{bmatrix} \mathrm{P}_h^n \\ \hat{\mathrm{P}}_h^n \end{bmatrix} = \begin{bmatrix} \mathrm{F}_h \\ \hat{\mathrm{F}}_h \end{bmatrix} + \frac{1}{\triangle t} \begin{bmatrix} \mathbb{A}\mathrm{P}_h^{n-1} \\ \mathbb{E}\hat{\mathrm{P}}_h^{n-1} \end{bmatrix}, \tag{5.30}$$

where Υ^n denotes the temporal approximation, at the time-level n, for any vector-valued function $\Upsilon(\cdot, t^n)$. From (5.30), the interior-based unknowns can be expressed in terms of skeleton-based unknowns such that,

$$\mathrm{P}_h^n = (\mathbb{A} + \mathbb{B}^{(\varepsilon)})^{-1} \left[\mathrm{F}_h + \frac{1}{\triangle t} \mathbb{A}\mathrm{P}_h^{n-1} - \mathbb{C}\hat{\mathrm{P}}_h^n \right]. \tag{5.31}$$

Thus, by inserting the relation (5.33) into the second line of the matrix system (5.30), we determine the reduced system :

$$\mathbb{K}\hat{\mathrm{P}}_h^n = \mathbb{F}^{n-1}, \tag{5.32}$$

which depends only on the skeleton-based unknowns. The sparse matrix $\mathbb{K}$ on the left-hand-side denotes the Schur complement is defined as

$$\mathbb{K} := \left[\mathbb{E} + \mathbb{D} - \mathbb{C}^{(\varepsilon)} (\mathbb{A} + \mathbb{B})^{-1} \mathbb{C} \right], \tag{5.33}$$

and the auxiliary vector of the right-hand side is given by

$$\mathbb{F}^{n-1} := \hat{\mathrm{F}}_h + \frac{1}{\triangle t} \mathbb{E}\hat{\mathrm{P}}_h^{n-1} - \mathbb{C}^{(\varepsilon)} (\mathbb{A} + \mathbb{B})^{-1} \left[\mathrm{F}_h + \frac{1}{\triangle t} \mathbb{A}\mathrm{P}_h^{n-1} \right]. \tag{5.34}$$

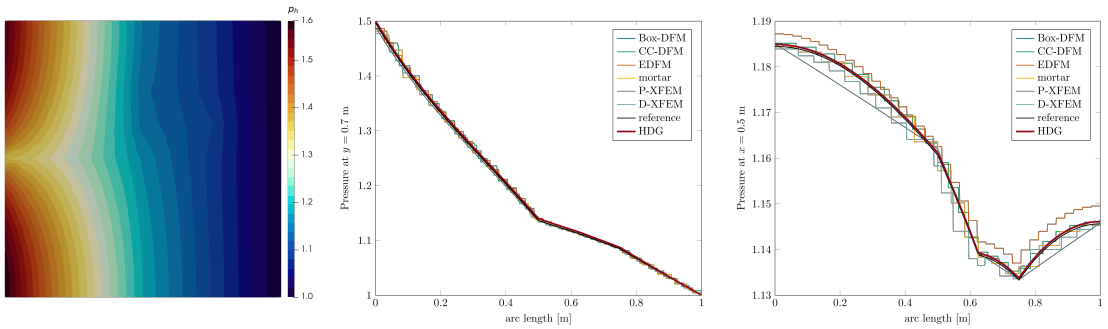
Finally, the interior-based unknowns P_h^n can be reconstructed at the element-level using the element reconstruction procedure of (5.31), for the time-step t^n.

5.7 Appendix B - Supplementary data

5.7.1 B1 - Regular fracture network

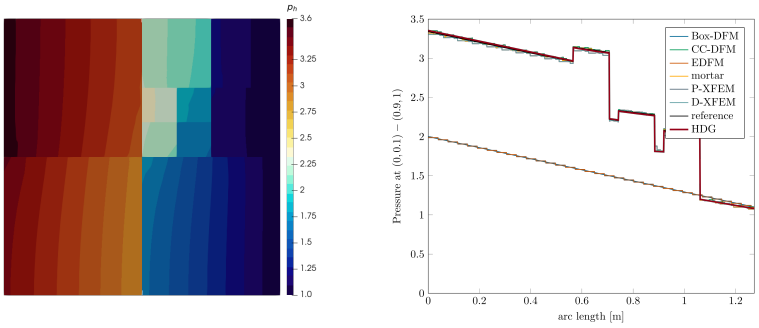
This appendix provided supplementary results for the regular test case of the benchmark [Flemisch 2018]. We refer the reader to this study for a complete description of the geometry and the boundary conditions.

Here, the grid is made up of 1024 rectangular elements and 112 fracture elements. The reference solution is computed with mimetic finite differences using a very fine mesh. The bulk permeability is set to $\boldsymbol{\alpha}_{\mathrm{B}} = \boldsymbol{I}_2$ m^2. The tangential and normal fracture permeability are given by $\boldsymbol{\alpha}_\gamma^{\boldsymbol{\tau}} = \alpha_\gamma^{\boldsymbol{n}} = \alpha_\gamma$ m^2. Two configurations are considered : (i) a permeable case for $\alpha_\gamma = 10^4$, and (ii) an impermeable case for $\alpha_\gamma = 10^{-4}$. Figures 5.22 and 5.23 display the discrete pressure for both the permeable and impermeable cases, respectively. We recognize that the results are in agreement with the existing solutions of the benchmark study.



(a) (b) (c)

FIGURE 5.22 – Permeable case - Illustration of the pressure distribution for the regular benchmark (a) with structured rectangular grid ($h = 1/32$) and linear functions. Comparison of the pressure profile with referenced solution of the literature along the lines $y = 0.7$ (b) and $x = 0.5$ (c).

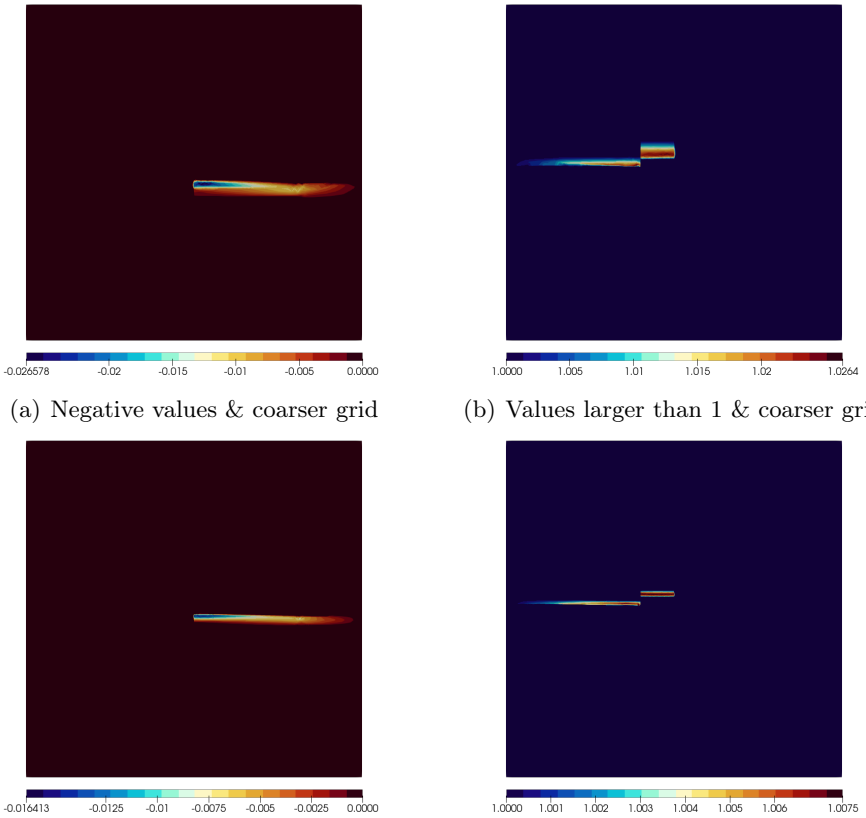


(a) (b)

FIGURE 5.23 – Impermeable fractures - Illustration of the pressure distribution for the regular benchmark (a) with structured rectangular grid ($h = 1/32$) and linear functions. Comparison of the pressure profile with referenced solution of the literature along the lines $(0, 0.1) - (0.9, 1)$ (b).

5.7.2 B2 - Values out of the bounds for the example C

Figure 5.24 illustrates the values out of the bounds for the FE solution, which appears mainly along the central x-axis.



(a) Negative values & coarser grid

(b) Values larger than 1 & coarser grid

(c) Negative values & finer grid

(d) Values larger than 1 & finer grid

FIGURE 5.24 – Appendix B - Illustration of the values out of the bounds for the FE solution. The values are negatives on the left and larger than 1 on the right.

5.7.3 B3 - Coordinates of the fracture networks of example D

The exact coordinates of the three fracture networks are listed below.

Case 1			**Case 2 and $M = 2$**			**Case 2 and $M = 4$**					
pt	x [m]	y [m]	pt	x [m]	y [m]	pt	x [m]	y [m]	pt	x [m]	y [m]
i_1	0.75	1	i_1	0.5	1	i_1	0.5	1	i_5	1	1
i_2	1.25	0	i_2	1	0	i_2	0.75	0	i_6	1.25	0
			i_3	1	1	i_3	0.75	1	i_7	1.25	1
			i_4	1.5	0	i_4	1	0	i_8	1.5	0

TABLE 5.2 – Test D - Coordinates of the numbered points of the fractures in the bulk.

Chapitre 6

Conclusion générale et Perspectives

Le travail retranscrit dans ce manuscrit de thèse se focalise sur la modélisation numérique de l'écoulement et du transport de masse dans les milieux poreux fracturés. En pratique, ces phénomènes sont formalisés par des lois de conservation, dont la résolution nécessite *a priori* une réflexion sur la mise en place de méthodes de discrétisation adaptées. Ce mémoire aborde précisément le développement et l'implémentation d'une classe d'approximation visant à améliorer la simulation des transferts de masse dans les milieux poreux fortement hétérogènes, anisotropes et fracturés. La première difficulté soulevée concerne l'hétérogénéité de la roche, dont l'amplitude peut varier de plusieurs ordres de grandeur d'une région à l'autre de l'espace. Bien que la solution du problème est continue, l'ambivalence avec la discontinuité naturelle de la roche est susceptible d'altérer la qualité des résultats (e.g., oscillations parasites, violation du principe du maximum). Lorsque ces discontinuités structurelles sont des fractures, leurs présences intrinsèque affectent sévèrement les différents processus physiques en intéraction. Pour ce cas particulier, les failles sont alors décrites individuellement dans la matrice poreuse comme des *interfaces* de dimension–$(d-1)$ entre les sous-domaines. Le transport de masse soulève principalement des difficultés concernant la nature mathématique des EDPs à résoudre, qui peuvent dégénérer dans le domaine d'étude. Ce phénomène physique est représenté par une équation de convection-diffusion-réaction, dans laquelle la prédominance du terme de convection (nombre de Péclet) peut entrainer des conséquences sur l'approximation de la solution (i.e., discontinuité à l'interface hyperbolique-elliptique). Parmi la grande variété de méthodes de discrétisation existantes dans la littérature, nous nous sommes focalisés sur la famille des méthodes Hybride de Galerkine Discontinues (HDG). Plus précisément, nous avons adopté le formalisme primal de Pénalisation Intérieure Hybride de Galerkine Discontinue (H-IP) pour le traitement unifié des opérateurs de diffusion et de convection. Cette classe fut introduite pour la première fois en 2009 dans le cadre des problèmes elliptiques du second-ordre. Depuis lors, elles connaissent un succès certain auprès de la communauté scientifique pour la résolution des EDPs (non)-linéaires (e.g., mécanique des fluides, mécanique du solide, propation des ondes). Grâce à son formalisme flexible, la méthode H-IP bénéficie d'un certain nombre d'avantages : elligibilité à la condensation statique, flexibilité aux mailles complèxes, approximation polynomiale élevé, et flux localement conservatifs. Par ailleurs, elles facilitent le traitement des hétérogénéités et anisotropies du milieu, tout en garantissant une convergence optimale de la solution. Cependant, la validité et la robustesse de la méthode H-IP repose sur la sélection du paramètre de pénalisation, qui va directement impacter la précision, la stabilité, et la convergence du schéma numérique.

Les différents développements et résultats de ce travail de recherche s'orientent autour de trois princiaples thématiques, que nous trouverons ci-après un bref récapitulatif.

6.1 Thématiques abordées

- **Anisotopies et hétérogénéités**
  Dans un premier temps, nous avons présenté au chapitre 2 le formalisme général de la méthode H-IP pour les équations elliptiques du second-ordre hétérogène et anisotrope. Nous débutons par la description générale de la méthode et de ses trois variantes à savoir, Incomplète (I-HIP), Non-Symétrique (N-HIP) et Symétrique (S-HIP). Puis, nous étudions le caractère bien-posé du problème discrêt, avant de procéder aux estimations de l'erreur en employant une fonction de pénalisation de la forme $\mathcal{O}(1/h^{\delta})$. Cette analyse révèle que le schéma numérique converge toujours de manière optimale suivant (i) la norme-énergie pour toutes les valeurs de $\delta \geq 0$ et $\varepsilon = \{0, \pm 1\}$, et (ii) la norme-L^2 pour $\varepsilon = 1$, indépendamment du paramètre δ. Cependant, pour $\varepsilon = \{-1, 0\}$, ce critère de convergence optimale n'est atteint que pour $\delta \geq 2$. En particulier, dans le cas d'une grille $\boldsymbol{\alpha}$–orthogonale, nous avons obtenu la superconvergence du schéma H-SIP sans l'usage d'une procédure de post-traitement. En pratique, le formalisme H-IP fait parti de la grande famille des méthodes IP, dont nous évaluons les performances numérique au chapitre 3. Nous distinguons les formes Hybride (H-IP), Embedded (E-IP), et Pondéré-IP (W-IP), dont la construction s'articule autour d'un mécanisme de pénalisation unique, qui est dépend de la diffusivité du milieu. Néanmoins, ces trois classes d'approximation demeurent fondamentalement différentes. En effet, les deux méthodes *Hybrides* (H-IP et E-IP) bénéficient de la procédure de condensation statique, permettant de réduire de moitié ou du quart les temps CPU par rapport au schéma W-IP. Le contraste est d'autant plus flagrant que les expérimentations numériques révèlent la supériorité des deux versions Hybrides sur tous les aspects numériques. Enfin, nous devons souligner la propriété d'équivalence totale entre les deux variantes H-IIP et W-IIP, qui établit de ce fait une connexion entre les familles de méthodes IP.

- **Diffusion localement évanescente**
  Au chapitre 4, nous avons étendu la méthode H-IP aux problèmes elliptiques dégénérés (du second-ordre). Cette méthodologie enrichie s'appuie sur une stratégie de pénalisation adaptative dépendante du nombre de Péclet. Dans la région elliptique, nous favorisons une technique basée sur la fonction de *Scharffeter-Gummel*, tandis qu'une stratégie *upwind* couplée à une correction de Lax-Friedrichs est employée dans la région hyperbolique. En fonction, de la prédominance des mécanismes de diffusion et convection, la méthode H-IP permet de traiter de façon efficace plusieurs situations physiques distinctes, à savoir non-dégénérée, entièrement dégénérée et localement dégénérée. L'étude du caractère bien-posé du schéma numérique permet de vérifier l'existence et l'unicité de la solution discrète. À travers plusieurs expérimentations, nous avons vérifié la flexibilité et la convergence optimale du formalisme H-IP, pour les trois régimes physiques mentionnées précédemment. Nous remarquons que la méthode capture aisément la discontinuité de la solution le long de l'interface *hyperbolique-elliptique*.

- **Traitement des fractures discretes**
  Finalement, dans le chapitre 5, nous avons adapté le formalisme H-IP à la résolution des écoulements dans les milieux poreux fracturés. Plus précisément, nous avons couplé la

classe H-IP dans la matrice poreuse avec la méthode des Éléments Finis dans le réseau de fractures discrètes. Il a été montré que cette méthodologie dispose d'un intérêt certain pour la résolution, puisqu'elle bénéficie d'un système linéaire compacte dont les degrés de liberté sont localisés sur les interfaces des éléments. Bien que l'analyse des estimations d'erreurs ne soit pas exposée, le schéma HIP/EF [1] converge de manière optimale. La flexibilité de ce formalisme permet de traiter les géométries les plus complèxes, telles que les réseaux de fractures fortement interconnectés et les failles partiellement immergées.

6.2 Perspectives de recherche

Ce travail de thèse ne constitue que la première étape d'un projet de développement global d'un modèle numérique robuste, visant à représenter de manière unifiée différents processus physiques dans les milieux poreux et fracturés. À cet effet, plusieurs pistes de recherche sont envisagées pour la poursuite des travaux initiés dans ce manuscrit.

1. Le formalisme HDG est étroitement lié à la classe des Éléments Finis standard et la famille des méthodes DG. Bien que ces dernières soient davantage employées pour la résolution des phénomènes dans les milieux poreux, elles possèdent certaines limitations. D'une part, la méthode EF ne dispose pas de propriétés de conservation locale de la masse, et ne traite pas efficacement l'opérateur hyperbolique. D'autre part, les schémas DG standards sont moins compétitifs que leurs homologues Hybrides. En général, le formalisme HDG préconise une définition de la trace discrète plus riche que la plupart des méthodes DG stabilisées. Nous pourrions reformuler la méthode H-IP dans le cadre des méthodes DG, à la fois pour l'écoulement et le transport de masse. Cette étude permettrait d'établir une nouvelle formulation de type Pénalisation Intérieure, qui serait *totalement équivalente* à la méthode H-IP présentée dans cette thèse (i.e., $u_h^{\text{H-IP}} \equiv u_h^{\text{IP}}$). À l'instar des schémas IP traditionnels, ce nouveau formalisme introduit la pénalisation de la composante normale du flux $[\![\boldsymbol{\alpha}\boldsymbol{\nabla}_h u_h]\!]$, indépendamment de la pénalisation de la variable d'état $[\![u_h]\!]$. Une analyse théorique de cette nouvelle classe d'approximation permettrait de garantir l'existence, l'unicité et la convergence de ce schéma numérique, au même titre que la méthode H-IP.

2. Des passerelles sont envisageables pour rendre la méthode EF localement conservative, en incluant la pénalisation du saut-DG de la dérivée normale de la variable d'état u_h. Cette étude nécessite *a priori* la connaissance de la pénalisation du flux $[\![\boldsymbol{\alpha}\boldsymbol{\nabla}_h u_h]\!]$, provennant de la réduction IP de la méthode H-IP. En restaurant la conservation locale du bilan massique, il devient alors possible de recupérer la propriété de superconvergence $(k+2)$ du schéma H-SIP pour une valeur de τ_{opt} donnée.

3. En parallèle, il semble pertinent d'inclure un opérateur de projection de la variable d'état u_h, afin d'améliorer la résolution des schémas H-IP. Cette démarche s'appuie sur l'utilisation d'un opérateur de *lifting* local, $\boldsymbol{R}_h^{\partial E} : L^2(\partial E) \to \mathcal{P}_k(E)^d$, dans le mécanisme de pénalisation. Il a été montré que l'ajout de cette composante établissait

1. Dans le chapitre 5, nous couplons la méthode H-IP dans la matrice poreuse avec le schéma EF (ou CG en anglais) dans les fractures pour la simulation de l'écoulement.

des connexions spécifiques entre les formalismes Hybride de Pénalisation Intérieure (H-IP), Hybride Locale de Galerkine Discontinue (H-LDG) [Cockburn 2009b] et Hybride d'Ordre-Élevé (HHO) [Boffi 2018] (voir aussi [Cockburn 2016b]).

4. Dans les applications consacrées au couplage de l'écoulement et du transport de masse, l'évaluation précise des flux qui lie ces deux phénomènes est essentielle. En pratique, nous appliquons une procédure locale de traitement *a posteriori* du flux discrèt $\boldsymbol{\sigma}_h = -\kappa \boldsymbol{\nabla}_h u_h$, afin de respecter les propriétés mathématiques du problème de départ. L'idée consiste à reconstruire un flux $\boldsymbol{\sigma}_h^*(\boldsymbol{u}_h)$, en s'appuyant sur les caractéristiques de l'espace $H(\mathrm{div}, \mathcal{T}_h) \in [L^2(\mathcal{T}_h)]^d$. L'intérêt de cet espace discret réside dans la continuité de sa composante normale, garantissant de ce fait l'unicité de $\boldsymbol{\sigma}_h^*(\boldsymbol{u}_h)$ sur l'ensemble du squelette de la partition.

5. Dans le contexte des milieux poreux fracturés, le couplage des méthodes H-IP et EF ne représente qu'une seule possibilité, parmi les nombreuses combinaisons imaginables pour la modélisation de l'écoulement. En réalité, le choix de la discrétisation dans la roche et les fractures reste libre, suivant les besoins de l'utilisateur. Il semble donc intéressant de procéder à des combinaisons entre les formulations H-IP, IP standard (ou pondérée), et/ou EF pour le problème couplé *matrice-fractures*. À court terme, il convient d'analyser le caractère *bien-posé* de ces différentes combinaisons numériques, afin de garantir l'unicité et l'existence d'une solution discrète.

6. La modélisation du transport de masse dans les milieux poreux fracturés correspond à la suite logique des différents développements exposés précédemment. Pour ce phénomène physique, la pénalisation adaptative que nous avons introduite dans le cadre des problèmes dégénérés peut jouer un rôle primordiale pour la résolution. En effet, le long de l'interface *matrice-fracture*, les conditions de couplage introduisent un mécanisme *upwind* pour caractériser l'échange de flux convectif entre les domaines.

7. Le couplage Écoulement/Transport dans les milieux poreux fracturés, incluant un contraste de masse volumique, offre de nombreuses possibilités. À l'échelle locale, de l'île de La Réunion, le phénomène d'intrusion d'eau salée dans les aquifères littoraux est un processus physique complèxe et non-linéaire. La modélisation requiert une méthode robuste et stable pour limiter les oscillations, les diffusions, et les dispersions numériques. En général, l'emploi des méthodes de Pénalisation Intérieure nécessitent une analyse de la compatibilité pour éviter les pertes de précision lors du couplage de l'écoulement et du transport. Cependant, il est possible de contourner cette restriction en employant une méthode H-IP pour la résolution. Finalement, l'application de ces mécanismes couplés en présence de fractures reste encore à ce jour très peu usitée. Elle constitue, de ce fait, une piste d'ores et déjà envisable dans la poursuite de cette thèse.

Toutes ces thématiques, constituent à ce jour les principaux axes auxquelles se rattachent nos travaux de recherches.

Annexes

Annexe A

Modèles de fractures discrètes

Dans cette annexe, nous proposons d'étendre les mécanismes d'écoulement et de transport de masse dans un milieu poreux fracturé, en employant la description *codimensionnelle* pour représenter le réseau de fractures. Avant d'établir la réduction du modèle DFM, nous établissons une courte synthèse non-exhaustive des différentes représentations de fractures de la littérature.

A.1 Modèles conceptuels pour les systèmes fracturés

La nécessité d'établir une représentation adéquate du réseau de fractures passe par la construction de différentes classes de modèles conceptuels (voir les travaux de [Tatomir 2012, Berre 2019, Tarhuni 2021] pour une description plus détaillée). En fonction du degré de fracturation du milieu et de l'echelle d'observation, on peut répertorier le milieu fracturé suivant l'une des deux classes suivantes : les modèles de fractures *continues* et *discrètes*. Dans la première approche, les fractures sont moyennées par différentes techniques d'homogénisation ou de mise à l'echelle – du terme *upscaling* en anglais ([Wen 1996, Christie 1996, Farmer 2002]), tandis que chaque fractures sont décrites individuellement dans la seconde description, [Dietrich 2005, Berre 2019].

A.1.1 Modèles continus

L'approche *continue*, ou implicite, est particulièrement adaptée pour les problèmes de grandes échelles (m-km), où l'identification explicite de toutes les fractures est fortement limitée par la dimension du système, [Farmer 2002]. Durant la fin du XXème siècle, cette méthodologie reçoit une attention particulière dans le secteur industriel, dont les modèles dotés d'une seule perméabilité (single-permeability) ou d'une double porosité (double-porosity) furent largement exploités dans les simulateurs commerciaux [Tatomir 2012]. Par ailleurs, la résolution s'effectue par des méthodes numériques standards, puisque le réservoir fracturé est matérialisé par un continuum. Malgré cet avantage certain, déterminer le plus fidèlement les propriétés du milieu nécessite d'imposer des conditions particulières, qui sont souvent dépendantes des hypothèses données au VER. En particulier, la distribution, l'orientation, la longueur, l'ouverture (allant du micromètre au centimètre), et la rugosité des fissures sont autant de paramètres physiques à prendre en compte pour caractériser les milieu poreux fracturé. Pour une présentation plus générale de ces différentes notions, nous référons les travaux de [Tatomir 2012, Odsæter 2018] et leurs références qui décrivent en détail la variété de modèles continus développés dans la littérature.

A.1.2 Modèles discrèts

La représentation *discrète*, ou explicite, construit une description plus détaillée du milieu poreux fracturé. Dans cette approche, toutes les fractures du système sont identifiées individuellement, chacune disposant de caractéristiques qui leurs sont propres (longueur, épaisseur, perméabilité), [Adler 1999]. Généralement, on décompose les modèles explicites suivant l'une des trois catégories de repésentation suivantes :

- **Fracture simple et discrète** – Les mécanismes d'écoulement et de transport de masse sont étudiés dans un nombre très limité de fractures bien-définies [Tatomir 2012].
- **Réseau de fractures discrètes (DFN)** – Dans cette représentation, l'écoulement du fluide se déroule uniquement dans le réseau de fractures interconnectées [Fumagalli 2019, Odsæter 2018], tandis que les échanges massiques entre la matrice et les fractures sont négligés, [Berre 2019]. En particulier, la matrice poreuse se comporte comme un matériau imperméable dans la représentation DFN.
- **Matrice-Fracture discrète (DFM)** – Les mécanismes d'écoulement et de transport interviennent à la fois dans la matrice et le réseau de fractures, [Fumagalli 2013]. Dans l'approche la plus simple, mais également la plus ancienne, les fractures et la matrice poreuse sont décrites suivant la même dimension spatiale - ou modèle *équidimensionnel* (voir [Martin 2004] et la section A.2 pour la description mathématique). La résolution s'opère au même titre qu'une Décomposition de Domaine (DD) en maillant très finement à l'intérieur des fractures. Cependant, la précision de ce modèle se révèle être un challenge de taille : *(i)* qui requiert une quantité importante de stockage mémoire pour des réseaux de fractures complexes, et *(ii)* pouvant engendrer des temps de calculs (CPU) relativement longs, [Tatomir 2012, Odsæter 2018]. Une alternative consiste à modifier la description géométrique de la fracture, qui est assimilée à un objet de dimension-$(d-1)$ immergée dans la matrice poreuse [Martin 2005, Tunc 2012, Angot 2009]. Par exemple si la matrice est en deux dimensions (*resp.* en trois dimensions) d'espace, la fracture sera assimilée à un segment droite (*resp.* une surface plane). On parle alors d'approche *codimensionnelle*, ou de dimension-mixte [Nordbotten 2019].

L'identification des fractures dominantes par un modèle DFM n'exclut pas son couplage avec un modèle continu pour la prise en compte des micro-fractures. On parle dans ce cas d'approche *hybride*. Bine qu'elle ne soit pas appliquée dans ce manuscrit, la représentation hybride s'avère particulièrement pertinente pour représenter un milieu fracturé fortement hétérogène et anisotrope.

A.2 La représentation équidimensionnelle

Avant d'aborder la description du modèle codimensionnel, développé à la section A.3, nous devons introduire les principales notations et contraintes liées à la géométrie et aux variables du problème. Pour faciliter la compréhension, nous envisagerons la configuration la plus simple, à savoir le cas d'un domaine Ω entièrement découpé en deux parties par une

unique fracture Ω_f définit par :

$$\Omega_f := \left\{ \boldsymbol{x} \in \mathbb{R}^d \mid \boldsymbol{x} = \boldsymbol{s} + r\boldsymbol{\eta},\ \boldsymbol{s} \in \gamma,\ r \in \left(-\frac{\ell}{2}, \frac{\ell}{2}\right) \right\}, \tag{A.1}$$

où ℓ désigne la largeur constante de la fracture Ω_f, et $\boldsymbol{\eta}$ le vecteur unitaire normale pointant de l'extérieur de Ω_1 vers l'intérieur de Ω_2. Ici, γ représente l'interface qui passe par l'axe centrale de la fracture Ω_f (voir Figure A.1 pour la représentation géométrique du domaine et des diverses notations). La matrice poreuse résultante se décompose suivant les sous-régions

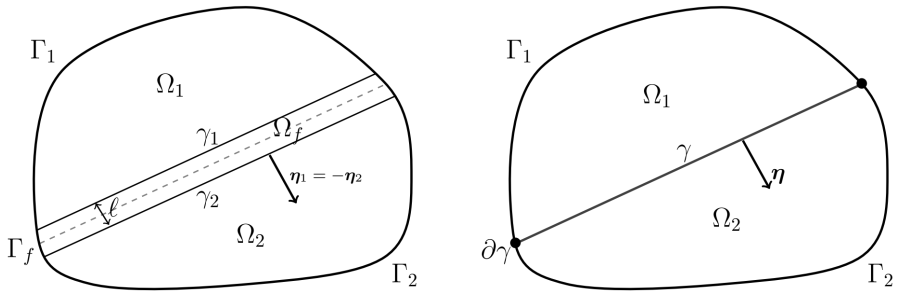


FIGURE A.1 – Description géométrique du domaine fracturé pour la représentation équi-dimensionnelle (à gauche) et co-dimensionnelle (à droite).

Ω_1 et Ω_2 tel que,

$$\Omega \backslash \bar{\Omega}_f = \Omega_1 \cup \Omega_2, \quad \text{et} \quad \Omega_1 \cap \Omega_2 = \emptyset.$$

Dans la suite, nous désignons par Γ_i la frontière du sous-domaine $\Omega_i = \partial\Omega_i \cap \partial\Omega$ qui est commune à la bordure extérieure de Ω, pour $i = 1, 2, f$. L'interface entre les sous-régions Ω_j et Ω_f, représentée par γ_j, se note

$$\gamma_j := \partial\Omega_j \cap \partial\Omega_f \cap \Omega, \quad j = 1, 2.$$

Désormais, nous supposerons que la fracture est remplie de débris, qui sont issus des déplacements de matières ou des réactions chimiques [Martin 2005, Angot 2009]. Cette hypothèse, couramment utilisée en hydrogéologie, permet d'assimiler la fracture à une matrice poreuse de faible dimension immergée dans la roche. De ce fait, les lois physiques gouvernant les échanges dans cette portion du domaine sont régis par les mêmes équations introduites à la section 1.2, i.e., l'écoulement du fluide et le transport de solutés. À partir de la section 1.3.1, nous définissons par $u_i := u_{|\Omega_i}$ et $\boldsymbol{\sigma}_i := \boldsymbol{\sigma}_{|\Omega_i}$ les restrictions de la variable d'état et du flux total $\boldsymbol{\sigma} = -\boldsymbol{\alpha}\nabla u + \boldsymbol{\beta} u$ dans la sous-région Ω_i, pour $i = 1, 2, f$. Puisqu'aucune confusion ne peut survenir, les fonctions $\boldsymbol{\alpha}_i$, $\boldsymbol{\beta}_i$, φ_i et f_i représentent également les restrictions de $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, φ et f dans Ω_i, $i = 1, 2, f$, respectivement. En utilisant les notations précédentes, nous pouvons réécrire le système (1.13) sous la formulation mixte suivante :

$$\boldsymbol{\sigma}_i = -\boldsymbol{\alpha}_i \nabla u_i + \boldsymbol{\beta} u_i \quad \text{dans } \Omega_i \times [0, T],\ i = 1, 2, f, \tag{A.2a}$$

$$\varphi_i \frac{\partial u_i}{\partial t} + \nabla \cdot \boldsymbol{\sigma}_i = f_i \quad \text{dans } \Omega_i \times [0, T],\ i = 1, 2, f, \tag{A.2b}$$

accompagnée des conditions de couplage (A.3) le long des interfaces matrice-fracture

$$\begin{aligned} u_j &= u_f && \text{sur } \gamma_j \times [0,T], \quad j=1,2, \\ \boldsymbol{\sigma}_j \cdot \boldsymbol{\eta} &= \boldsymbol{\sigma}_f \cdot \boldsymbol{\eta} && \text{sur } \gamma_j \times [0,T], \quad j=1,2, \end{aligned} \tag{A.3}$$

et munie de la condition initiale et des conditions aux limites pour $i = 1, 2, f$

$$\begin{aligned} u_i &= u_{i,0} && \text{dans } \Omega_i \times \{0\}, \\ u_i &= \bar{u}_i && \text{sur } \Gamma_i^D \times [0,T], \\ \boldsymbol{\sigma}_i \cdot \boldsymbol{n}_i &= \bar{\sigma}_i && \text{sur } \Gamma_i^N \times [0,T], \end{aligned} \tag{A.4}$$

avec Γ_i^D et Γ_i^N les portions de Γ_i qui correspondent aux frontières de Dirichlet et de Neumann, respectivement. Notons que, l'écriture mixte (A.2) peut être vue comme la représentation équidimensionnelle du milieu fracturé [Martin 2004]. La *continuité* des variables u et $\boldsymbol{\sigma}$ est imposée par les conditions de transmission (A.3) le long de l'interface γ_j, pour $j = 1, 2$.

Dans la prochaine section, nous présentons la dérivation du modèle réduit de l'équation de conservation (A.2) associé au problème couplé matrice-fracture. Pour une démonstration plus rigoureuse, nous mentionnons les travaux de [Martin 2005, Tunc 2012, Angot 2009, Fumagalli 2012, Fumagalli 2013], qui établissent différentes procédures de réduction.

A.3 Le modèle réduit

Le modèle codimensionnel consiste en une procédure de réduction des mécanismes agissant dans la faille Ω_f dans l'interface equivalente γ. Elle comporte deux éléments majeurs :

1. Une loi de comportement moyenne suivant la direction longitudinale de la fracture ;
2. Des conditions de raccordement (ou couplage) entre la matrice et la fracture.

Notons que dans les sous-régions Ω_1 et Ω_2, les processus physiques de l'équation (A.2) restent inchangés. À l'intérieur de la fracture Ω_f, nous supposerons que le tenseur de perméabilité $\boldsymbol{\alpha}_f$ vérifie les mêmes conditions de régularité de la matrice poreuse (1.14), qui dispose d'une strucutre bloc-diagonale représentée par,

$$\boldsymbol{\alpha}_f = \begin{bmatrix} \alpha_f^{\boldsymbol{n}} & 0 \\ 0 & \boldsymbol{\alpha}_f^{\boldsymbol{\tau}} \end{bmatrix}, \tag{A.5}$$

où les constantes $\alpha_f^{\boldsymbol{n}} : \Omega_f \to \mathbb{R}$ et $\boldsymbol{\alpha}_f^{\boldsymbol{\tau}} : \Omega_f \to \mathbb{R}^{(d-1),(d-1)}$ désignent les perméabilités normale et tangentielle le long de chaque section transversale de Ω_f, respectivement. On admettra également que le coefficient de stockage φ_f est constant dans Ω_f. De la même manière que [Martin 2005], nous définissons respectivement les opérateurs gradient et divergence tangentielle $\boldsymbol{\nabla}_{\boldsymbol{\tau}} := \boldsymbol{T}\nabla$ et $\boldsymbol{\nabla}_{\boldsymbol{\tau}} \cdot := \boldsymbol{T} : \nabla$, associés aux matrices de projection tangente et normale $\boldsymbol{T} := \boldsymbol{I}_d - \boldsymbol{N}$ et $\boldsymbol{N} := \boldsymbol{\eta} \otimes \boldsymbol{\eta}$. Nous rapportons les variables moyennes dans Ω_f de la variable d'état u_γ, la composante tangentielle de la vitesse $\boldsymbol{\sigma}_\gamma$ et le terme source f_γ,

$$u_\gamma = \frac{1}{\ell} \int_{-\ell/2}^{\ell/2} u_f \, d\boldsymbol{r}, \quad \boldsymbol{\sigma}_\gamma = \int_{-\ell/2}^{\ell/2} \boldsymbol{T} \cdot \boldsymbol{\sigma}_f \, d\boldsymbol{r}, \quad \text{et} \quad f_\gamma = \int_{-\ell/2}^{\ell/2} f_f \, d\boldsymbol{r},$$

pour tous points $\boldsymbol{r}$ appartenant à Ω_f.

A.3.1 Réduction de l'équation de conservation

Nous débutons par la décomposition de l'équation de conservation (A.2b) et du flux total $\boldsymbol{\sigma}$ dans le plan généré par l'interface γ, i.e., suivant ses composantes normale et tangentielle :

$$\text{Conservation :} \qquad \varphi_f \frac{\partial u_f}{\partial t} + \left[\boldsymbol{\nabla_n} \cdot \boldsymbol{\sigma}_f + \boldsymbol{\nabla_\tau} \cdot \boldsymbol{\sigma}_f\right] = f_f, \tag{A.6a}$$

$$\text{Flux total :} \qquad \boldsymbol{\sigma}_f = \boldsymbol{\sigma}_{f,\boldsymbol{\eta}} + \boldsymbol{\sigma}_{f,\boldsymbol{\tau}}, \tag{A.6b}$$

avec $\boldsymbol{\sigma}_{f,\boldsymbol{\eta}} = \boldsymbol{N}\boldsymbol{\sigma}_f$ et $\boldsymbol{\sigma}_{f,\boldsymbol{\tau}} = \boldsymbol{T}\boldsymbol{\sigma}_f$. En intégrant l'équation (A.6a) et la partie tangentielle du flux total sur chaque section transversale $[-\ell/2, \ell/2]$, nous obtenons l'expression du comportement moyen de la loi de conservation dans la fracture γ :

$$\varphi_f \ell \frac{\partial u_\gamma}{\partial t} + \boldsymbol{\nabla_\tau} \cdot \boldsymbol{\sigma}_\gamma = f_\gamma + (\boldsymbol{\sigma}_1^*|_\gamma - \boldsymbol{\sigma}_2^*|_\gamma) \cdot \boldsymbol{\eta} \qquad \text{dans } \gamma \times [0,T], \tag{A.7a}$$

$$\boldsymbol{\sigma}_\gamma = -\ell \boldsymbol{\alpha}_f^{\boldsymbol{\tau}} \boldsymbol{\nabla_\tau} u_\gamma + \ell \boldsymbol{\beta}_\Gamma u_\gamma \qquad \text{dans } \gamma \times [0,T], \tag{A.7b}$$

où la composante normale de la vitesse $\boldsymbol{\sigma}_j^*|_\gamma \cdot \boldsymbol{\eta} = -\boldsymbol{\alpha}_j \nabla u_j|_\gamma \cdot \boldsymbol{\eta} + (\boldsymbol{\beta}_j|_\gamma \cdot \boldsymbol{\eta}) u_j^*|_\gamma$. La première relation (A.7a) est obtenue en imposant la continuité de la composante normale du flux total (A.3) sur γ_j, pour $j = 1, 2$. Le saut du flux total $(\boldsymbol{\sigma}_1^*|_\gamma - \boldsymbol{\sigma}_2^*|_\gamma) \cdot \boldsymbol{\eta}$ décrit la contribution du flux massique provenant de la matrice dans l'écoulement de la fracture γ. La variable u^* se définit de la façon suivante,

$$u_j^*|_\gamma = \begin{cases} u_j|_\gamma & \text{si } (\boldsymbol{\beta}_j|_\gamma \cdot \boldsymbol{\eta}) > 0 \\ u_\gamma & \text{si } (\boldsymbol{\beta}_j|_\gamma \cdot \boldsymbol{\eta}) < 0 \end{cases} \quad \text{pour } j = 1, 2, \tag{A.8}$$

est évaluée sur la base d'un schéma *upwind* du premier ordre ([Fumagalli 2013]). Le dernier terme dans la relation (A.7b) provient de l'intégrale $\int_{-\ell/2}^{\ell/2} \boldsymbol{T}\boldsymbol{\beta}_f u_f \, d\boldsymbol{r} \approx \boldsymbol{\beta}_\gamma \int_{-\ell/2}^{\ell/2} u_f \, d\boldsymbol{r}$, qui est une condition suffisamment régulière, compte tenue de la faible dimension transversale de la fracture par rapport à la dimension du milieu.

A.3.2 Conditions de couplage

Pour fermer le problème couplé matrice-fracture, nous devons relier les différentes variables du problème ensemble (voir également [Fumagalli 2012, Formaggia 2014, Berge 2019, Fumagalli 2021] pour plus de détails). Tout d'abord, nous évaluons la composante normale du flux total $\boldsymbol{\sigma}_{f,\boldsymbol{\eta}} = -\alpha_f^{\boldsymbol{n}} \boldsymbol{\nabla_n} u_f + \boldsymbol{N}\boldsymbol{\beta}_f u_f$ sur chaque section transversale de Ω_f. En intégrant suivant la direction normale de la fracture, nous obtenons,

$$\begin{aligned} \int_{-\ell/2}^{\ell/2} \boldsymbol{\sigma}_{f,\boldsymbol{\eta}} \cdot \boldsymbol{\eta} \, d\boldsymbol{r} &= \int_{-\ell/2}^{\ell/2} -\alpha_f^{\boldsymbol{n}} \boldsymbol{\nabla_n} u_f \cdot \boldsymbol{\eta} \, d\boldsymbol{r} + \int_{-\ell/2}^{\ell/2} (\boldsymbol{\beta}_f \cdot \boldsymbol{\eta}) u_f \, d\boldsymbol{r}, \\ &= -\alpha_f^{\boldsymbol{n}} (u_f|_{\gamma_2} - u_f|_{\gamma_1}) + \int_{-\ell/2}^{\ell/2} (\boldsymbol{\beta}_f \cdot \boldsymbol{\eta}) u_f \, d\boldsymbol{r}, \end{aligned}$$

où le terme $\int_{-\ell/2}^{\ell/2} (\boldsymbol{\beta}_f \cdot \boldsymbol{\eta}) u_f \, d\boldsymbol{r}$ disparaît, car la vitesse d'écoulement est négligeable dans la direction normale de la fracture (i.e., l'écoulement s'opère majoritaire suivant la composante

tangentielle). En appliquant une règle trapézoidale sur l'intégrale du membre de gauche, nous déduisons $\int_{-\ell/2}^{\ell/2} \boldsymbol{\sigma}_{f,\boldsymbol{\eta}} \cdot \boldsymbol{\eta} \; d\boldsymbol{r} \approx \frac{\ell}{2} (\boldsymbol{u}_f|_{\gamma_1} \cdot \boldsymbol{\eta} + \boldsymbol{u}_f|_{\gamma_2} \cdot \boldsymbol{\eta})$. En regroupant les différentes contributions, nous retrouvons la première condition de couplage qui lie le flux total et la variable d'état de part et d'autre de l'interface γ :

$$(\boldsymbol{\sigma}_1|_\gamma + \boldsymbol{\sigma}_2|_\gamma) \cdot \boldsymbol{\eta} = -\frac{2\alpha_f^{\boldsymbol{n}}}{\ell}(u_2|_\gamma - u_1|_\gamma) \qquad \text{sur } \gamma \times [0,T]. \tag{A.9}$$

La relation (A.9) est obtenue en fixant la continuité de u et $\boldsymbol{\sigma} \cdot \boldsymbol{\eta}$ (voir condition (A.3)) sur les portions γ_1 et γ_2. La seconde condition de couplage résulte d'une justification plus complexe. À partir des variables continues définies sur les frontières de γ_1 et γ_2, nous approchons la valeur de u dans la fracture Ω_f par le développement de Taylor suivant :

$$u_f(\boldsymbol{x}) = u_f(\boldsymbol{x}_1) + \frac{\ell}{2} \nabla u_f(\boldsymbol{\theta}_1) \cdot \boldsymbol{\eta},$$

où $\boldsymbol{x}_1 = \boldsymbol{x} - (\ell/2)\boldsymbol{\eta}$ et $\boldsymbol{\theta}_1 = \boldsymbol{x} - \xi_1(\ell/2)\boldsymbol{\eta}$, pour tout $\xi_1, \in [0,1]$. Dans la seconde section transversale, la valeur de u dans la fracture est donnée par :

$$u_f(\boldsymbol{x}) = u_f(\boldsymbol{x}_2) - \frac{\ell}{2} \nabla u_f(\boldsymbol{\theta}_2) \cdot \boldsymbol{\eta},$$

où $\boldsymbol{x}_2 = \boldsymbol{x} + (\ell/2)\boldsymbol{\eta}$ et $\boldsymbol{\theta}_2 = \boldsymbol{x} + \xi_2(\ell/2)\boldsymbol{\eta}$, pour tout $\xi_2 \in [0,1]$. En utilisant la projection normale de la vitesse totale $\boldsymbol{N}\boldsymbol{\sigma}_f$, nous retrouvons la variation linéaire du flux total sur Ω_f,

$$\begin{cases} \boldsymbol{\sigma}_{f,\boldsymbol{\eta}}(\boldsymbol{\theta}_1) \cdot \boldsymbol{\eta} &= \xi_1 \boldsymbol{\sigma}_1|_{\gamma_1} \cdot \boldsymbol{\eta} + (1-\xi_1)\boldsymbol{\sigma}_2|_{\gamma_2} \cdot \boldsymbol{\eta}, \\ \boldsymbol{\sigma}_{f,\boldsymbol{\eta}}(\boldsymbol{\theta}_2) \cdot \boldsymbol{\eta} &= \xi_2 \boldsymbol{\sigma}_2|_{\gamma_2} \cdot \boldsymbol{\eta} + (1-\xi_2)\boldsymbol{\sigma}_1|_{\gamma_1} \cdot \boldsymbol{\eta}. \end{cases} \tag{A.10}$$

À partir de la définition (A.9), l'estimation de la variable d'état u_f se transforme en :

$$\begin{cases} u_f(\boldsymbol{x}) &= \dfrac{u_1|_{\gamma_1} + u_2|_{\gamma_2}}{2} - \dfrac{\ell}{\alpha_f^{\boldsymbol{n}}} \dfrac{(2\xi_1 - 1)}{4} (\boldsymbol{\sigma}_1|_{\gamma_1} + \boldsymbol{\sigma}_2|_{\gamma_2}) \cdot \boldsymbol{\eta}, \\ u_f(\boldsymbol{x}) &= \dfrac{u_1|_{\gamma_1} + u_2|_{\gamma_2}}{2} - \dfrac{\ell}{\alpha_f^{\boldsymbol{n}}} \dfrac{(2\xi_2 - 1)}{4} (\boldsymbol{\sigma}_1|_{\gamma_1} + \boldsymbol{\sigma}_2|_{\gamma_2}) \cdot \boldsymbol{\eta}. \end{cases}$$

Puisque l'expression de $u_f(\boldsymbol{x})$ est unique dans la fracture γ, la seule possibilité revient à choisir $\xi_1 = \xi_2 = \xi$. Finalement, en intégrant sur la section transversale $[-\ell/2, \ell/2]$ de Ω_f, nous obtenons la seconde condition de couplage :

$$u_\gamma = \frac{u_1|_\gamma + u_2|_\gamma}{2} - \frac{\ell}{\alpha_f^{\boldsymbol{n}}} \frac{(2\xi - 1)}{4} (\boldsymbol{\sigma}_1|_\gamma + \boldsymbol{\sigma}_2|_\gamma) \cdot \boldsymbol{\eta} \qquad \text{sur } \gamma \times [0,T]. \tag{A.11}$$

Généralement, les conditions de couplage (A.9) et (A.11) sont formulées par le système :

$$\xi \boldsymbol{\sigma}_1|_\gamma \cdot \boldsymbol{\eta}_1 + (1-\xi)\boldsymbol{\sigma}_2|_\gamma \cdot \boldsymbol{\eta}_2 = \frac{2\alpha_f^{\boldsymbol{n}}}{\ell}(u_1|_\gamma - u_\gamma) \tag{A.12a}$$

$$(1-\xi)\boldsymbol{\sigma}_1|_\gamma \cdot \boldsymbol{\eta}_1 + \xi \boldsymbol{\sigma}_2|_\gamma \cdot \boldsymbol{\eta}_2 = \frac{2\alpha_f^{\boldsymbol{n}}}{\ell}(u_\gamma - u_2|_\gamma) \tag{A.12b}$$

pour $\xi =]0.5, 1]$.

Bibliographie

[Ackerer 2004] Philippe Ackerer, Anis Younès et Martial Mancip. *A new coupling algorithm for density-driven flow in porous media.* Geophysical research letters, vol. 31, no. 12, 2004. (Cité en page 7.)

[Adler 1999] Pierre M Adler et J-F Thovert. Fractures and fracture networks, volume 15. Springer Science & Business Media, 1999. (Cité en pages 2, 4, 13 et 122.)

[Ahmed 2015] Raheel Ahmed, Michael G Edwards, Sadok Lamine, Bastiaan AH Huisman et Mayur Pal. *Control-volume distributed multi-point flux approximation coupled with a lower-dimensional fracture model.* Journal of Computational Physics, vol. 284, pages 462–489, 2015. (Cité en page 85.)

[Alabi 2011] Olusegun Olalekan Alabi. *Validity of Darcy's law in laminar regime.* The Electronic Journal of Geotechnical Engineering, vol. 16, pages 27–40, 2011. (Cité en page 6.)

[Angot 2009] Philippe Angot, Franck Boyer et Florence Hubert. *Asymptotic and numerical modelling of flows in fractured porous media.* ESAIM : Mathematical Modelling and Numerical Analysis, vol. 43, no. 2, pages 239–275, 2009. (Cité en pages 2, 84, 85, 90, 122, 123 et 124.)

[Antonietti 2019] Paola F Antonietti, Chiara Facciola, Alessandro Russo et Marco Verani. *Discontinuous Galerkin approximation of flows in fractured porous media on polytopic grids.* SIAM Journal on Scientific Computing, vol. 41, no. 1, pages A109–A138, 2019. (Cité en pages 85 et 96.)

[Antonietti 2020] Paola F Antonietti, Chiara Facciolà et Marco Verani. *Unified analysis of discontinuous Galerkin approximations of flows in fractured porous media on polygonal and polyhedral grids.* 2020. (Cité en pages 85 et 96.)

[Arnold 1982] Douglas N Arnold. *An interior penalty finite element method with discontinuous elements.* SIAM journal on numerical analysis, vol. 19, no. 4, pages 742–760, 1982. (Cité en pages 12, 18, 38, 43, 45 et 86.)

[Arnold 2002] Douglas N Arnold, Franco Brezzi, Bernardo Cockburn et L Donatella Marini. *Unified analysis of discontinuous Galerkin methods for elliptic problems.* SIAM journal on numerical analysis, vol. 39, no. 5, pages 1749–1779, 2002. (Cité en pages 10, 12, 18, 38, 41 et 59.)

[Barth 2003] Timothy Barth et Mario Ohlberger. *Finite volume methods : foundation and analysis.* 2003. (Cité en page 11.)

[Bassi 1997] Francesco Bassi et Stefano Rebay. *A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier–Stokes equations.* Journal of computational physics, vol. 131, no. 2, pages 267–279, 1997. (Cité en page 12.)

[Baumann 1999] Carlos Erik Baumann et J Tinsley Oden. *A discontinuous hp finite element method for convection—diffusion problems.* Computer Methods in Applied Mechanics and Engineering, vol. 175, no. 3-4, pages 311–341, 1999. (Cité en page 12.)

[Bear 1988] Jacob Bear. Dynamics of fluids in porous media. Courier Corporation, 1988. (Cité en pages 4 et 7.)

[Bear 2012] Jacob Bear. Hydraulics of groundwater. Courier Corporation, 2012. (Cité en pages 3 et 4.)

[Bear 2018] Jacob Bear. Modeling phenomena of flow and transport in porous media, volume 1. Springer, 2018. (Cité en pages 4, 6, 7 et 8.)

[Berge 2019] Runar Lie Berge. *Numerical methods for coupled processes in fractured porous media*. PhD thesis, The University of Bergen, 2019. (Cité en page 125.)

[Berkowitz 2002] Brian Berkowitz. *Characterizing flow and transport in fractured geological media : A review*. Advances in water resources, vol. 25, no. 8-12, pages 861–884, 2002. (Cité en page 2.)

[Berre 2019] Inga Berre, Florian Doster et Eirik Keilegavlen. *Flow in fractured porous media : a review of conceptual models and discretization approaches*. Transport in Porous Media, vol. 130, no. 1, pages 215–236, 2019. (Cité en pages 121 et 122.)

[Berre 2021] Inga Berre, Wietse M Boon, Bernd Flemisch, Alessio Fumagalli, Dennis Gläser, Eirik Keilegavlen, Anna Scotti, Ivar Stefansson, Alexandru Tatomir, Konstantin Brenner*et al.* *Verification benchmarks for single-phase flow in three-dimensional fractured porous media*. Advances in Water Resources, vol. 147, page 103759, 2021. (Cité en pages xii, 86, 107 et 110.)

[Boffi 2018] Daniele Boffi et Daniele A Di Pietro. *Unified formulation and analysis of mixed and primal discontinuous skeletal methods on polytopal meshes*. ESAIM : Mathematical Modelling and Numerical Analysis, vol. 52, no. 1, pages 1–28, 2018. (Cité en page 118.)

[Boon 2018] Wietse M Boon, Jan M Nordbotten et Ivan Yotov. *Robust discretization of flow in fractured porous media*. SIAM Journal on Numerical Analysis, vol. 56, no. 4, pages 2203–2233, 2018. (Cité en pages 84 et 85.)

[Bourhane 2014] Anli Bourhane. *Méthodes d'investigation de l'intrusion marine dans les aquifères volcaniques (La Réunion et La Grande Comore)*. La Réunion, Université de La Réunion-Laboratoire GéoSciences Réunion. 187p, 2014. (Cité en pages 2 et 7.)

[Brezzi 2004] F. Brezzi, L. D. Marini et E. Süli. *DISCONTINUOUS GALERKIN METHODS FOR FIRST-ORDER HYPERBOLIC PROBLEMS*. Mathematical Models and Methods in Applied Sciences, vol. 14, no. 12, pages 1893–1903, 2004. (Cité en page 57.)

[Brezzi 2006] F. Brezzi, L.D. Marini, S. Micheletti, P. Pietra et R. Sacco. *Stability and error analysis of mixed finite-volume methods for advection dominated problems*. Computers & Mathematics with Applications, vol. 51, no. 5, pages 681–696, 2006. (Cité en page 57.)

[Brezzi 2012] Franco Brezzi et Michel Fortin. Mixed and hybrid finite element methods, volume 15. Springer Science & Business Media, 2012. (Cité en page 13.)

[Budiša 2021] Ana Budiša et Xiaozhe Hu. *Block preconditioners for mixed-dimensional discretization of flow in fractured porous media*. Computational Geosciences, vol. 25, no. 2, pages 671–686, 2021. (Cité en page 85.)

[Burman 2006] Erik Burman et Paolo Zunino. *A domain decomposition method based on weighted interior penalties for advection-diffusion-reaction problems.* SIAM Journal on Numerical Analysis, vol. 44, no. 4, pages 1612–1638, 2006. (Cité en pages 12, 38 et 43.)

[Camargo 2022] Murilo Camargo, Pedro R Cleto, Michael A Maedo, Eduardo A Rodrigues, Luís AG Bitencourt Jr et Osvaldo L Manzoli. *Modeling the hydrodynamic behavior of fractures and barriers in porous media using coupling finite elements.* Journal of Petroleum Science and Engineering, vol. 208, page 109700, 2022. (Cité en page 85.)

[Castillo 2002] Paul Castillo. *Performance of discontinuous Galerkin methods for elliptic PDEs.* SIAM Journal on Scientific Computing, vol. 24, no. 2, pages 524–547, 2002. (Cité en page 12.)

[Castillo 2006] Paul Castillo. *A review of the local discontinuous Galerkin (LDG) method applied to elliptic problems.* Applied numerical mathematics, vol. 56, no. 10-11, pages 1307–1313, 2006. (Cité en page 12.)

[Chauveteau 1967] G Chauveteau et CL Thirriot. *Régimes d'écoulement en milieu poreux et limite de la loi de Darcy.* La Houille Blanche, no. 2, pages 141–148, 1967. (Cité en page 6.)

[Chave 2018] Florent Chave, Daniele A Di Pietro et Luca Formaggia. *A hybrid high-order method for Darcy flows in fractured porous media.* SIAM Journal on Scientific Computing, vol. 40, no. 2, pages A1063–A1094, 2018. (Cité en pages 86 et 96.)

[Chave 2019] Florent Chave, Daniele A Di Pietro et Luca Formaggia. *A Hybrid High-Order method for passive transport in fractured porous media.* GEM-International Journal on Geomathematics, vol. 10, no. 1, pages 1–34, 2019. (Cité en page 86.)

[Chavent 1986] Guy Chavent et Jérôme Jaffré. Mathematical models and finite elements for reservoir simulation : single phase, multiphase and multicomponent flows through porous media. Elsevier, 1986. (Cité en pages 3 et 7.)

[Christie 1996] Michael A Christie. *Upscaling for reservoir simulation.* Journal of petroleum technology, vol. 48, no. 11, pages 1004–1010, 1996. (Cité en page 121.)

[Ciarlet 1991] Philippe G Ciarlet. *Basic error estimates for elliptic problems.* 1991. (Cité en pages 21, 26 et 29.)

[Cockburn 1999] Bernardo Cockburn. *Discontinuous Galerkin methods for convection-dominated problems.* In High-order methods for computational physics, pages 69–224. Springer, 1999. (Cité en page 12.)

[Cockburn 2007] Bernardo Cockburn, Jayadeep Gopalakrishnan et Haiying Wang. *Locally conservative fluxes for the continuous Galerkin method.* SIAM journal on numerical analysis, vol. 45, no. 4, pages 1742–1776, 2007. (Cité en page 11.)

[Cockburn 2008] Bernardo Cockburn, Bo Dong et Johnny Guzmán. *A superconvergent LDG-hybridizable Galerkin method for second-order elliptic problems.* Mathematics of Computation, vol. 77, no. 264, pages 1887–1916, 2008. (Cité en page 80.)

[Cockburn 2009a] Bernardo Cockburn, Bo Dong, Johnny Guzmán, Marco Restelli et Riccardo Sacco. *A hybridizable discontinuous Galerkin method for steady-state convection-*

diffusion-reaction problems. SIAM Journal on Scientific Computing, vol. 31, no. 5, pages 3827–3846, 2009. (Cité en pages 39, 42 et 43.)

[Cockburn 2009b] Bernardo Cockburn, Jayadeep Gopalakrishnan et Raytcho Lazarov. *Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems.* SIAM Journal on Numerical Analysis, vol. 47, no. 2, pages 1319–1365, 2009. (Cité en pages 13, 18, 38, 41, 42, 52, 56 et 118.)

[Cockburn 2009c] Bernardo Cockburn, Johnny Guzmán, See-Chew Soon et Henry K Stolarski. *An analysis of the embedded discontinuous Galerkin method for second-order elliptic problems.* SIAM journal on numerical analysis, vol. 47, no. 4, pages 2686–2707, 2009. (Cité en pages 39, 42, 43 et 52.)

[Cockburn 2012a] Bernardo Cockburn, George E Karniadakis et Chi-Wang Shu. Discontinuous galerkin methods : theory, computation and applications, volume 11. Springer Science & Business Media, 2012. (Cité en page 12.)

[Cockburn 2012b] Bernardo Cockburn, Weifeng Qiu et Ke Shi. *Conditions for superconvergence of HDG methods for second-order elliptic problems.* Mathematics of Computation, vol. 81, no. 279, pages 1327–1353, 2012. (Cité en page 57.)

[Cockburn 2012c] Bernardo Cockburn, Weifeng Qiu et Ke Shi. *Superconvergent HDG methods on isoparametric elements for second-order elliptic problems.* SIAM Journal on Numerical Analysis, vol. 50, no. 3, pages 1417–1432, 2012. (Cité en pages 13, 39, 43 et 86.)

[Cockburn 2016a] Bernardo Cockburn. Static condensation, hybridization, and the devising of the hdg methods, chapitre 3, pages 129–177. Springer International Publishing, Cham, 2016. (Cité en pages 57 et 65.)

[Cockburn 2016b] Bernardo Cockburn, Daniele Antonio Di Pietro et Alexandre Ern. *Bridging the hybrid high-order and hybridizable discontinuous Galerkin methods.* ESAIM : M2AN, vol. 50, no. 3, pages 635–650, 2016. (Cité en pages 57, 65 et 118.)

[Croisille 2005] J-P Croisille, Alexandre Ern, Tony Lelièvre et Jennifer Proft. *Analysis and simulation of a coupled hyperbolic/parabolic model problem.* 2005. (Cité en page 12.)

[Croucher 1995] AE Croucher et MJ O'sullivan. *The Henry problem for saltwater intrusion.* Water Resources Research, vol. 31, no. 7, pages 1809–1814, 1995. (Cité en pages 2 et 7.)

[Darcy 1856] Henry Darcy. Les fontaines publiques de la ville de dijon : Exposition et application des principes à suivre et des formules à employer dans les questions de distribution d'eau : Ouvrage terminé par un appendice relatif aux fournitures d'eau de plusieurs villes, au filtrage des eaux et à la fabrication des tuyaux de fonte, de plomb, de tôle et de bitume, volume 2. V. Dalmont, 1856. (Cité en page 6.)

[Di Pietro 2008] Daniele A Di Pietro, Alexandre Ern et Jean-Luc Guermond. *Discontinuous Galerkin methods for anisotropic semidefinite diffusion with advection.* SIAM Journal on Numerical Analysis, vol. 46, no. 2, pages 805–831, 2008. (Cité en pages 12, 38, 56 et 67.)

[Di Pietro 2011] Daniele Antonio Di Pietro et Alexandre Ern. Mathematical aspects of discontinuous galerkin methods, volume 69. Springer Science & Business Media, 2011. (Cité en pages 12, 18, 19, 21, 26, 28, 43 et 86.)

[Di Pietro 2014] Daniele A Di Pietro et Martin Vohralík. *A review of recent advances in discretization methods, a posteriori error analysis, and adaptive algorithms for numerical modeling in geosciences.* Oil & Gas Science and Technology–Revue d'IFP Energies nouvelles, vol. 69, no. 4, pages 701–729, 2014. (Cité en pages 3, 10 et 11.)

[Di Pietro 2015a] Daniele A. Di Pietro et Alexandre Ern. *A hybrid high-order locking-free method for linear elasticity on general meshes.* Computer Methods in Applied Mechanics and Engineering, vol. 283, pages 1–21, 2015. (Cité en page 57.)

[Di Pietro 2015b] Daniele Antonio Di Pietro, Jerome Droniou et Alexandre Ern. *A discontinuous-skeletal method for advection-diffusion-reaction on general meshes.* SIAM Journal on Numerical Analysis, vol. 53, no. 5, pages 2135–2157, Septembre 2015. (Cité en pages 56, 57, 58 et 80.)

[Diersch 2013] Hans-Jörg G Diersch. Feflow : finite element modeling of flow, mass and heat transport in porous and fractured media. Springer Science & Business Media, 2013. (Cité en pages 1, 6, 7, 8, 10 et 11.)

[Dietrich 2005] Peter Dietrich, Rainer Helmig, Martin Sauter, Heinz Hötzl, Jürgen Köngeter et Georg Teutsch. Flow and transport in fractured porous media. Springer, 2005. (Cité en pages 2, 3 et 121.)

[Dijoux 2019a] Loïc Dijoux. *Simulation numérique des phénomènes d'écoulement et de transport de masse en milieu poreux.* PhD thesis, Université de la Réunion, 2019. (Cité en page 2.)

[Dijoux 2019b] Loic Dijoux, Vincent Fontaine et Thierry Alex Mara. *A projective hybridizable discontinuous Galerkin mixed method for second-order diffusion problems.* Applied Mathematical Modelling, vol. 75, pages 663–677, 2019. (Cité en pages 18, 39 et 57.)

[Douglas 1976] Jim Douglas et Todd Dupont. *Interior penalty procedures for elliptic and parabolic Galerkin methods.* In Computing methods in applied sciences, pages 207–216. Springer, 1976. (Cité en page 12.)

[Dryja 2003] Maksymilian Dryja. *On discontinuous Galerkin methods for elliptic problems with discontinuous coefficients.* Computational Methods in Applied Mathematics, vol. 3, no. 1, pages 76–85, 2003. (Cité en pages 12 et 43.)

[D'Angelo 2012] Carlo D'Angelo et Anna Scotti. *A mixed finite element method for Darcy flow in fractured porous media with non-matching grids.* ESAIM : Mathematical Modelling and Numerical Analysis, vol. 46, no. 2, pages 465–489, 2012. (Cité en page 85.)

[Egger 2010] Herbert Egger et Joachim Schöberl. *A hybrid mixed discontinuous Galerkin finite-element method for convection–diffusion problems.* IMA Journal of Numerical Analysis, vol. 30, no. 4, pages 1206–1234, 2010. (Cité en pages 13, 18 et 86.)

[Ern 2004] Alexandre Ern et Jean-Luc Guermond. Theory and practice of finite elements, volume 159. Springer, 2004. (Cité en page 11.)

[Ern 2009] Alexandre Ern, Annette F Stephansen et Paolo Zunino. *A discontinuous Galerkin method with weighted averages for advection–diffusion equations with locally small and anisotropic diffusivity.* IMA Journal of Numerical Analysis, vol. 29, no. 2, pages 235–256, 2009. (Cité en pages 12, 28, 38, 56 et 70.)

[Etangsale 2021a] Gregory Etangsale, Marwan Fahs, Vincent Fontaine et AR Isa-Abadi. *Families of hybridizable interior penalty discontinuous Galerkin methods for degenerate advection-diffusion-reaction problems.* arXiv preprint arXiv :2106.00226, 2021. (Cité en page 18.)

[Etangsale 2021b] Grégory Etangsale, Vincent Fontaine et Nalitiana Rajaonison. *Performances of Hybridized-, Embedded-, and Weighted-Interior Penalty Discontinuous Galerkin Methods for Heterogeneous and Anisotropic Diffusion Problems.* Frontiers in Water, page 141, 2021. (Cité en page 95.)

[Etangsale 2022] Grégory Etangsale, Marwan Fahs, Vincent Fontaine et Nalitiana Rajaonison. *Improved error estimates of hybridizable interior penalty methods using a variable penalty for highly anisotropic diffusion problems.* Computers & Mathematics with Applications, vol. 119, pages 89–99, 2022. (Cité en pages 39, 42, 57, 76 et 86.)

[Eymard 2000] Robert Eymard, Thierry Gallouët et Raphaèle Herbin. *Finite volume methods.* Handbook of numerical analysis, vol. 7, pages 713–1018, 2000. (Cité en pages 10 et 11.)

[Fabien 2020a] Maurice S Fabien, Matthew Knepley et Beatrice Riviere. *A high order hybridizable discontinuous Galerkin method for incompressible miscible displacement in heterogeneous media.* Results in Applied Mathematics, vol. 8, page 100089, 2020. (Cité en page 18.)

[Fabien 2020b] Maurice S Fabien, Matthew G Knepley et Beatrice M Riviere. *Families of interior penalty hybridizable discontinuous galerkin methods for second order elliptic problems.* Journal of Numerical Mathematics, vol. 28, no. 3, pages 161–174, 2020. (Cité en pages 13, 18, 19, 28, 30, 31, 39, 42, 45, 52, 57, 76 et 86.)

[Farmer 2002] CL Farmer. *Upscaling : a review.* International journal for numerical methods in fluids, vol. 40, no. 1-2, pages 63–78, 2002. (Cité en page 121.)

[Fetter 2017] Charles Willard Fetter, Thomas Boving et David Kreamer. Contaminant hydrogeology. Waveland Press, 2017. (Cité en pages 1 et 7.)

[Fidkowski 2016] Krzysztof J Fidkowski. *A hybridized discontinuous Galerkin method on mapped deforming domains.* Computers & Fluids, vol. 139, pages 80–91, 2016. (Cité en page 39.)

[Fidkowski 2019] Krzysztof J Fidkowski. *Comparison of hybrid and standard discontinuous Galerkin methods in a mesh-optimisation setting.* International Journal of Computational Fluid Dynamics, vol. 33, no. 1-2, pages 34–42, 2019. (Cité en page 39.)

[Flemisch 2016] Bernd Flemisch, Alessio Fumagalli et Anna Scotti. *A review of the XFEM-based approximation of flow in fractured porous media.* Advances in discretization methods, pages 47–76, 2016. (Cité en pages 84 et 85.)

[Flemisch 2018] Bernd Flemisch, Inga Berre, Wietse Boon, Alessio Fumagalli, Nicolas Schwenck, Anna Scotti, Ivar Stefansson et Alexandru Tatomir. *Benchmarks for single-phase flow in fractured porous media.* Advances in Water Resources, vol. 111, pages 239–258, 2018. (Cité en pages xii, 86, 98, 106, 107, 110 et 113.)

[Fontaine 2008] Vincent Fontaine. *Quelques méthodes numériques robustes pour les modèles de transfert diffusif en milieu poreux.* PhD thesis, La Réunion, 2008. (Cité en pages 2 et 11.)

[Formaggia 2014] Luca Formaggia, Alessio Fumagalli, Anna Scotti et Paolo Ruffo. *A reduced model for Darcy's problem in networks of fractures.* ESAIM : Mathematical Modelling and Numerical Analysis, vol. 48, no. 4, pages 1089–1116, 2014. (Cité en page 125.)

[Formaggia 2018] Luca Formaggia, Anna Scotti et Federica Sottocasa. *Analysis of a mimetic finite difference approximation of flows in fractured porous media.* ESAIM : Mathematical Modelling and Numerical Analysis, vol. 52, no. 2, pages 595–630, 2018. (Cité en page 85.)

[Fumagalli 2012] Alessio Fumagalli. *Numerical modelling of flows in fractured porous media by the XFEM method.* PhD thesis, Politecnico di Milano, 2012. (Cité en pages 85, 89, 124 et 125.)

[Fumagalli 2013] Alessio Fumagalli et Anna Scotti. *A reduced model for flow and transport in fractured porous media with non-matching grids.* In Numerical mathematics and advanced applications 2011, pages 499–507. Springer, 2013. (Cité en pages 122, 124 et 125.)

[Fumagalli 2019] Alessio Fumagalli, Eirik Keilegavlen et Stefano Scialò. *Conforming, non-conforming and non-matching discretization couplings in discrete fracture network simulations.* Journal of Computational Physics, vol. 376, pages 694–712, 2019. (Cité en page 122.)

[Fumagalli 2021] Alessio Fumagalli et Anna Scotti. *A mathematical model for thermal single-phase flow and reactive transport in fractured porous media.* Journal of Computational Physics, vol. 434, page 110205, 2021. (Cité en page 125.)

[Gastaldi 1989] Fabio Gastaldi et Alfio Quarteroni. *On the Coupling of Hyperbolic and Parabolic Systems : Analitical and Numerical Approach.* In Proceedings of the Third German-Italian Symposium Applications of Mathematics in Industry and Technology, pages 123–165. Springer, 1989. (Cité en pages 12, 56 et 65.)

[Geuzaine 2009] Christophe Geuzaine et Jean-François Remacle. *Gmsh : A 3-D finite element mesh generator with built-in pre-and post-processing facilities.* International journal for numerical methods in engineering, vol. 79, no. 11, pages 1309–1331, 2009. (Cité en page 109.)

[Gläser 2019] Dennis Gläser, Bernd Flemisch, Rainer Helmig et Holger Class. *A hybrid-dimensional discrete fracture model for non-isothermal two-phase flow in fractured porous media.* GEM-International Journal on Geomathematics, vol. 10, no. 1, pages 1–25, 2019. (Cité en page 84.)

[Gong 2020] Linxian Gong, Lei Nie et Yan Xu. *Geometrical and topological analysis of pore space in sandstones based on x-ray computed tomography.* Energies, vol. 13, no. 15, page 3774, 2020. (Cité en pages ix et 4.)

[Guzmán 2009] Johnny Guzmán et Béatrice Rivière. *Sub-optimal convergence of non-symmetric discontinuous Galerkin methods for odd polynomial approximations.* Journal of Scientific Computing, vol. 40, no. 1, pages 273–280, 2009. (Cité en page 19.)

[Heße 2013] Falk Heße, Vladyslav Prykhodko et Sabine Attinger. *Assessing the validity of a lower-dimensional representation of fractures for numerical and analytical investigations.* Advances in water resources, vol. 56, pages 35–48, 2013. (Cité en page 2.)

[Hong 2019] Qingguo Hong, Fei Wang, Shuonan Wu et Jinchao Xu. *A unified study of continuous and discontinuous Galerkin methods.* Science China Mathematics, vol. 62, no. 1, pages 1–32, 2019. (Cité en page 12.)

[Hoteit 2005] Hussein Hoteit et Abbas Firoozabadi. *Multicomponent fluid flow by discontinuous Galerkin and mixed methods in unfractured and fractured media.* Water Resources Research, vol. 41, no. 11, 2005. (Cité en page 84.)

[Houston 2002] Paul Houston, Christoph Schwab et Endre Süli. *Discontinuous hp-finite element methods for advection-diffusion-reaction problems.* SIAM Journal on Numerical Analysis, vol. 39, no. 6, pages 2133–2163, 2002. (Cité en page 56.)

[Kadeethum 2020] T Kadeethum, HM Nick, Sanghyun Lee et F Ballarin. *Flow in porous media with low dimensional fractures by employing enriched Galerkin method.* Advances in Water Resources, vol. 142, page 103620, 2020. (Cité en page 85.)

[Khoei 2016] AR Khoei, N Hosseini et T Mohammadnejad. *Numerical modeling of two-phase fluid flow in deformable fractured porous media using the extended finite element method and an equivalent continuum model.* Advances in water resources, vol. 94, pages 510–528, 2016. (Cité en page 84.)

[Kirby 2012] Robert M Kirby, Spencer J Sherwin et Bernardo Cockburn. *To CG or to HDG : a comparative study.* Journal of Scientific Computing, vol. 51, no. 1, pages 183–212, 2012. (Cité en pages 11, 13, 18, 57 et 86.)

[Kirk 2019] Keegan LA Kirk et Sander Rhebergen. *Analysis of a pressure-robust hybridized discontinuous Galerkin method for the stationary Navier-Stokes equations.* Journal of Scientific Computing, vol. 81, no. 2, pages 881–897, 2019. (Cité en pages 18, 19 et 57.)

[Koohbor 2020a] Behshad Koohbor. *Modeling water flow and mass transport in fractured porous media : application to seawater intrusion and unsaturated zone.* PhD thesis, Strasbourg, 2020. (Cité en pages 2 et 8.)

[Koohbor 2020b] Behshad Koohbor, Marwan Fahs, Hussein Hoteit, Joanna Doummar, Anis Younes et Benjamin Belfort. *An advanced discrete fracture model for variably saturated flow in fractured porous media.* Advances in Water Resources, vol. 140, page 103602, 2020. (Cité en pages 84 et 85.)

[Köppel 2019a] Markus Köppel, Vincent Martin, Jérôme Jaffré et Jean E Roberts. *A Lagrange multiplier method for a discrete fracture model for flow in porous media.* Computational Geosciences, vol. 23, no. 2, pages 239–253, 2019. (Cité en page 86.)

[Köppel 2019b] Markus Köppel, Vincent Martin et Jean E Roberts. *A stabilized Lagrange multiplier finite-element method for flow in porous media with fractures.* GEM-International Journal on Geomathematics, vol. 10, no. 1, pages 1–29, 2019. (Cité en page 86.)

[Lampe 2013] Victor Lampe. Modelling fluid flow and heat transport in fractured porous media. Master's thesis, The University of Bergen, 2013. (Cité en pages xiii et 5.)

[Ledoux 2003] Emmanuel Ledoux. Modèles mathématiques en hydrogéologie. Centre d'Informatique Géologique, Ecole Nationale Supérieure des Mines de Paris, Janvier 2003. (Cité en pages 6 et 7.)

[Lee 2019] Jeonghun J Lee, Stephen J Shannon, Tan Bui-Thanh et John N Shadid. *Analysis of an HDG method for linearized incompressible resistive MHD equations.* SIAM Journal on Numerical Analysis, vol. 57, no. 4, pages 1697–1722, 2019. (Cité en page 18.)

[Lehrenfeld 2010] Christoph Lehrenfeld. *Hybrid discontinuous Galerkin methods for solving incompressible flow problems.* Rheinisch-Westfalischen Technischen Hochschule Aachen, vol. 111, 2010. (Cité en pages 11, 18, 22, 24, 57 et 86.)

[Lehrenfeld 2016] Christoph Lehrenfeld et Joachim Schöberl. *High order exactly divergence-free hybrid discontinuous Galerkin methods for unsteady incompressible flows.* Computer Methods in Applied Mechanics and Engineering, vol. 307, pages 339–361, 2016. (Cité en page 39.)

[Lesaint 1974] Pierre Lesaint et Pierre-Arnaud Raviart. *On a finite element method for solving the neutron transport equation.* Publications mathématiques et informatique de Rennes, no. S4, pages 1–40, 1974. (Cité en page 12.)

[LeVeque 1998] Randall J LeVeque. *Finite difference methods for differential equations.* Draft version for use in AMath, vol. 585, no. 6, page 112, 1998. (Cité en page 10.)

[Lin 2014] Guang Lin, Jiangguo Liu, Lin Mu et Xiu Ye. *Weak Galerkin finite element methods for Darcy flow : Anisotropy and heterogeneity.* Journal of computational physics, vol. 276, pages 422–437, 2014. (Cité en page 11.)

[Lin 2015] Guang Lin, Jiangguo Liu et Farrah Sadre-Marandi. *A comparative study on the weak Galerkin, discontinuous Galerkin, and mixed finite element methods.* Journal of Computational and Applied Mathematics, vol. 273, pages 346–362, 2015. (Cité en page 11.)

[Ma 2021] Tianran Ma, Keni Zhang, Weijun Shen, Chaobin Guo et Hao Xu. *Discontinuous and continuous Galerkin methods for compressible single-phase and two-phase flow in fractured porous media.* Advances in Water Resources, vol. 156, page 104039, 2021. (Cité en page 85.)

[Manzoli 2021] Osvaldo L Manzoli, Livia FA Borges, Eduardo A Rodrigues, Pedro R Cleto, Michael A Maedo et Luis AG Bitencourt Jr. *A new discrete fracture approach based on the use of coupling finite elements for modeling fluid transport in naturally fractured porous media.* Computer Methods in Applied Mechanics and Engineering, vol. 386, page 114112, 2021. (Cité en page 85.)

[Martin 2004] Vincent Martin. *Simulations multidomaines des écoulements en milieu poreux.* PhD thesis, Université Paris Dauphine-Paris IX, 2004. (Cité en pages 122 et 124.)

[Martin 2005] Vincent Martin, Jérôme Jaffré et Jean E Roberts. *Modeling fractures and barriers as interfaces for flow in porous media.* SIAM Journal on Scientific Computing, vol. 26, no. 5, pages 1667–1691, 2005. (Cité en pages 2, 84, 85, 87, 88, 90, 122, 123 et 124.)

[Meerschaert 2004] Mark M Meerschaert et Charles Tadjeran. *Finite difference approximations for fractional advection–dispersion flow equations.* Journal of computational and applied mathematics, vol. 172, no. 1, pages 65–77, 2004. (Cité en pages 10 et 11.)

[Mozolevski 2021] Igor Mozolevski, Marcio A Murad et Luciane A Schuh. *High order discontinuous Galerkin method for reduced flow models in fractured porous media.* Mathematics and Computers in Simulation, vol. 190, pages 1317–1341, 2021. (Cité en page 85.)

[Mu 2014] Lin Mu, Junping Wang, Yanqiu Wang et Xiu Ye. *Interior penalty discontinuous Galerkin method on very general polygonal and polyhedral meshes.* Journal of computational and applied mathematics, vol. 255, pages 432–440, 2014. (Cité en page 12.)

[Nguyen 2009] Ngoc Cuong Nguyen, Jaume Peraire et Bernardo Cockburn. *An implicit high-order hybridizable discontinuous Galerkin method for linear convection–diffusion equations.* Journal of Computational Physics, vol. 228, no. 9, pages 3232–3254, 2009. (Cité en pages 13, 18, 39, 57 et 86.)

[Nguyen 2011] Ngoc Cuong Nguyen, J. Peraire et Bernardo Cockburn. *An Implicit High-order Hybridizable Discontinuous Galerkin Method for the Incompressible Navier-Stokes Equations.* J. Comput. Phys., vol. 230, no. 4, pages 1147–1170, F ?rier 2011. (Cité en page 57.)

[Nikitin 2020] Kirill D Nikitin et Ruslan M Yanbarisov. *Monotone embedded discrete fractures method for flows in porous media.* Journal of Computational and Applied Mathematics, vol. 364, page 112353, 2020. (Cité en page 85.)

[Nitsche 1972] J Nitsche. *On Dirichlet problems using subspaces with nearly zero boundary conditions.* In The mathematical foundations of the finite element method with applications to partial differential equations, pages 603–627. Elsevier, 1972. (Cité en page 12.)

[Nordbotten 2019] Jan Martin Nordbotten, Wietse M Boon, Alessio Fumagalli et Eirik Keilegavlen. *Unified approach to discretization of flow in fractured porous media.* Computational Geosciences, vol. 23, no. 2, pages 225–237, 2019. (Cité en pages 85 et 122.)

[Odsæter 2018] Lars Hov Odsæter. *Reservoir Simulation in Heterogeneous and Fractured Reservoirs.* PhD thesis, NTNU, 2018. (Cité en pages 121 et 122.)

[Odsæter 2019] Lars H Odsæter, Trond Kvamsdal et Mats G Larson. *A simple embedded discrete fracture-matrix model for a coupled flow and transport problem in porous media.* Computer methods in applied mechanics and engineering, vol. 343, pages 572–601, 2019. (Cité en page 85.)

[Oikawa 2015] Issei Oikawa. *A Hybridized Discontinuous Galerkin Method with Reduced Stabilization.* Journal of Scientific Computing, vol. 65, no. 1, pages 327–340, 2015. (Cité en page 57.)

[Oikawa 2017] Issei Oikawa. *Hdg methods for second-order elliptic problems (numerical analysis : New developments for elucidating interdisciplinary problems ii)*. RIMS Kokyuroku, vol. 2037, pages 61–74, 2017. (Cité en page 19.)

[Oleinik 1973] O. Oleinik, E.V. Radkevich, J.V. Radkevic et P.C. Fife. Second-order equations with nonnegative characteristic form. Springer US, 1973. (Cité en pages 56, 57, 58 et 59.)

[Peraire 2008] Jaime Peraire et P-O Persson. *The compact discontinuous Galerkin (CDG) method for elliptic problems*. SIAM Journal on Scientific Computing, vol. 30, no. 4, pages 1806–1824, 2008. (Cité en pages 12 et 38.)

[Proft 2009] J Proft et B Riviere. *DISCONTINUOUS GALERKIN METHODS FOR CONVECTION-DIFFUSION EQUATIONS FOR VARYING AND VANISHING DIFFUSIVITY*. International Journal of Numerical Analysis & Modeling, vol. 6, no. 4, 2009. (Cité en page 12.)

[Raeisi Isa-Abadi 2020] Ali Raeisi Isa-Abadi, Vincent Fontaine, Hamid-Reza Ghafouri, Anis Younes et Marwan Fahs. *A fully interior penalty discontinuous Galerkin method for variable density groundwater flow problems*. Computers & Fluids, vol. 213, page 104744, 2020. (Cité en page 56.)

[Ramasomanana 2012] Fanilo Heninkaja Ramasomanana. *Modélisation du transport réactif dans les eaux souterraines : généralisation des méthodes ELLAM :(Eulerian-Lagrangian Localized Adjoint Method)*. PhD thesis, Strasbourg, 2012. (Cité en pages 10 et 11.)

[Reed 1973] William H Reed et Thomas R Hill. *Triangular mesh methods for the neutron transport equation*. Rapport technique, Los Alamos Scientific Lab., N. Mex.(USA), 1973. (Cité en pages 12 et 38.)

[Rivière 1999] Béatrice Rivière, Mary F Wheeler et Vivette Girault. *Improved energy estimates for interior penalty, constrained and discontinuous Galerkin methods for elliptic problems*. Computational Geosciences, vol. 3, no. 3, pages 337–360, 1999. (Cité en page 19.)

[Rivière 2008] Béatrice Rivière. Discontinuous galerkin methods for solving elliptic and parabolic equations : theory and implementation. SIAM, 2008. (Cité en pages 12, 18, 19, 21, 38, 42, 45 et 86.)

[Rubin 1983] Jacob Rubin. *Transport of reacting solutes in porous media : Relation between mathematical nature of problem formulation and chemical nature of reactions*. Water resources research, vol. 19, no. 5, pages 1231–1252, 1983. (Cité en page 7.)

[Salinas 2018] Pablo Salinas, Dimitris Pavlidis, Ziqing Xie, H Osman, Christopher C Pain et Matthew D Jackson. *A discontinuous control volume finite element method for multi-phase flow in heterogeneous porous media*. Journal of Computational Physics, vol. 352, pages 602–614, 2018. (Cité en page 85.)

[Sánchez 2017] Manuel A Sánchez, C Ciuca, Ngoc Cuong Nguyen, Jaime Peraire et Bernardo Cockburn. *Symplectic Hamiltonian HDG methods for wave propagation phenomena*. Journal of Computational Physics, vol. 350, pages 951–973, 2017. (Cité en page 18.)

[Scharfetter 1969] Donald L Scharfetter et Hermann K Gummel. *Large-signal analysis of a silicon read diode oscillator*. IEEE Transactions on electron devices, vol. 16, no. 1, pages 64–77, 1969. (Cité en page 74.)

[Schöberl 2014] Joachim Schöberl. *C++ 11 implementation of finite elements in NGSolve*. Institute for Analysis and Scientific Computing, Vienna University of Technology, 2014. (Cité en pages 30, 47 et 75.)

[Schwenck 2015] Nicolas Schwenck. An xfem-based model for fluid flow in fractured porous media. Universität Stuttgart, 2015. (Cité en pages ix et 5.)

[Sevilla 2019] Ruben Sevilla. *HDG-NEFEM for two dimensional linear elasticity*. Computers & Structures, vol. 220, pages 69–80, 2019. (Cité en page 18.)

[Shin 2015] Dongwook Shin, Youngmok Jeon et Eun-Jae Park. *A hybrid discontinuous Galerkin method for advection–diffusion–reaction problems*. Applied Numerical Mathematics, vol. 95, pages 292–303, 2015. (Cité en page 76.)

[Stephansen 2007] Annette Fagerhaug Stephansen. *Discountinuous Galerkin methods and posteriori error analysis for heterogeneous diffusion problems*. PhD thesis, Ecole des Ponts ParisTech, 2007. (Cité en page 12.)

[Sun 2009] Shuyu Sun et Jiangguo Liu. *A locally conservative finite element method based on piecewise constant enrichment of the continuous Galerkin method*. SIAM Journal on Scientific Computing, vol. 31, no. 4, pages 2528–2548, 2009. (Cité en page 11.)

[Tarhuni 2021] MN Tarhuni, WR Sulaiman, MZ Jaafar, M Milad et AM Alghol. *A Review of the Dynamic Modeling Approaches for Characterizing Fluid Flow in Naturally Fractured Reservoirs*. Energy Engineering, 2021. (Cité en page 121.)

[Tatomir 2012] Alexandru-Bogdan Tatomir. *From discrete to continuum concepts of flow in fractured porous media*. PhD thesis, Universität Stuttgart, 2012. (Cité en pages 85, 121 et 122.)

[Tunc 2012] Xavier Tunc, Isabelle Faille, Thierry Gallouët, Marie Christine Cacas et Pascal Havé. *A model for conductive faults with non-matching grids*. Computational Geosciences, vol. 16, no. 2, pages 277–296, 2012. (Cité en pages 85, 122 et 124.)

[Vasilyeva 2019] Maria Vasilyeva, Eric T Chung, Wing Tat Leung et Valentin Alekseev. *Non-local multicontinuum (NLMC) upscaling of mixed dimensional coupled flow problem for embedded and discrete fracture models*. GEM-International Journal on Geomathematics, vol. 10, no. 1, pages 1–23, 2019. (Cité en page 85.)

[Watanabe 2012] Nori Watanabe, W Wang, J Taron, UJ Görke et O Kolditz. *Lower-dimensional interface elements with local enrichment : application to coupled hydro-mechanical problems in discretely fractured porous media*. International Journal for Numerical Methods in Engineering, vol. 90, no. 8, pages 1010–1034, 2012. (Cité en page 85.)

[Wells 2011] Garth N Wells. *Analysis of an interface stabilized finite element method : the advection-diffusion-reaction equation*. SIAM journal on numerical analysis, vol. 49, no. 1, pages 87–109, 2011. (Cité en pages 13, 18, 22, 26, 27, 30, 39, 57, 67, 70 et 86.)

[Wen 1996] Xian-Huan Wen et J Jaime Gómez-Hernández. *Upscaling hydraulic conductivities in heterogeneous media : An overview.* Journal of Hydrology, vol. 183, no. 1-2, pages ix–xxxii, 1996. (Cité en page 121.)

[Wheeler 1978] Mary Fanett Wheeler. *An elliptic collocation-finite element method with interior penalties.* SIAM Journal on Numerical Analysis, vol. 15, no. 1, pages 152–161, 1978. (Cité en pages 12 et 38.)

[Woopen 2014] Michael Woopen, Aravind Balan, Georg May et Jochen Schütz. *A comparison of hybridized and standard DG methods for target-based hp-adaptive simulation of compressible flow.* Computers & Fluids, vol. 98, pages 3–16, 2014. (Cité en page 39.)

[Wu 2013] Wenjuan Wu, Xinlong Feng et Demin Liu. *The local discontinuous Galerkin finite element method for a class of convection–diffusion equations.* Nonlinear Analysis : Real World Applications, vol. 14, no. 1, pages 734–752, 2013. (Cité en page 12.)

[Younes 2003] Anis Younes. *On modelling the multidimensional coupled fluid flow and heat or mass transport in porous media.* International Journal of Heat and Mass Transfer, vol. 46, no. 2, pages 367–379, 2003. (Cité en pages 7 et 11.)

[Younes 2023] Anis Younes, Behshad Koohbor, Marwan Fahs et Hussein Hoteit. *An efficient discontinuous Galerkin - mixed finite element model for variable density flow in fractured porous media.* Journal of Computational Physics, vol. 477, page 111937, 2023. (Cité en page 85.)

[Zhang 2019] Xiao Zhang, Xiaoping Xie et Shiquan Zhang. *An optimal embedded discontinuous galerkin method for second-order elliptic problems.* Computational Methods in Applied Mathematics, vol. 19, no. 4, pages 849–861, 2019. (Cité en page 39.)

[Zhao 2022] Lina Zhao, Dohyun Kim, Eun-Jae Park et Eric Chung. *Staggered DG method with small edges for Darcy flows in fractured porous media.* Journal of Scientific Computing, vol. 90, no. 3, pages 1–24, 2022. (Cité en page 85.)

[Zulian 2022] Patrick Zulian, Philipp Schädle, Liudmila Karagyaur et Maria GC Nestola. *Comparison and application of non-conforming mesh models for flow in fractured porous media using dual Lagrange multipliers.* Journal of Computational Physics, vol. 449, page 110773, 2022. (Cité en page 86.)