HDG and HDG+ methods for harmonic wave problems with convection

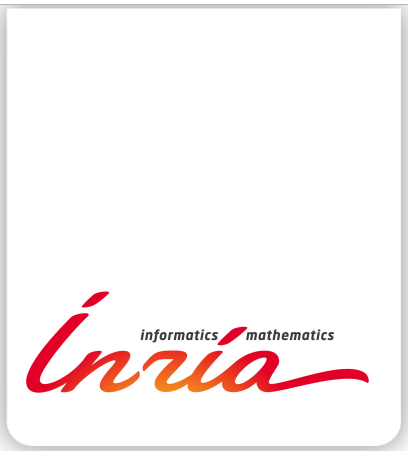
Hélène Barucq, Nathan Rouxelin, Sébastien Tordeux

HDG and HDG+ methods for harmonic wave problems with convection

Hélène Barucq, Nathan Rouxelin, Sébastien Tordeux

RESEARCH REPORT
N° 9410
June 2021
Project-Team Makutu

ISSN 0249-6399 ISRN INRIA/RR--9410--FR+ENG



HDG and HDG+ methods for harmonic wave problems with convection

Hélène Barucq*, Nathan Rouxelin*, Sébastien Tordeux*

Project-Team Makutu

Research Report n° 9410 — June 2021 — 88 pages

Abstract: In this report, we introduce three variants of the HDG method based on two weak formulations of the convected Helmholtz equation. Two of them are standard HDG methods with the same interpolation degree for all the unknowns and one of them uses a higher interpolation degree for the volumetric scalar unknown. For those three numerical methods, a detailed analysis including local and global well-posedness, as well as convergence estimates is carried out. We then provide implementation details and numerical experiments to illustrate our theoretical results.

Key-words: Hybridizable Discontinuous Galerkin Method (HDG), aeroacoustics, convected Helmholtz equation, harmonic regime

* Makutu, Inria, e2s-UPPA. Emails : {helene.barucq, nathan.rouxelin, sebastien.tordeux}@inria.fr

RESEARCH CENTRE
BORDEAUX – SUD-OUEST

200 avenue de la Vieille Tour
33405 Talence Cedex

Méthodes HDG et HDG+ pour des problèmes d'ondes convectées en régime harmonique

Résumé : Dans ce rapport, nous construisons trois variantes de la méthode HDG, basées sur deux formulations faibles de l'équation d'Helmholtz convectée. Deux de ces méthodes sont des méthodes HDG standard qui utilisent le même degré d'interpolation polynomiale pour toutes les inconnues. La troisième méthode, quant à elle, utilise un degré d'interpolation plus élevé pour l'inconnue scalaire volumique, à l'instar des méthodes HDG+. Pour toutes ces méthodes, une analyse détaillée a été effectuée, elle inclut des résultats d'existence et unicité locale et globale ainsi qu'une étude de convergence. Pour finir, nous présentons les détails de l'implémentation de ces méthodes et des expériences numériques qui illustrent nos résultats théoriques.

Mots-clés : Méthode de Galerkine discontinue hybridisable, aéroacoustique, équation d'Helmholtz convectée, domaine fréquentiel

Contents

Introduction **4**

1 Model problem **5**
1.1 First-order formulations 6

2 Notations **8**
2.1 Approximation spaces 8
2.2 Hermitian products and norms 9
2.3 Faces, jumps and averages 10

3 HDG method for the total flux formulation **11**
3.1 Constructing the formulation 11
3.2 Choice of penalization parameter 14
3.3 Local solvability 21
3.4 Error analysis 24
3.5 Global solvability 26

4 HDG(+) methods for the diffusive flux formulation **27**
4.1 Construction of the method 28
4.2 Local solvability 31
4.3 Error analysis of the HDG+ method 34
4.4 Error analysis of the HDG method with diffusive flux 47

5 Implementation **48**
5.1 Framework and notations 49
5.2 Implementation of the diffusive flux HDG method 50
5.3 Implementation of the total flux HDG method 53
5.4 Implementation of the HDG+ method 54
5.5 Comparison of the cost of the HDG and HDG+ methods 60

6 Numerical experiments **61**
6.1 Convergence rate 61
6.2 A posteriori error estimate 72
6.3 Is the upwinding mechanism necessary ? 76
6.4 Point-sources in a uniform flow 78
6.5 Gaussian jet 81

Conclusion **82**

A Intermediate results for the error analysis of the HDG method with diffusive flux **83**

References **88**

Introduction

Nowadays the solar interior is studied by considering the propagation of aeroacoustic waves in time-harmonic domain. Realistic models of solar oscillations require to approximate non-standard Hilbert settings leading to non-conforming methods, such as *Discontinuous Galerkin Methods*. As those methods have a very important numerical cost, we consider the so-called *Hybridizable Discontinuous Galerkin Methods* (HDG), which relies on a static condensation process to reduce the number of degrees of freedom.

As a first step towards the use of HDG in helioseismology, we construct and study HDG for the simplest aeroacoustic model : the *convected Helmholtz equation*.

HDG have been used and validated by numerous authors for various problems such as elliptic equations in [CGL09, CDG$^+$09, CC12, CC14], acoustic wave propagation in [GM11, GSV18, NPRC15], elastic wave propagation in [HPS17, BDMP21, CS13, FCS15, BCDL15], Maxwell equations in [CQSS17, CQS18, CLOS20]. These methods have also been used to implement the forward propagator in the context of quantitative inverse problems in [FS20] where a specific formulation of the adjoint method is developed. In this paper, we will consider the HDG+ variant of HDG, introduced in [Leh10], where different polynomial degrees are used for the different unknowns. This HDG+ has been considered for various applications in [CQSS17, Oik14, Oik16, Oik18, QSS16, QS16a, QS16b, Hun19] and to the best of our knowledge, the case of the convected Helmholtz equation has not been addressed yet.

Theory for HDGs is rather similar to the one for mixed finite elements and the actual connection was first established by Cockburn and his coworkers in [CGS10]. For a self-contained introduction to the theory of HDG, we refer to [DS19]. For a historical perspective on HDG, we refer to [Coc14].

For a comparison between HDG and Continuous Galerkin methods, we refer to [KSC12, YMKS16]. The relationship between HDG and HHO (Hybrid High-Order, another new generation of high-order face-based finite element method) has been studied in [CDPE16].

Main results: We construct three variants of the HDG method for convected acoustics in the frequency domain. Our main results include a detailed analysis of those methods where the most important properties of the method are proved including local and global solvability, convergence rate for regular solutions. The choice of the penalization parameter is also discussed. Finally, those three methods were implemented in `hawen` (see [Fau21]) and we also provide numerical experiments. Using those numerical experiments, we can conclude that the HDG+ and HDG-$\boldsymbol{\sigma}_h$ methods should be preferred to the HDG-$\boldsymbol{q}_h$ method as they seem more robust.

Organization of this paper: This work is organized as follows:

- in Section 1: we present the convected Helmholtz equation and recall some results on this equation, we also present two ways to reach a first-order in space formulation;
- in Section 2: we introduce some notations and the approximation spaces needed to construct HDGs developed in this paper;
- in Section 3: we construct the HDG-$\boldsymbol{\sigma}_h$ method based on the *total flux formulation* of the convected Helmholtz equation, we also provide theoretical results and discuss the optimal choice of penalization parameter for this method;
- in Section 4: we construct the HDG-$\boldsymbol{q}_h$ and HDG+ methods based on the *diffusive flux formulation* of the convected Helmholtz equation, we also provide detailed analysis of those methods;

- in Section 5: we give details on how those methods can be implemented in a nodal settings;
- in Section 6: we present numerical experiments to illustrate our theoretical results, as well as some illustrative examples.

1 Model problem

As a model problem we consider the so-called *convected Helmholtz equation*

$$\rho_0 \left(-\omega^2 p - 2i\omega \boldsymbol{v_0} \cdot \nabla p + \boldsymbol{v_0} \cdot \nabla(\boldsymbol{v_0} \cdot \nabla p)\right) - \operatorname{div}\left(\rho_0 c_0^2 \nabla p\right) = s \tag{1}$$

where ω is the angular frequency, ρ_0 is the density of the fluid, $\boldsymbol{v_0}$ is the velocity of the fluid, c_0 is the adiabatic sound speed, and s is the acoustic source.

Validity of this equation: Equation (1) is the simplest aeroacoustic models and therefore has a limited validity. This equation can be used for

- a *uniform background flow*, in this case the unknown p can be interpreted as a pressure perturbation,
- a *potential background flow*, in this case the unknown p should be interpreted as an *acoustic potential* and the physical quantities can be retrieved using the following identities

$$\begin{aligned} &\text{Pressure perturbation:} && p' = -\rho_0 c_0(-i\omega + \boldsymbol{v_0} \cdot \nabla)p, \\ &\text{Velocity perturbation:} && \boldsymbol{v}' = -c_0 \nabla p, \end{aligned}$$

see [Pie90, Sec. II.].

Combining the second-order differential operators: We will assume that the background flow is incompressible which leads to the following local mass conservation equation

$$\operatorname{div}(\rho_0 \boldsymbol{v_0}) = 0.$$

With this assumption, we have

$$\begin{aligned} \rho_0 \boldsymbol{v_0} \cdot \nabla(\boldsymbol{v_0} \cdot \nabla p) &= \operatorname{div}(\rho_0 (\boldsymbol{v_0} \cdot \nabla p)\boldsymbol{v_0}) - (\boldsymbol{v_0} \cdot \nabla p)\cancel{\operatorname{div}(\rho_0 \boldsymbol{v_0})} \\ &= \operatorname{div}(\rho_0 (\boldsymbol{v_0} \cdot \nabla p)\boldsymbol{v_0}) \\ &= \operatorname{div}\left(\rho_0 \boldsymbol{v_0} \boldsymbol{v_0}^T \nabla p\right) \end{aligned}$$

Leading to

$$\rho_0 \left(-\omega^2 p - 2i\omega \boldsymbol{v_0} \cdot \nabla p\right) - \operatorname{div}(\boldsymbol{K_0} \nabla p) = s \tag{2}$$

where $\boldsymbol{K_0} = \rho_0 \left(c_0^2 \mathbf{Id} - \boldsymbol{v_0} \boldsymbol{v_0}^T\right)$.
It is easy to prove that

Lemma 1.1:
$\boldsymbol{K_0}$ is symmetric positive-definite and

$$\operatorname{Sp}(\boldsymbol{K_0}) = \left\{\rho_0 c_0^2, \rho_0 (c_0^2 - |\boldsymbol{v_0}|^2)\right\}$$

Proof: $\boldsymbol{K_0} \boldsymbol{v_0} = \rho_0 \left(c_0^2 - |\boldsymbol{v_0}|^2\right) \boldsymbol{v_0}$ and $\boldsymbol{K_0} \boldsymbol{u} = \rho_0 c_0^2 \boldsymbol{u}$ for all $\boldsymbol{u} \in \boldsymbol{v_0}^\perp$.

Fredholm type : If the background flow is subsonic, ie.

$$\inf_{\mathcal{O}} \left(c_0^2 - |\boldsymbol{v_0}|^2\right) > 0, \tag{3}$$

then (2) leads to a problem of Fredholm type. Indeed, by using LEMMA 1.1 we can conclude that $-\mathrm{div}\,(\boldsymbol{K_0}\nabla p)$ is a coercive operator, and that the convected Helmholtz equation therefore has a *coercive + compact* structure.

Boundary conditions: Let Γ be the boundary of the domain $\mathcal{O}$ and let $\boldsymbol{n}$ be the outward-facing normal vector.
We will use the following boundary conditions

$$\text{Neumann:} \quad (\boldsymbol{K_0}\nabla p)\cdot\boldsymbol{n} + 2i\omega(\rho_0\boldsymbol{v_0}\cdot\boldsymbol{n})p = g_N \quad \text{on } \Gamma_N \tag{4a}$$

$$\text{Dirichlet:} \quad p = g_D \quad \text{on } \Gamma_D \tag{4b}$$

$$\text{Impedance:} \quad (\boldsymbol{K_0}\nabla p)\cdot\boldsymbol{n} + \mathcal{Z}p = g_I \quad \text{on } \Gamma_I \tag{4c}$$

and

$$\Gamma = \Gamma_N \cup \Gamma_D \quad \Gamma_D \cap \Gamma_N = \varnothing.$$

Remark 1.1: In this report, we will only consider Dirichlet (4b) and Neumann (4a) boundary conditions. Impedance boundary condition (4c) is useful to consider local *absorbing boundary conditions* which will be considered in a future work.

1.1 First-order formulations

As it is usually done in the framework of HDG methods, we will rewrite (2) as a first-order in space system. Notice that we have chosen to keep a second-order dependance in frequency. Adaptation of our method to a first-oder in frequency formulation is straightforward.
We will compare two different ways to reach a first-order in space formulation.

To lighten the notations in the remaining of this paper, we introduce the following vector field

$$\boldsymbol{b_0} := \rho_0\boldsymbol{v_0},$$

that satisfies the following mass conservation equation

$$\mathrm{div}\,(\boldsymbol{b_0}) = 0.$$

1.1.1 Diffusive flux formulation:

We begin by introducing the *diffusive flux*

$$\boldsymbol{q} := -\boldsymbol{K_0}\nabla p$$

as a new unknown, leading to the following first-order in space system

$$\boldsymbol{W_0 q} + \nabla p = 0 \tag{5a}$$

$$-\rho_0\omega^2 p - 2i\omega\boldsymbol{b_0}\cdot\nabla p + \mathrm{div}\,(\boldsymbol{q}) = s \tag{5b}$$

where

$$\boldsymbol{W_0} := \boldsymbol{K_0}^{-1} = \frac{1}{\rho_0 c_0^2}\left[\mathbf{Id} + \frac{\boldsymbol{v_0}\boldsymbol{v_0}^T}{c_0^2 - |\boldsymbol{v_0}|^2}\right]. \tag{6}$$

Note that $\boldsymbol{K_0}$ is always invertible thanks to (3), indeed we have

$$\det \boldsymbol{K_0} = \rho_0 c_0^2\left(c_0^2 - |\boldsymbol{v_0}|^2\right) \neq 0.$$

The second equality in (6) comes from the *Sherman-Morrison formula*, see [SM50] :

Lemma 1.2:
If $\boldsymbol{A} \in GL_n(\mathbb{R})$ and $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$, then $\boldsymbol{A} + \boldsymbol{u}\boldsymbol{v}^T$ is invertible if and only if $1 + \boldsymbol{v}^T\boldsymbol{A}^{-1}\boldsymbol{u} \neq 0$ and

$$\left(\boldsymbol{A} + \boldsymbol{u}\boldsymbol{v}^T\right)^{-1} = \boldsymbol{A}^{-1} - \frac{\boldsymbol{A}^{-1}\boldsymbol{u}\boldsymbol{v}^T\boldsymbol{A}^{-1}}{1 + \boldsymbol{v}^T\boldsymbol{A}^{-1}\boldsymbol{u}}.$$

With this formulation, the Neumann boundary condition (4a) becomes

$$\boldsymbol{q}\cdot\boldsymbol{n} - 2i\omega(\boldsymbol{b_0}\cdot\boldsymbol{n})p = -g_N.$$

Variational formulation: We can now write a variational formulation for (5a)–(5b) : Seek $(\boldsymbol{q}, p) \in \boldsymbol{H}_{\mathrm{div}}(\mathcal{O}) \times H^1(\mathcal{O})$ such that for all $(\boldsymbol{r}, w) \in \boldsymbol{H}_{\mathrm{div}}(\mathcal{O}) \times H^1(\mathcal{O})$

$$\int_{\mathcal{O}} \boldsymbol{W_0}\boldsymbol{q}\cdot\boldsymbol{r}^*\mathrm{d}\boldsymbol{x} - \int_{\mathcal{O}} p\,\mathrm{div}\,(\boldsymbol{r}^*)\,\mathrm{d}\boldsymbol{x} + \int_{\partial\mathcal{O}} p\boldsymbol{r}^*\cdot\boldsymbol{n}\mathrm{d}\sigma = 0 \tag{7a}$$

$$-\omega^2\int_{\mathcal{O}}\rho_0 p w^*\mathrm{d}\boldsymbol{x} + 2i\omega\int_{\mathcal{O}} p\boldsymbol{b_0}\cdot\nabla w^*\mathrm{d}\boldsymbol{x} - \int_{\mathcal{O}}\boldsymbol{q}\cdot\nabla w^*\mathrm{d}\boldsymbol{x} + \int_{\partial\mathcal{O}} w^*\boldsymbol{q}\cdot\boldsymbol{n} - 2i\omega p w^*\boldsymbol{b_0}\cdot\boldsymbol{n}\mathrm{d}\sigma = \int_{\mathcal{O}} s w^*\mathrm{d}\boldsymbol{x} \tag{7b}$$

where the boundary integrals should formally be interpreted as the duality bracket $\langle\cdot,\cdot\rangle_{H^{-\frac{1}{2}}(\partial\mathcal{O}),H^{\frac{1}{2}}(\partial\mathcal{O})}$ between $H^{-\frac{1}{2}}(\partial\mathcal{O})$ and $H^{\frac{1}{2}}(\partial\mathcal{O})$.

1.1.2 Total flux formulation:

As $\mathrm{div}\,(\boldsymbol{b_0}) = 0$, we notice that

$$2i\omega\boldsymbol{b_0}\cdot\nabla p = \mathrm{div}\,(2i\omega p\boldsymbol{b_0}),$$

and we can therefore rewrite (2) as

$$-\rho_0\omega^2 p - \mathrm{div}\,(\boldsymbol{K_0}\nabla p + 2i\omega p\boldsymbol{b_0}) = s.$$

This leads to another possible first-order in space formulation. We introduce the *total flux*

$$\boldsymbol{\sigma} := -\boldsymbol{K_0}\nabla p - 2i\omega p\boldsymbol{b_0},$$

leading to the following system

$$\boldsymbol{W_0}\boldsymbol{\sigma} + \nabla p + 2i\omega p\boldsymbol{W_0}\boldsymbol{b_0} = 0, \tag{8a}$$

$$-\rho_0\omega^2 p + \mathrm{div}\,(\boldsymbol{\sigma}) = s. \tag{8b}$$

With this formulation, the Neumann boundary condition (4a) becomes

$$\boldsymbol{\sigma}\cdot\boldsymbol{n} = -g_N.$$

Variational formulation: We can now write a variational formulation for (8a)–(8b) : Seek $(\boldsymbol{\sigma}, p) \in \boldsymbol{H}_{\mathrm{div}}(\mathcal{O}) \times H^1(\mathcal{O})$ such that for all $(\boldsymbol{r}, w) \in \boldsymbol{H}_{\mathrm{div}}(\mathcal{O}) \times H^1(\mathcal{O})$

$$\int_{\mathcal{O}} \boldsymbol{W_0}\boldsymbol{\sigma}\cdot\boldsymbol{r}^*\mathrm{d}\boldsymbol{x} - \int_{\mathcal{O}} p\,\mathrm{div}\,(\boldsymbol{r}^*)\,\mathrm{d}\boldsymbol{x} + 2i\omega\int_{\mathcal{O}} p\boldsymbol{W_0}\boldsymbol{b_0}\cdot\boldsymbol{r}^*\mathrm{d}\boldsymbol{x} + \int_{\partial\mathcal{O}} p\boldsymbol{r}^*\cdot\boldsymbol{n}\mathrm{d}\sigma = 0 \tag{9a}$$

$$-\omega^2\int_{\mathcal{O}}\rho_0 p w^*\mathrm{d}\boldsymbol{x} - \int_{\mathcal{O}}\boldsymbol{\sigma}\cdot\nabla w^*\mathrm{d}\boldsymbol{x} + \int_{\partial\mathcal{O}} w^*\boldsymbol{\sigma}\cdot\boldsymbol{n}\mathrm{d}\sigma = \int_{\mathcal{O}} s w^*\mathrm{d}\boldsymbol{x} \tag{9b}$$

where the boundary integrals should formally be interpreted as the duality bracket $\langle\cdot,\cdot\rangle_{H^{-\frac{1}{2}}(\partial\mathcal{O}),H^{\frac{1}{2}}(\partial\mathcal{O})}$ between $H^{-\frac{1}{2}}(\partial\mathcal{O})$ and $H^{\frac{1}{2}}(\partial\mathcal{O})$.

2 Notations

In this section, we introduce the notations and approximation spaces that will be used to construct the HDG methods considered in this paper.

2.1 Approximation spaces

We consider a mesh $\mathcal{T}_h$ of the domain $\mathcal{O}$ of dimension n. For an element $K \in \mathcal{T}_h$, we denote by $\mathcal{E}(K)$ the set of its edges. We also consider

$$\begin{aligned} \text{The set of boundary edges:} \quad & \mathcal{E}_h^b := \{e = \partial K \cap \Gamma \mid K \in \mathcal{T}_h\}, \\ \text{The set of interior edges:} \quad & \mathcal{E}_h^i := \{e = \partial K_+ \cap \partial K_- \mid K_+, K_- \in \mathcal{T}_h\}, \\ \text{The set of all edges:} \quad & \mathcal{E}_h := \mathcal{E}_h^b \cup \mathcal{E}_h^i. \end{aligned}$$

To study the convergence of the methods, we will assume that the mesg has the usual *shape-regularity* property, see [EG04, Def. 1.107].

For $K \in \mathcal{T}_h$, we denote by $\mathcal{P}_k(K)$ the space of polynomials of total degree at most k defined on K. We will also use the space of vectorial polynomials $\boldsymbol{\mathcal{P}}_k(K) = \mathcal{P}_k(K)^n$. Even if those spaces can be defined for $k > 0$, in this paper we will usually assume that $k \geqslant 2$ as HDG method of lower order have no interest from a computational point of view.

On each element $K \in \mathcal{T}_h$, we introduce the following approximation spaces for the pressure and the flux

$$\begin{aligned} \boldsymbol{V}_h(K) &:= \left\{ \boldsymbol{q} \in \boldsymbol{L}^2(\mathcal{O}) \;\middle|\; \boldsymbol{q}|_K \in \boldsymbol{\mathcal{P}}_k(K) \right\} && \text{for the flux } \boldsymbol{q}_h \text{ or } \boldsymbol{\sigma}_h, \\ W_h(K) &:= \left\{ p \in L^2(\mathcal{O}) \;\middle|\; p \in \mathcal{P}_\ell(K) \right\} && \text{for the pressure } p_h, \end{aligned}$$

where ℓ can be equal to k or $k+1$ depending on the formulation.

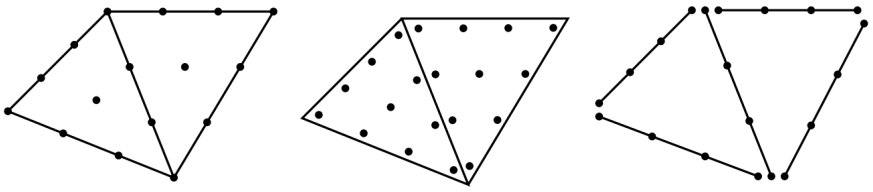
To construct HDG formulations, we will need to add a surfacic unknown, called the *numerical trace* and denoted by $\widehat{p_h}$, to the problem. This unknown will be the main unknown of the method as the *static condensation* process will allow to eliminate the volumetric unknowns to obtain a so-called *global problem*. To approximate this new unknown we introduce the following space for $e \in \mathcal{E}(K)$

$$M_h(e) := \left\{ \mu \in L^2(\mathcal{E}_h) \;\middle|\; \mu|_e \in \mathcal{P}_k(e) \right\}.$$

As those approximation spaces are discontinuous, we can construct the *global approximation spaces* as the cartesian product of the local ones

$$\begin{aligned} \boldsymbol{V}_h &:= \prod_{K \in \mathcal{T}_h} \boldsymbol{V}_h(K) && \text{for the flux } \boldsymbol{q}_h \text{ or } \boldsymbol{\sigma}_h, \\ W_h &:= \prod_{K \in \mathcal{T}_h} W_h(K) && \text{for the pressure } p_h, \\ M_h &:= \prod_{e \in \mathcal{E}_h} M_h(e) && \text{for the trace } \widehat{p_h}. \end{aligned}$$

In Figure 1, we have depicted the differences in the *degrees of freedom* for the continuous (CG), discontinuous (DG) and hybridizable discontinuous (HDG) Galerkin methods. The degrees of freedom of the HDG methods are the ones associated with the numerical trace $\widehat{p_h}$. As the numerical cost of the method is directly linked to the number of degrees of freedom, we can clearly see that the HDG method is less expensive than the DG method.



(a) CG (b) DG (c) HDG

Figure 1: Polynomial interpolation of degree 3

In TABLE 1, we give a summary of the choice of local spaces for the different variations of the HDG method considered in this report.

Variable	Space	HDG $(p_h, \boldsymbol{\sigma})$	HDG $(p_h, \boldsymbol{q}_h)$	HDG+ $(p_h, \boldsymbol{q}_h)$
Pressure p_h	$W_h(K)$	$\mathcal{P}_k(K)$		$\mathcal{P}_{k+1}(K)$
Flux $\boldsymbol{q}_h$ or $\boldsymbol{\sigma}_h$	$\boldsymbol{V}_h(K)$	$\boldsymbol{\mathcal{P}}_k(K)$		
Trace $\widehat{p_h}$	$M_h(e)$	$\mathcal{P}_k(e)$		

Table 1: Choice of local spaces and penalization parameter for the different methods

For the HDG+ method, we will also need the orthogonal projection on the space of piecewise polynomial functions on the edges

$$P_M : \prod_{K\in\mathcal{T}_h} L^2(\partial K) \longrightarrow \prod_{K\in\mathcal{T}_h} \prod_{e\in\mathcal{E}(K)} \mathcal{P}_k(e).$$

It is important to emphasize the difference between this space and M_h. Indeed as

$$M_h = \prod_{e\in\mathcal{E}_h} \mathcal{P}_k(e) \neq \prod_{K\in\mathcal{T}_h} \prod_{e\in\mathcal{E}(K)} \mathcal{P}_k(e),$$

the functions in M_h are single-valued on the skeleton of the mesh, whereas the functions in the other space are multi-valued on the interior edges. Functions in both spaces are discontinuous at the vertices.

Remark 2.1: It is also possible to choose a continuous space for $\widehat{p_h}$, this leads to the so-called *Locally Discontinuous but Globally Continuous method* (LDGC), see *eg.* [ALA13, FLd14]. However this choice does not seem to improve the convergence rate of the method.

2.2 Hermitian products and norms

For an element $K \in \mathcal{T}_h$, we denote the standard L^2-hermitian product[1] and its associated norm by

$$(u, v)_K := \int_K u \cdot v^* \mathrm{d}\boldsymbol{x} \quad \text{and} \quad \|u\|_K^2 := (u, u)_K ,$$

we then introduce the broken hermitian product and norm

$$(u, v)_{\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h} (u, v)_K \quad \text{and} \quad \|u\|_{\mathcal{T}_h}^2 := \sum_{K\in\mathcal{T}_h} \|u\|_K^2 .$$

[1]For vector fields, the $\mathbb{R}^n$ dot-product is used inside the integral as the conjugate is already applied.

On the boundary of an element K, we also denote the local hermitian product by

$$\langle u, v\rangle_{\partial K} := \sum_{e\in\mathcal{E}(K)} \int_e u \cdot v^* \mathrm{d}\sigma \quad \text{and} \quad \|u\|^2_{\partial K} := \langle u, u\rangle_{\partial K}\,,$$

and the broken hermitian product is denoted by

$$\langle u, v\rangle_{\partial \mathcal{T}_h} := \sum_{K\in\mathcal{T}_h} \langle u, v\rangle_{\partial K} \quad \text{and} \quad \|u\|^2_{\partial \mathcal{T}_h} := \sum_{K\in\mathcal{T}_h} \|u\|^2_{\partial K}\,.$$

We have chosen to use angle brackets $\langle\cdot,\cdot\rangle$ to denote the boundary integrals as they should formally be interpreted as the duality bracket between $H^{-\frac{1}{2}}(\partial K)$ and $H^{\frac{1}{2}}(\partial K)$.
We also define the following weighted norms

$$\begin{aligned} \|u\|^2_{\rho_0,K} &:= (\rho_0 u, u)_K && \text{which satisfies} && \|u\|_{\rho_0,K} \leqslant \|\rho_0\|^{\frac{1}{2}}_{L^\infty(K)} \|u\|_K \\ \|\boldsymbol{q}\|^2_{\boldsymbol{W_0},K} &:= (\boldsymbol{W_0 q}, \boldsymbol{q})_K && \text{which satisfies} && \|\boldsymbol{q}\|_{\boldsymbol{W_0},K} \leqslant C_{\boldsymbol{W_0},K} \|\boldsymbol{q}\|_K \end{aligned}$$

where

$$C_{\boldsymbol{W_0},K} = \left(\max_K \frac{1}{\rho_0\left(c_0^2 - |\boldsymbol{v_0}|^2\right)}\right)^{\frac{1}{2}}$$

is the largest eigenvalue of $\boldsymbol{W_0}$ in K, see Lemma 1.1.

2.3 Faces, jumps and averages

In this subsection, we will introduce notations for the faces quantities. As usual with methods belonging to the DG family, we will need to define jumps and averages which link the unknowns between two elements.

Faces and normals: For an interior face $\mathcal{E}_h^i \ni e = \partial K_+ \cap \partial K_-$, we denote by $\boldsymbol{n}^+$ (resp. $\boldsymbol{n}^-$) a unitary outgoing normal vector of ∂K_+ (resp. ∂K_-). We will always assume that the flow $\boldsymbol{v_0}$ goes from K_- to K_+, as depicted on Figure 2.

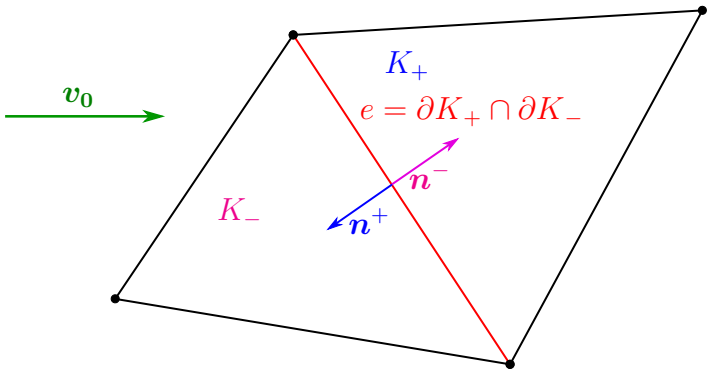


Figure 2: Normal vectors on an interior face

When the orientation of the face does not matter, we will denote by $\boldsymbol{n}$ any unitary normal vector to e.
If e is a boundary edge, then $\boldsymbol{n}$ denotes the outward-pointing unitary normal vector.

Jumps and averages: We will often use the *average operator* defined by

$$\begin{aligned} &\text{On } \mathcal{E}_h^i \ni e = \partial K_+ \cap \partial K_-, && \{\!\{\varphi\}\!\}_e := \frac{1}{2}\left(\varphi^+ + \varphi^-\right), \\ &\text{On } \mathcal{E}_h^b \ni e = \partial K \cap \Gamma, && \{\!\{\varphi\}\!\}_e := \frac{1}{2}\varphi, \end{aligned}$$

where φ can either be a scalar or vectorial quantity.
We will also make frequent use of the *jump operator* defined by

$$\begin{aligned} &\text{On } \mathcal{E}_h^i \ni e = \partial K_+ \cap \partial K_-, && [\![\boldsymbol{q}]\!]_e := \boldsymbol{q}^+ \cdot \boldsymbol{n}^+ + \boldsymbol{q}^- \cdot \boldsymbol{n}^-, \\ &\text{On } \mathcal{E}_h^b \ni e = \partial K \cap \Gamma, && [\![\boldsymbol{q}]\!]_e := \boldsymbol{q} \cdot \boldsymbol{n}, \end{aligned}$$

for a vectorial quantity. Notice that with this definition, the jump operator only controls the normal part of the vector. For a scalar quantity, the *jump operator* is defined by

$$\begin{aligned} &\text{On } \mathcal{E}_h^i \ni e = \partial K_+ \cap \partial K_-, && [\![p]\!]_e := p^+ - p^-, \\ &\text{On } \mathcal{E}_h^b \ni e = \partial K \cap \Gamma, && [\![p]\!]_e := p, \end{aligned}$$

for a scalar quantity. A sketch of those quantities is given in FIGURE 3.

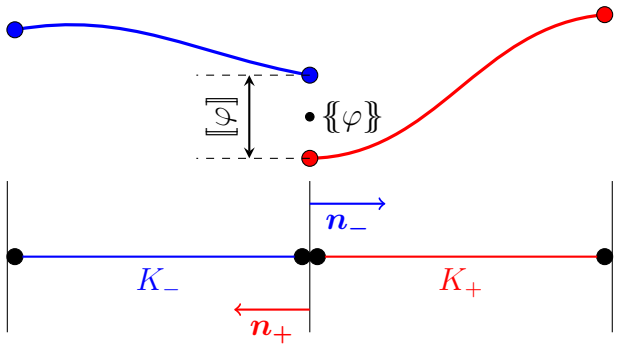


Figure 3: 1D-sketch of the jump and average on an interior node

3 HDG method for the total flux formulation

In this section, we will focus on the *total flux formulation*. We will first construct the HDG method and we will then discuss its most important properties.

3.1 Constructing the formulation

On an element $K \in \mathcal{T}_h$, we recall that the weak formulation (9a)–(9b) reads : seek $(\boldsymbol{\sigma}, p) \in \boldsymbol{H}_{\text{div}}(\mathcal{O}) \times H^1(\mathcal{O})$ such that for all $(\boldsymbol{r}, w) \in \boldsymbol{H}_{\text{div}}(\mathcal{O}) \times H^1(\mathcal{O})$

$$\int_K \boldsymbol{W_0}\boldsymbol{\sigma} \cdot \boldsymbol{r}^* \mathrm{d}\boldsymbol{x} - \int_K p \mathrm{div}\,(\boldsymbol{r}^*)\,\mathrm{d}\boldsymbol{x} + 2i\omega \int_K p \boldsymbol{W_0}\boldsymbol{b_0} \cdot \boldsymbol{r}^* \mathrm{d}\boldsymbol{x} + \int_{\partial K} p \boldsymbol{r}^* \cdot \boldsymbol{n} \mathrm{d}\sigma = 0, \tag{11a}$$

$$-\omega^2 \int_K \rho_0 p w^* \mathrm{d}\boldsymbol{x} + \int_K \mathrm{div}\,(\boldsymbol{\sigma})\, w^* \mathrm{d}\boldsymbol{x} = \int_K s w^* \mathrm{d}\boldsymbol{x} \tag{11b}$$

Choice of approximation spaces: We denote by $\boldsymbol{\sigma}_h$ and p_h the approximations of $\boldsymbol{\sigma}$ and p on K.
For this method, we choose to use the following local approximation spaces

$$\begin{aligned} \boldsymbol{V}_h(K) &= \boldsymbol{\mathcal{P}}_k(K) && \text{for the flux } \boldsymbol{\sigma}_h, \\ W_h(K) &= \mathcal{P}_k(K) && \text{for the pressurs } p_h, \end{aligned}$$

where $k \geqslant 3$ is the degree of the method. We recall that if $k \leqslant 2$, then there are no interior degrees of freedom and the HDG method has no interest over the DG methods from a computational point of view. Notice that we use the same interpolation degree for both unknowns, which may lead to unstable continuous Galerkin methods, see *eg.* [EG04, "Checkerboard-like instablitiy" p.188]. As we will discuss later, this is not a problem for HDG methods.

Introduction of the hybrid unknown: To reach a HDG formulation, we introduce a new unknown $\widehat{p_h}$ which is an approximation of p on $\mathcal{E}_h$, the skeleton of the mesh $\mathcal{T}_h$. We will usually refer to $\widehat{p_h}$ as the *numerical trace*. This unknown is the main unknown of the HDG method. Indeed, we will be able to use a *static condensation process* to eliminate the interior degrees of freedom and to obtain a so-called *global problem* for $\widehat{p_h}$ only. To introduce this unknown in the formulation, the boundary integral in (11a) is discretized as follows

$$\int_{\partial K} p\boldsymbol{r}^* \cdot \boldsymbol{n} \mathrm{d}\sigma \quad \text{becomes} \quad \int_{\partial K} \widehat{p_h}\boldsymbol{r}_h^* \cdot \boldsymbol{n} \mathrm{d}\sigma.$$

For this new unknown, we will use the following approximation space

$$M_h(e) = \mathcal{P}_k(e), \quad \forall e \in \mathcal{E}(K).$$

Penalization parameter: The unknown $\widehat{p_h}$ is often called a *Lagrange multiplier*. Indeed, when going from a continuous Galerkin method to a HDG one, the continuity of the numerical solution is not strongly enforced anymore and it is is added in the method as a constraint. The quantity $\widehat{p_h}$ is therefore the Lagrange multiplier that enforces this weak continuity requirement. To enforce this constrain, we introduce a *penalization parameter* denoted by τ. and the following boundary term

$$\langle \tau(p_h - \widehat{p_h}), w_h \rangle_{\partial K}$$

will be added to the local problem. This boundary term can be interpreted as a weak enforcement of the following Dirichlet boundary condition

$$p_h = \widehat{p_h}, \quad \text{on } \partial K.$$

Practical choice of τ will be discussed later.

Local problem: We approximate the variational formulation (11a)–(11b) on an element $K \in \mathcal{T}_h$ leading to the *local problem* : seek $(\boldsymbol{\sigma}_h, p_h) \in \boldsymbol{V}_h(K) \times W_h(K)$ such that

$$(\boldsymbol{W_0}\boldsymbol{\sigma}_h, \boldsymbol{r}_h)_K - (p_h, \operatorname{div}(\boldsymbol{r}_h))_K + 2i\omega\,(p_h \boldsymbol{W_0}\boldsymbol{b_0}, \boldsymbol{r}_h)_K + \langle \widehat{p_h}, \boldsymbol{r}_h \cdot \boldsymbol{n} \rangle_{\partial K} = 0, \tag{12a}$$

$$-\omega^2\,(\rho_0 p_h, w_h)_K + (\operatorname{div}(\boldsymbol{\sigma}_h), w_h)_K + i\omega\,\langle \tau(p_h - \widehat{p_h}), w_h \rangle_{\partial K} = (s, w_h)_K, \tag{12b}$$

for all $(\boldsymbol{r}_h, w_h) \in \boldsymbol{V}_h(K) \times W_h(K)$.
Notice that (12a)–(12b) is the variational formulation of the convected Helmholtz equation on K with weak Dirichlet boundary conditions on ∂K.

Transmission condition: Due to the discontinuous nature of the approximation spaces, we need to link all the local problems together. To this end, we introduce the *numerical flux* for $\boldsymbol{\sigma}_h$

$$\widehat{\boldsymbol{\sigma}_h} \cdot \boldsymbol{n} := \boldsymbol{\sigma}_h \cdot \boldsymbol{n} + i\omega\tau(p_h - \widehat{p_h}), \tag{13}$$

which satisfies the following *transmission condition*

$$\langle \widehat{\boldsymbol{\sigma}_h}, \mu_h \rangle_{\partial\mathcal{T}_h \setminus \Gamma_D} + \langle \widehat{p_h} - g_D, \mu_h \rangle_{\Gamma_D} = \langle g_N, \mu_h \rangle_{\Gamma_N} \tag{14}$$

for all $\mu_h \in M_h$.
Notice that (14) enforces the normal continuity of $\widehat{\boldsymbol{\sigma}_h}$ on the interior faces as well as the Neumann and Dirichlet boundary conditions on Γ_N and Γ_D.
Indeed on the interior faces we have

$$\langle \widehat{\boldsymbol{\sigma}_h}, \mu_h \rangle_{\partial\mathcal{T}_h \setminus \Gamma_D} = \sum_{e \in \mathcal{E}_h^i} \langle [\![\widehat{\boldsymbol{\sigma}_h}]\!], \mu_h \rangle_e = 0,$$

as $\boldsymbol{\sigma}_h \in \boldsymbol{V}_h$, $p_h \in W_h$ and $\widehat{p_h} \in M_h$, we have $[\![\widehat{\boldsymbol{\sigma}_h}]\!] \in M_h$ and we therefore conclude that $[\![\widehat{\boldsymbol{\sigma}_h}]\!] = 0$, as $[\![\widehat{\boldsymbol{\sigma}_h}]\!]$ is a polynomial of degree up to k orthogonal to all polynomials of degree up to k. We recall that on an interior edge $\mathcal{E}_h^i \ni e = \partial K_+ \cap \partial K_-$, the jump operator is defined as

$$[\![\boldsymbol{\sigma}_h]\!] := \boldsymbol{\sigma}_h^+ \cdot \boldsymbol{n}^+ + \boldsymbol{\sigma}_h^- \cdot \boldsymbol{n}^-.$$

Remark 3.1: The transmission condition (14) can be understood as a weak requirement of $\boldsymbol{H}_{\mathrm{div}}(\mathcal{O})$-conformity. Indeed it is shown in [PE12, Lemma 1.2.4] that $\boldsymbol{\sigma}_h \in \boldsymbol{H}_{\mathrm{div}}(\mathcal{O})$ means

$$\forall K \in \mathcal{T}_h,\ \boldsymbol{\sigma}_h^K \in \boldsymbol{H}_{\mathrm{div}}(K) \quad \text{and} \quad \forall e \in \mathcal{E}_h^i,\ [\![\boldsymbol{\sigma}_h]\!]_e \equiv 0.$$

The former is a consequence of the polynomial nature of the approximation spaces, and we will now focus and the latter. Owing to the transmission condition, we have

$$\forall e \in \mathcal{E}_h^i,\ 0 = [\![\widehat{\boldsymbol{\sigma}_h}]\!] = [\![\boldsymbol{\sigma}_h]\!] + i\omega [\![\tau(p_h - \widehat{p_h})]\!].$$

As p_h and $\widehat{p_h}$ are two approximations of the same unknown p, the quantity $p_h - \widehat{p_h}$ is expected to be small. We can therefore conclude that $[\![\boldsymbol{\sigma}_h]\!]$ is small and that

$$[\![\boldsymbol{\sigma}_h]\!] \underset{h_K \to 0}{\longrightarrow} 0.$$

For applications where a precise approximation of the flux is required, it is possible to post-process $\boldsymbol{\sigma}_h$ to obtain a new approximate $\widetilde{\boldsymbol{\sigma}_h}$ with strong $\boldsymbol{H}_{\mathrm{div}}$-conformity, see [CGS10, Sec. 5.1].

3.1.1 Compact formulation:

HDG methods are usually stated in a compact form that can be obtained by summing the local problems (12a)–(12b) over the mesh elements and by adding the transmission condition (14). This formulation reads : seek $(\boldsymbol{\sigma}_h, p_h, \widehat{p_h}) \in \boldsymbol{V}_h \times W_h \times M_h$ such that

$$(\boldsymbol{W_0}\boldsymbol{\sigma}_h, \boldsymbol{r}_h)_{\mathcal{T}_h} - (p_h, \mathrm{div}\,(\boldsymbol{r}_h))_{\mathcal{T}_h} + 2i\omega\,(p_h \boldsymbol{W_0}\boldsymbol{b_0}, \boldsymbol{r}_h)_{\mathcal{T}_h} + \langle \widehat{p_h}, \boldsymbol{r}_h \cdot \boldsymbol{n} \rangle_{\partial\mathcal{T}_h} = 0, \tag{15a}$$

$$-\omega^2\,(\rho_0 p_h, w_h)_{\mathcal{T}_h} + (\mathrm{div}\,(\boldsymbol{\sigma}_h)\,, w_h)_{\mathcal{T}_h} + i\omega\,\langle \tau(p_h - \widehat{p_h}), w_h \rangle_{\partial\mathcal{T}_h} = (s, w_h)_{\mathcal{T}_h}\,, \tag{15b}$$

$$\langle \boldsymbol{\sigma}_h \cdot \boldsymbol{n} + i\omega\tau(p_h - \widehat{p_h}), \mu_h \rangle_{\partial\mathcal{T}_h \setminus \Gamma_D} + \langle \widehat{p_h} - g_D, \mu_h \rangle_{\Gamma_D} = \langle g_N, \mu_h \rangle_{\Gamma_N}\,, \tag{15c}$$

for all $(\boldsymbol{r}_h, w_h, \mu_h) \in \boldsymbol{V}_h \times W_h \times M_h$. This formulation will be useful to perform the numerical analysis of the method.

Remark 3.2: At this point, to completely define the HDG method, it only remains to choose the penalization parameter τ, this will be done in the next section.

3.1.2 Condensed variational formulation

The compact formulation (15a)–(15b)–(15c) cannot directly be used to efficiently implement the HDG method. Indeed it is not clear how a formulation involving only $\widehat{p_h}$ can be reached. To emphasize how it can be done, we will now write a *condensed* variational formulation for $\widehat{p_h}$ only.
We introduce the so-called *local solvers*

$$\begin{aligned}
\mathsf{P}^K &: (\widehat{p_h}, s) \longmapsto p_h^K, \\
\mathsf{\Sigma}^K &: (\widehat{p_h}, s) \longmapsto \boldsymbol{\sigma}_h^K, \\
\widehat{\mathsf{\Sigma}}^K &: (\widehat{p_h}, s) \longmapsto \widehat{\boldsymbol{\sigma}_h}^K,
\end{aligned}$$

where $(\boldsymbol{\sigma}_h^K, p_h^K)$ is the solution of (12a)–(12b) and $\widehat{\boldsymbol{\sigma}_h}^K$ is defined by (13).
We can therefore rewrite the transmission condition (15c) as

$$a_h(\widehat{p_h}, \mu) = \ell_h(\mu), \tag{16}$$

where

$$\begin{aligned}
a_h(\widehat{p_h}, \mu_h) &:= \left\langle \mathsf{\Sigma}^K(\widehat{p_h}, s) \cdot \boldsymbol{n} + i\omega\tau(\mathsf{P}^K(\widehat{p_h}, s) - \widehat{p_h}), \mu_h \right\rangle_{\partial\mathcal{T}_h \setminus \Gamma_D} + \langle \widehat{p_h}, \mu_h \rangle_{\Gamma_D}, \\
\ell_h(\mu_h) &:= \langle g_N, \mu_h \rangle_{\Gamma_N} + \langle g_D, \mu_h \rangle_{\Gamma_D}.
\end{aligned}$$

Equation (16) is the so-called *global problem* and is the main equation of the HDG method. From a computational point of view, we proceed as described in Algorithm 1.

Algorithm 1: Solving HDG-$\boldsymbol{\sigma}_h$

for $K \in \mathcal{T}_h$ **do**
  Construct the local solvers P^K, $\mathsf{\Sigma}^K$, $\widehat{\mathsf{\Sigma}}^K$
  Add local contribution to the global problem (16)
Solve (16) for $\widehat{p_h}$ `// This is the main step`
for $K \in \mathcal{T}_h$ **do**
  Reconstruct the local unknowns $p_h^K = \mathsf{P}^k(\widehat{p_h}, s)$ and $\boldsymbol{\sigma}_h^K = \mathsf{\Sigma}(\widehat{p_h}, s)$

This algorithm is the blueprint of the practical implementation of the HDG method which will be discussed in Section 5.

3.2 Choice of penalization parameter

In this section, we will show how the penalization parameter can be chosen to obtain an upwinding mechanism with physical meaning. To do that, we will first need to rewrite the HDG method as a DG one, we will then solve a Riemann problem to obtain the value of τ.

3.2.1 DG formulation

In this section, we will rewrite the HDG method (15a)–(15b)–(15c) as a standard discontinuous Galerkin method.

We introduce the following bilinear form

$$\begin{aligned}\mathcal{B}_h([\boldsymbol{\sigma}_h,p_h];[\boldsymbol{r}_h,w_h]) := &(\boldsymbol{W_0}\boldsymbol{\sigma}_h,\boldsymbol{r}_h)_{\mathcal{T}_h} - (p_h,\operatorname{div}(\boldsymbol{r}_h))_{\mathcal{T}_h} + 2i\omega\,(p_h\boldsymbol{W_0}\boldsymbol{b_0},\boldsymbol{r}_h)_{\mathcal{T}_h}\\ &- \omega^2\,(\rho_0 p_h,w_h)_{\mathcal{T}_h} + (\operatorname{div}(\boldsymbol{\sigma}_h),w_h)_{\mathcal{T}_h}\\ &+ \sum_{e\in\mathcal{E}_h}\Big(\langle\widehat{p_h},[\![\boldsymbol{r}_h]\!]\rangle_e + \langle\widehat{\boldsymbol{\sigma}_h}\cdot\boldsymbol{n},[\![w_h]\!]\rangle_e\Big),\end{aligned} \tag{17}$$

from which all the mixed DG methods can be generated by choosing $\widehat{p_h}$ and $\widehat{\boldsymbol{\sigma}_h}$. Notice that $\widehat{p_h}$ was an unknown of the HDG method whereas it now should be chosen by the user of the method.
For example, the LDG method is obtained by choosing

$$\widehat{p_h} = \{\!\{p_h\}\!\} + \alpha[\![p_h]\!] \quad \text{and} \quad \widehat{\boldsymbol{\sigma}_h} = \{\!\{\boldsymbol{\sigma}_h\}\!\} + \beta[\![\boldsymbol{\sigma}_h]\!]\boldsymbol{n} + \gamma[\![p_h]\!]\boldsymbol{n},$$

and the DG method with central flux is obtained by choosing

$$\widehat{p_h} = \{\!\{p_h\}\!\} \quad \text{and} \quad \widehat{\boldsymbol{\sigma}_h} = \{\!\{\boldsymbol{\sigma}\}\!\} - \alpha[\![p_h]\!].$$

Proposition 3.1:
The HDG method (15a)–(15b)–(15c) and the DG method associated to the bilinear form (17) are equivalent if and only if

$$\widehat{p_h} = \{\!\{p_h\}\!\} + \frac{\tau^+-\tau^-}{2(\tau^++\tau^-)}[\![p_h]\!] + \frac{1}{i\omega(\tau^++\tau^-)}[\![\boldsymbol{\sigma}_h]\!], \tag{18a}$$

$$\widehat{\boldsymbol{\sigma}_h}\cdot\boldsymbol{n} = \{\!\{\boldsymbol{\sigma}_h\}\!\}\cdot\boldsymbol{n} + i\omega\frac{\tau^+\tau^-}{\tau^++\tau^-}[\![p_h]\!] - \frac{\tau^+-\tau^-}{2(\tau^++\tau^-)}[\![\boldsymbol{\sigma}_h]\!], \tag{18b}$$

for all interior edges $\mathcal{E}_h^b \ni e = \partial K_+\cap\partial K_-$ and where $\tau^\pm = \tau|_{\partial K_\pm}$.

We recall the convention for the labelling $\partial K_\pm$: $\boldsymbol{v_0}$ is directed from ∂K_- toward ∂K_+ and we denote by $\boldsymbol{n}$ any normal vector to the face when the orientation does not matter.
Proof: Writing down the transmission condition (15c) on an interior face e, we have

$$[\![\boldsymbol{\sigma}_h]\!] + 2i\omega\left(\{\!\{\tau\}\!\}\{\!\{p_h\}\!\} + \frac{1}{4}[\![\tau]\!][\![p_h]\!]\right) - 2i\omega\{\!\{\tau\}\!\}\widehat{p_h} = 0,$$

which leads to

$$\widehat{p_h} = \{\!\{p_h\}\!\} + \frac{[\![\tau]\!]}{4\{\!\{\tau\}\!\}}[\![p_h]\!] + \frac{1}{2i\omega\{\!\{\tau\}\!\}}[\![\boldsymbol{\sigma}_h]\!],$$

and we obtain (18a) by developing the jumps and average terms.
As the numerical flux $\widehat{\boldsymbol{\sigma}_h}$ is continuous across the interface, we have

$$\begin{aligned}\widehat{\boldsymbol{\sigma}_h}\cdot\boldsymbol{n} &= \{\!\{\widehat{\boldsymbol{\sigma}_h}\}\!\}\cdot\boldsymbol{n}\\ &= \{\!\{\boldsymbol{\sigma}_h\}\!\}\cdot\boldsymbol{n} + \frac{i\omega}{2}[\![\tau]\!]\left(\{\!\{p_h\}\!\} - \widehat{p_h}\right) + \frac{i\omega}{2}\{\!\{\tau\}\!\}[\![p_h]\!]\\ &= \{\!\{\boldsymbol{\sigma}_h\}\!\}\cdot\boldsymbol{n} - \frac{i\omega}{2}[\![\tau]\!]\left(\frac{[\![\tau]\!]}{4\{\!\{\tau\}\!\}}[\![p_h]\!] + \frac{1}{2i\omega\{\!\{\tau\}\!\}}[\![\boldsymbol{\sigma}_h]\!]\right) + \frac{i\omega}{2}\{\!\{\tau\}\!\}[\![p_h]\!]\end{aligned}$$

and we obtain (18b) as

$$-\frac{[\![\tau]\!]^2}{4\{\!\{\tau\}\!\}} + \{\!\{\tau\}\!\} = -\frac{(\tau^+)^2-2\tau^+\tau^-+(\tau^-)^2}{2(\tau^++\tau^-)} + \frac{(\tau^+)^2+2\tau^+\tau^-+(\tau^-)^2}{2(\tau^++\tau^-)} = 2\frac{\tau^+\tau^-}{\tau^++\tau^-}.$$

We now have to show that this choice of numerical flux $\widehat{\boldsymbol{\sigma}_h}$ is compatible with the HDG method. Starting from (18b) we have

$$\begin{aligned}\widehat{\boldsymbol{\sigma}_h}\cdot\boldsymbol{n}^+ &= \boldsymbol{\sigma}_h^+\cdot\boldsymbol{n}^+ - \frac{1}{2}\left(1+\frac{\tau^+-\tau^-}{\tau^++\tau^-}\right)[\![\boldsymbol{\sigma}_h]\!] + i\omega\frac{\tau^+\tau^-}{\tau^++\tau^-}[\![p_h]\!]\\ &= \boldsymbol{\sigma}_h^+\cdot\boldsymbol{n}^+ - \frac{\tau^+}{\tau^++\tau^-}[\![\boldsymbol{\sigma}_h]\!] + i\omega\frac{\tau^+\tau^-}{\tau^++\tau^-}[\![p_h]\!],\end{aligned}$$

on the other hand, rewriting (18a) gives

$$\begin{aligned}-\frac{\tau^+}{\tau^++\tau^-}[\![\boldsymbol{\sigma}_h]\!] &= i\omega\tau^+\left[\{\!\{p_h\}\!\} - \widehat{p_h} + \frac{\tau^+-\tau^-}{2(\tau^++\tau^-)}[\![p_h]\!]\right]\\ &= i\omega\tau^+(p_h^+-\widehat{p_h}) + i\omega\frac{\tau^+}{2}\left(\frac{\tau^+-\tau^-}{\tau^++\tau^-}-1\right)[\![p_h]\!]\\ &= i\omega\tau^+(p_h^+-\widehat{p_h}) - i\omega\frac{\tau^+\tau^-}{\tau^++\tau^-}[\![p_h]\!],\end{aligned}$$

so we finally have

$$\widehat{\boldsymbol{\sigma}_h}\cdot\boldsymbol{n}^+ = \boldsymbol{\sigma}_h^+\cdot\boldsymbol{n}^+ + i\omega\tau^+(p_h^+-\widehat{p_h}).$$

Similar computations can be carried out on ∂K_-. ∎

Particular form of the HDG fluxes: We would like to point out that $\widehat{p_h}$ depends on $[\![\boldsymbol{\sigma}_h]\!]$, this is a distinctive feature of HDG methods among the family of DG methods. To understand this, let us consider DG method with the following fluxes

$$\widehat{p_h} = \{\!\{p_h\}\!\} + \alpha[\![p_h]\!] \quad \text{and} \quad \widehat{\boldsymbol{\sigma}_h} = \{\!\{\boldsymbol{\sigma}_h\}\!\} + \beta[\![\boldsymbol{\sigma}_h]\!]\boldsymbol{n} + \gamma[\![p_h]\!]\boldsymbol{n}, \tag{19}$$

where α,β,γ are arbitrary constants. This construction is adapted from [HW08, Sec 7.2.2]. Testing (17) with $[\boldsymbol{r}_h, 0]$ leads to

$$(\boldsymbol{W_0}\boldsymbol{\sigma}_h,\boldsymbol{r}_h)_{\mathcal{T}_h} - (p_h,\operatorname{div}(\boldsymbol{r}_h))_{\mathcal{T}_h} + 2i\omega\,(p_h\boldsymbol{W_0}\boldsymbol{b_0},\boldsymbol{r}_h)_{\mathcal{T}_h} + \sum_{e\in\mathcal{E}_h}\langle\widehat{p_h},[\![\boldsymbol{r}_h]\!]\rangle_e = 0.$$

Integrating by parts leads to

$$\begin{aligned}(\boldsymbol{W_0}\boldsymbol{\sigma}_h,\boldsymbol{r}_h)_{\mathcal{T}_h} = -(\nabla p_h,\boldsymbol{r}_h)_{\mathcal{T}_h} - 2i\omega\,(p_h\boldsymbol{W_0}\boldsymbol{b_0},\boldsymbol{r}_h)_{\mathcal{T}_h} &+ \sum_{e\in\mathcal{E}_h}\left[\langle[\![p_h]\!]\boldsymbol{n},\{\!\{\boldsymbol{r}_h\}\!\}\rangle_e - \langle\widehat{p_h},[\![\boldsymbol{r}_h]\!]\rangle_e\right]\\ &+ \sum_{e\in\mathcal{E}_h^i}\langle\{\!\{p_h\}\!\},[\![\boldsymbol{r}_h]\!]\rangle_e\,,\end{aligned} \tag{20}$$

where we used the identity

$$\langle p_h,\boldsymbol{r}_h\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = \sum_{e\in\mathcal{E}_h}\langle[\![p_h]\!]\boldsymbol{n},\{\!\{\boldsymbol{r}_h\}\!\}\rangle_e + \sum_{e\in\mathcal{E}_h^i}\langle\{\!\{p_h\}\!\},[\![\boldsymbol{r}_h]\!]\rangle_e\,,$$

coming from [HW08, Lemma 7.9]. Using the definition of $\widehat{p_h}$ given in (19), the surfacic terms in (20) become

$$-\sum_{e\in\mathcal{E}_h^i}\langle[\![p_h]\!],\{\!\{\boldsymbol{r}_h\}\!\}\cdot\boldsymbol{n}-\alpha[\![\boldsymbol{r}_h]\!]\rangle_e - \sum_{e\in\mathcal{E}_h^b}\langle p_h,\boldsymbol{r}_h\cdot\boldsymbol{n}\rangle_e\,.$$

We now introduce the *lifting operator* $\mathcal{L}$ defined by

$$(\mathcal{L}(p_h),\boldsymbol{r}_h)_{\mathcal{T}_h} = \sum_{e\in\mathcal{E}_h^i}\langle[\![p_h]\!],\{\!\{\boldsymbol{r}_h\}\!\}\cdot\boldsymbol{n}-\alpha[\![\boldsymbol{r}_h]\!]\rangle_e + \sum_{e\in\mathcal{E}_h^b}\langle p_h,\boldsymbol{r}_h\cdot\boldsymbol{n}\rangle_e\,,\quad\forall\boldsymbol{r}_h\in\boldsymbol{V}_h,$$

and (20) becomes

$$(\boldsymbol{W_0}\boldsymbol{\sigma}_h, \boldsymbol{r}_h)_{\mathcal{T}_h} = (-\nabla p_h - 2i\omega p_h \boldsymbol{W_0}\boldsymbol{b_0} - \mathcal{L}(p_h), \boldsymbol{r}_h)_{\mathcal{T}_h}\,.$$

We can see that $\boldsymbol{\sigma}_h$ is completely defined in terms of p_h and it is therefore not possible to add a transmission condition, which is required to allow the static condensation process.
On the other hand, if the HDG flux

$$\widehat{p_h} = \{\!\{p_h\}\!\} + \alpha[\![p_h]\!] + \delta[\![\boldsymbol{\sigma}_h]\!]$$

is used, we will obtain an expression of $\boldsymbol{\sigma}_h$ in terms of p_h and $[\![\boldsymbol{\sigma}_h]\!]$. We will therefore need to add the transmission condition to close the discrete system and it will be possible to perform the static condensation.

3.2.2 Computing the penalization parameter

Proposition 3.2:
On an interior face $\mathcal{E}_h^i \ni e = \partial K_+ \cap \partial K_-$ the following choice of penalization parameter

$$\tau^{\pm} = \rho_0(c_0 + \boldsymbol{v_0}\cdot\boldsymbol{n}^{\pm}), \tag{21}$$

where $\tau^{\pm} = \tau|_{\partial K_{\pm}}$, leads to an upwinding mechanism.

To prove this proposition, we will need to solve a Riemann problem and compare its solution with Proposition 3.1 to obtain a value for $\tau^{\pm}$ with physical meaning. The first step to be able to solve the Riemann problem is to rewrite the original equation as a time-domain hyperbolic system.

Hyperbolic system: We start from the convected acoustic wave equation

$$\rho_0\left(\frac{\partial}{\partial t} + \boldsymbol{v_0}\cdot\nabla\right)^2 p - \operatorname{div}\left(\rho_0 c_0^2 \nabla p\right) = 0$$

and we write it as a hyperbolic system. First we have

$$\rho_0\frac{\partial^2 p}{\partial t^2} - \operatorname{div}\left(\boldsymbol{K_0}\nabla p - 2\rho_0\frac{\partial p}{\partial t}\boldsymbol{v_0}\right) = 0,$$

we therefore introduce the *total flux*

$$\frac{\partial\widetilde{\boldsymbol{\sigma}}}{\partial t} = -\boldsymbol{K_0}\nabla p + 2\rho_0\frac{\partial p}{\partial t}\boldsymbol{v_0},$$

leading to the following first-order formulation

$$\frac{\partial p}{\partial t} = -\frac{1}{\rho_0}\operatorname{div}(\widetilde{\boldsymbol{\sigma}})\,, \tag{22a}$$

$$\frac{\partial\widetilde{\boldsymbol{\sigma}}}{\partial t} = -\boldsymbol{K_0}\nabla p + 2\rho_0\frac{\partial p}{\partial t}\boldsymbol{v_0}. \tag{22b}$$

However this formulation does not have the form of a hyperbolic system.

Using (22a) in (22b), we have

$$\frac{\partial p}{\partial t} = -\frac{1}{\rho_0}\mathrm{div}\left(\widetilde{\boldsymbol{\sigma}}\right), \tag{23a}$$

$$\frac{\partial \widetilde{\boldsymbol{\sigma}}}{\partial t} = -\boldsymbol{K_0}\nabla p - 2\mathrm{div}\left(\widetilde{\boldsymbol{\sigma}}\right)\boldsymbol{v_0}. \tag{23b}$$

Notice that we need to work with a first-order in time formulation whereas our methods are written for second-order in time (or equivalently in frequency) formulations. However we have the following relationship between $\boldsymbol{\sigma}$ and $\widetilde{\boldsymbol{\sigma}}$

$$\boldsymbol{\sigma} = i\omega\widetilde{\boldsymbol{\sigma}},$$

making it possible to go back to a second-order formulation.
The system (23a)–(23b) can be written as

$$\frac{\partial \mathbf{U}}{\partial t} = \mathbb{A}_x\frac{\partial \mathbf{U}}{\partial x} + \mathbb{A}_y\frac{\partial \mathbf{U}}{\partial y}, \tag{24}$$

where

$$\mathbf{U} := \begin{bmatrix} p \\ \widetilde{\boldsymbol{\sigma}} \end{bmatrix} \quad ; \quad \mathbb{A}_x := \begin{bmatrix} 0 & -\frac{1}{\rho_0} & 0 \\ -M_{0,xx} & -2v_{0,x} & 0 \\ -M_{0,yx} & -2v_{0,y} & 0 \end{bmatrix} \quad ; \quad \mathbb{A}_y := \begin{bmatrix} 0 & 0 & -\frac{1}{\rho_0} \\ -M_{0,xy} & 0 & -2v_{0,x} \\ -M_{0,yy} & 0 & -2v_{0,y} \end{bmatrix}.$$

To check that (24) is a hyperbolic system, one needs to show that for all $\alpha, \beta \in \mathbb{R}$ the matrix

$$\mathbb{A}_{\alpha,\beta} := \alpha\mathbb{A}_x + \beta\mathbb{A}_y = -\begin{bmatrix} 0 & \frac{\alpha}{\rho_0} & \frac{\beta}{\rho_0} \\ \alpha M_{0,xx} + \beta M_{0,xy} & 2\alpha v_{0,x} & 2\beta v_{0,x} \\ \alpha M_{0,yx} + \beta M_{0,yy} & 2\alpha v_{0,y} & 2\beta v_{0,y} \end{bmatrix}$$

is diagonalizable with real eigenvalues, see Figure 4.

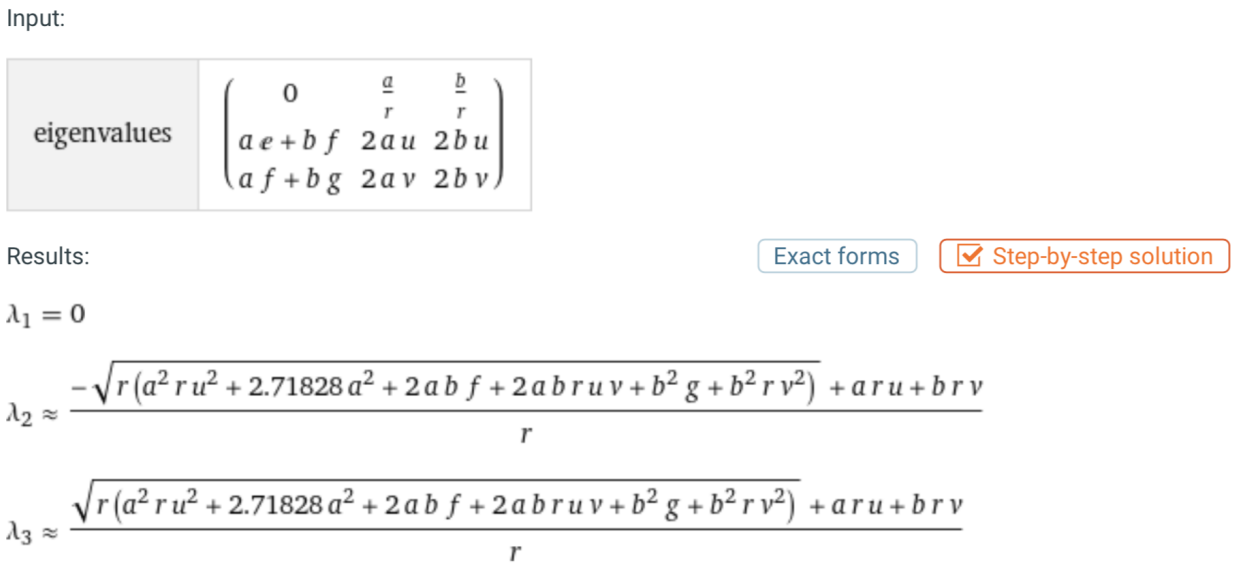


Figure 4: Computation of the eigenvalues of $\mathbb{A}_{\alpha,\beta}$ with WolframAlpha

Riemann solver: To compute the upwind penalization parameters, we consider a vertical interface located at $x = 0$ and we assume that the background flow is uniform.
We will solve the problem (24) with the following initial condition

$$\begin{aligned} \mathbf{U}(x, y, 0) &= \mathbf{U}^+, && \text{if } x > 0, \\ \mathbf{U}(x, y, 0) &= \mathbf{U}^-, && \text{if } x < 0. \end{aligned}$$

With this choice of initial condition, we obtain a well-posed problem which is invariant with respect to y.
Our goal is to compute $\mathbf{U}$ at $x = 0$.
Due to the invariance with respect to y, we can rewrite (24) as

$$\frac{\partial \mathbf{U}}{\partial t} = \mathbb{A}_x \frac{\partial \mathbf{U}}{\partial x}.$$

Furthermore, we can obtain the following system for $[p, \widetilde{\sigma}_x]^T$ only

$$\frac{\partial}{\partial t}\begin{bmatrix} p \\ \widetilde{\sigma}_x \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & -\dfrac{1}{\rho_0} \\ -M_{0,xx} & -2v_{0,x} \end{bmatrix}}_{=:\mathbb{A}} \frac{\partial}{\partial x}\begin{bmatrix} p \\ \widetilde{\sigma}_x \end{bmatrix},$$

as the DG method is only written in terms of $\widetilde{\boldsymbol{\sigma}} \cdot \boldsymbol{n} = \widetilde{\sigma}_x$.
To compute the eigenvalues of $\mathbb{A}$, we need to solve

$$\begin{vmatrix} -\lambda & -\dfrac{1}{\rho_0} \\ -M_{0,xx} & -2v_{0,x} - \lambda \end{vmatrix} = 0 \iff \lambda^2 + 2v_{0,x}\lambda - \frac{M_{0,xx}}{\rho_0} = 0.$$

Recalling that

$$M_{0,xx} = \rho_0 c_0^2 - \rho_0 v_{0,x}^2,$$

we obtain the two following eigenvalues

$$\begin{aligned} \lambda_1 &= -\left(c_0 + v_{0,x}\right), \\ \lambda_2 &= c_0 - v_{0,x}, \end{aligned}$$

and the associated eigenvectors are

$$\boldsymbol{w}_1 := \begin{bmatrix} 1 \\ \rho_0(c_0 + v_{0,x}) \end{bmatrix} \quad \text{and} \quad \boldsymbol{w}_2 := \begin{bmatrix} 1 \\ \rho_0(v_{0,x} - c_0). \end{bmatrix}$$

We can now define

$$\boldsymbol{W} := \begin{bmatrix} 1 & 1 \\ \rho_0(c_0 + v_{0,x}) & \rho_0(v_{0,x} - c_0) \end{bmatrix},$$

and therefore

$$\boldsymbol{W}^{-1} = \frac{1}{2\rho_0 c_0}\begin{bmatrix} \rho_0(c_0 - v_{0,x}) & 1 \\ \rho_0(c_0 + v_{0,x}) & -1 \end{bmatrix} = \begin{bmatrix} \ell_1 \\ \ell_2 \end{bmatrix}.$$

We have

$$\begin{aligned} \begin{bmatrix} p \\ \widetilde{\sigma}_x \end{bmatrix}(0, t) &= \ell_1 \begin{bmatrix} p^+ \\ \widetilde{\sigma}_x^+ \end{bmatrix} \boldsymbol{w}_1 + \ell_2 \begin{bmatrix} p^- \\ \widetilde{\sigma}_x^- \end{bmatrix} \boldsymbol{w}_2 \\ &= \frac{\rho_0(c_0 - v_{0,x})p^+ + \widetilde{\sigma}_x^+}{2\rho_0 c_0} \begin{bmatrix} 1 \\ \rho_0(c_0 + v_{0,x}) \end{bmatrix} + \frac{\rho_0(c_0 + v_{0,x})p^- - \widetilde{\sigma}_x^-}{2\rho_0 c_0} \begin{bmatrix} 1 \\ \rho_0(v_{0,x} - c_0) \end{bmatrix}, \end{aligned}$$

therefore

$$\widehat{p} = \frac{1}{2}\left(p^+ + p^-\right) - \frac{v_{0,x}}{2c_0}\left(p^+ - p^-\right) + \frac{1}{2\rho_0 c_0}\left(\widetilde{\sigma}_x^+ - \widetilde{\sigma}_x^-\right),$$
$$\widehat{\widetilde{\sigma}_x} = \frac{1}{2}\left(\widetilde{\sigma}_x^+ + \widetilde{\sigma}_x^-\right) + \frac{v_{0,x}}{2c_0}\left(\widetilde{\sigma}_x^+ - \widetilde{\sigma}_x^-\right) + \rho_0\frac{c_0^2 - v_{0,x}^2}{2c_0}\left(p^+ - p^-\right).$$

Finally, we can infer the form of the DG flux for a generic interface

$$\widehat{p} = \{\!\{p\}\!\} - \frac{\boldsymbol{v_0}\cdot\boldsymbol{n}^-}{2c_0}[\![p]\!] + \frac{1}{2\rho_0 c_0}[\![\widetilde{\boldsymbol{\sigma}}]\!], \tag{26a}$$
$$\widehat{\widetilde{\boldsymbol{\sigma}}}\cdot\boldsymbol{n}^- = \{\!\{\widetilde{\boldsymbol{\sigma}}\}\!\}\cdot\boldsymbol{n}^- + \frac{\boldsymbol{v_0}\cdot\boldsymbol{n}^-}{2c_0}[\![\widetilde{\boldsymbol{\sigma}}]\!] + \rho_0\frac{c_0^2 - (\boldsymbol{v_0}\cdot\boldsymbol{n})^2}{2c_0}[\![p]\!]. \tag{26b}$$

Notice that we had to chose an orientation of the normal vector. Following our convention, we have chosen to use $\boldsymbol{n}^-$ as it has the same orientation as $\boldsymbol{v_0}$.
Rewriting (26a)–(26b) in terms of $\boldsymbol{\sigma}$ instead of $\widetilde{\boldsymbol{\sigma}}$ leads to

$$\widehat{p_h} = \{\!\{p\}\!\} - \frac{\boldsymbol{v_0}\cdot\boldsymbol{n}^-}{2c_0}[\![p]\!] + \frac{1}{2i\omega\rho_0 c_0}[\![\boldsymbol{\sigma}]\!] \tag{27a}$$
$$\widehat{\boldsymbol{\sigma}}\cdot\boldsymbol{n}^- = \{\!\{\widetilde{\boldsymbol{\sigma}}\}\!\}\cdot\boldsymbol{n}^- + \frac{\boldsymbol{v_0}\cdot\boldsymbol{n}^-}{2c_0}[\![\widetilde{\boldsymbol{\sigma}}]\!] + i\omega\rho_0\frac{c_0^2 - (\boldsymbol{v_0}\cdot\boldsymbol{n}^-)^2}{2c_0}[\![p]\!]. \tag{27b}$$

Comparing (27a)–(27b) with (18a)–(18b) , we see that

$$\tau^+ + \tau^- = 2\rho_0 c_0, \tag{28a}$$
$$\frac{\tau^+\tau^-}{\tau^+ + \tau^-} = \rho_0\frac{c_0^2 - (\boldsymbol{v_0}\cdot\boldsymbol{n}^-)^2}{2c_0}. \tag{28b}$$

The system (28a)–(28b) leads to the following second-order equation

$$(\tau^+)^2 - 2\rho_0 c_0\tau^+ + \rho_0^2\left(c_0^2 - (\boldsymbol{v_0}\cdot\boldsymbol{n}^-)^2\right) = 0,$$

and to the two following families for $\tau^\pm$

$$\tau_1^+ = \rho_0(c_0 + \boldsymbol{v_0}\cdot\boldsymbol{n}^-), \qquad \tau_1^- = \rho_0(c_0 - \boldsymbol{v_0}\cdot\boldsymbol{n}^-),$$
$$\tau_2^+ = \rho_0(c_0 - \boldsymbol{v_0}\cdot\boldsymbol{n}^-), \qquad \tau_2^- = \rho_0(c_0 + \boldsymbol{v_0}\cdot\boldsymbol{n}^-).$$

To discriminate between $\tau_1^\pm$ and $\tau_2^\pm$ we once again go back to (18a)–(18b) and we see that the solution must satisfy

$$\frac{\tau^+ - \tau^-}{2(\tau^+ + \tau^-)} = -\frac{\boldsymbol{v_0}\cdot\boldsymbol{n}^-}{2c_0}.$$

We can therefore conclude that the upwind fluxes are obtained by using the $\tau_2^\pm$ solution. We can make this choice independent of the orientation convention by noticing that $\boldsymbol{n}^+ = -\boldsymbol{n}^-$, leading to

$$\tau_2^\pm = \rho_0(c_0 + \boldsymbol{v_0}\cdot\boldsymbol{n}^\pm).$$

Remark 3.3: To keep polynomial fluxes on the interfaces, the background quantities will be approximated by their value at the center of the interface.

Remark 3.4: In the context of DG and HDG methods, τ is usually chosen to be of the «order of unity» to ensure optimal convergence rate. In the error analysis of the method, we allow the dependency to the background coefficient to be hidden in the constants, so the choice (21) is actually possible.

3.3 Local solvability

We will now show the local solvability for the *total flux* formulation. Proving the well-posedness of the local problems is always very important when working with HDG methods. For the strongly coercive problems, for which HDG methods were initially designed, this property usually comes directly from the continuous problem. However for harmonic wave equations, which are only weakly coercive, things are more complicated : indeed solving the local problem amounts to solving a wave problem with Dirichlet boundary conditions. We therefore need to ensure that the local problem does not introduce resonance into the method, which is the case when the elements are small enough. In this section, we will prove that the static condensation process is well-defined when the mesh is fine enough.

Notice that in this case the proof relies on an *absorption technique* and is therefore very technical. Readers who are not familiar with HDG theory should probably begin with the proof for the *diffusive flux* formulation which is easier. It will be detailed in Subsection 4.2.

First, we need to show the

Lemma 3.1:
For $p_h \in \mathcal{P}_k(K)$ with $k > 0$, the following inverse inequality holds

$$\|\nabla p_h \cdot \boldsymbol{n}\|_{\partial K} \lesssim \|\nabla p_h\|_{\partial K} \lesssim h_K^{-\frac{1}{2}} \|\nabla p_h\|_K .$$

Proof:
First, we notice that if p_h is constant the desired inequality reduces to $0 \lesssim 0$. We therefore only consider non-constant p_h.
Let $\widetilde{K}$ be the reference unit element. We consider the map $F : \widetilde{K} \longrightarrow K$. We use $\widetilde{\cdot}$ to denote quantities on the reference element instead of the more standard notation $\hat{\cdot}$ to avoid confusion, as we already used $\hat{\cdot}$ to denote the numerical fluxes.
Let $\tilde{\gamma}^1 : H^2(\widetilde{K}) \longrightarrow L^2(\partial\widetilde{K})$ be the normal derivative operator in the reference element. As $\tilde{\gamma}^1$ is continuous, we have

$$\left\|\tilde{\gamma}^1(\widetilde{p_h})\right\|_{\partial\widetilde{K}} \lesssim \|\widetilde{p_h}\|_{2,\widetilde{K}} \lesssim |\widetilde{p_h}|_{1,\widetilde{K}}$$

The second inequality holds as $\widetilde{p_h} \in \mathcal{P}_k(\widetilde{K})$ which is a finite-dimensional vector-space on which all the norms are equivalent and p_h is not constant.
We now recall the following scaling inequalities, see [DS19, Eq (1.6), (1.7) & (1.8)]

$$|\widetilde{p_h}|_{1,\widetilde{K}} \lesssim |\det \mathrm{Jac}(F)|^{-\frac{1}{2}} \|\mathrm{Jac}(F)\| \, |p_h|_{1,K} \lesssim |p_h|_{1,K};$$
$$h_K^{\frac{1}{2}} \|\mu_h\|_{\partial K} \lesssim \|\widetilde{\mu_h}\|_{\partial\widetilde{K}}$$

Due to the regularity of the mesh, we have

$$h_K^{\frac{1}{2}} \|\nabla p_h \cdot \boldsymbol{n}\|_{\partial K} \lesssim \|\nabla p_h\|_K .$$

■

Theorem 1 : *Local solvability for the total flux HDG method*

If τ is chosen such that

$$\exists \tau_0 > 0, \quad \forall e \in \mathcal{E}(K), \quad 0 < \tau_0 \leqslant \tau + \boldsymbol{b_0} \cdot \boldsymbol{n},$$

then there exists a constant $\alpha_+ > 0$ such that the local problem is well-posed if $\omega h_K < \alpha_+$.

Proof: As (12a)–(12b) is a finite-dimensional problem, we only need to prove uniqueness of the solution. We therefore assume that $\widehat{p_h} = 0$ and $s = 0$, and we need to show that the system

$$\left(\boldsymbol{W_0}\boldsymbol{\sigma}_h, \boldsymbol{r}_h\right)_K - \left(p_h, \operatorname{div}\left(\boldsymbol{r}_h\right)\right)_K + 2i\omega\left(p_h\boldsymbol{W_0}\boldsymbol{b_0}, \boldsymbol{r}_h\right)_K = 0, \qquad \forall \boldsymbol{r}_h \in \boldsymbol{V}_h(K) \tag{30a}$$

$$-\omega^2\left(\rho_0 p_h, w_h\right)_K + \left(\operatorname{div}\left(\boldsymbol{\sigma}_h\right), w_h\right)_K + i\omega\left\langle \tau p_h, w_h\right\rangle_{\partial K} = 0, \qquad \forall w_h \in W_h(K). \tag{30b}$$

has only one solution $(\boldsymbol{\sigma}_h, p_h) = (\boldsymbol{0}, 0)$.
We will prove the theorem by contradiction. We therefore assume that there is a non-zero solution $(\boldsymbol{\sigma}_h, p_h)$ to (30a)–(30b).
Step 1: Energy-like identity.
We test (30a) with $\boldsymbol{r}_h = \boldsymbol{\sigma}_h$ and conjugate the resulting equation, we then test (30b) with $w_h = p_h$ and add the two resulting equations leading to

$$\left\|\boldsymbol{\sigma}_h\right\|^2_{\boldsymbol{W_0},K} - \omega^2\left\|p_h\right\|^2_{\rho_0,K} - 2i\omega\left(\boldsymbol{\sigma}_h, p_h\boldsymbol{W_0}\boldsymbol{b_0}\right)_K + i\omega\left\langle \tau p_h, p_h\right\rangle_{\partial K} = 0.$$

We then focus on the third term, as $\boldsymbol{W_0}$ is real and symmetric we have

$$\left(\boldsymbol{\sigma}_h, p_h\boldsymbol{W_0}\boldsymbol{b_0}\right)_K = \left(\boldsymbol{W_0}\boldsymbol{\sigma}_h, p_h\boldsymbol{b_0}\right)_K.$$

Taking $\boldsymbol{r}_h = p_h\left\{\boldsymbol{b_0}\right\}$, where $\left\{\boldsymbol{b_0}\right\}$ is the average of $\boldsymbol{b_0}$ on K, in (30a), we have

$$\left(\boldsymbol{W_0}\boldsymbol{\sigma}_h, p_h\left\{\boldsymbol{b_0}\right\}\right)_K - \left(p_h, \left\{\boldsymbol{b_0}\right\}\cdot\nabla p_h\right)_K + 2i\omega\left(p_h\boldsymbol{W_0}\boldsymbol{b_0}, p_h\left\{\boldsymbol{b_0}\right\}\right)_K = 0,$$

leading to

$$-\left(\boldsymbol{W_0}\boldsymbol{\sigma}_h, p_h\boldsymbol{b_0}\right)_K + \left(p_h, \boldsymbol{b_0}\cdot\nabla p_h\right)_K - 2i\omega\left(p_h\boldsymbol{W_0}\boldsymbol{b_0}, p_h\boldsymbol{b_0}\right)_K = -\varepsilon,$$

where

$$\varepsilon := \left(\boldsymbol{W_0}\boldsymbol{\sigma}_h, p_h(\boldsymbol{b_0} - \left\{\boldsymbol{b_0}\right\})\right)_K - \left(p_h, (\boldsymbol{b_0} - \left\{\boldsymbol{b_0}\right\})\cdot\nabla p_h\right)_K + 2i\omega\left(p_h\boldsymbol{W_0}\boldsymbol{b_0}, p_h(\boldsymbol{b_0} - \left\{\boldsymbol{b_0}\right\})\right)_K.$$

We chose the notation ε to emphasize that this quantity is small, as it will be discussed in *Step 3*.
So we have

$$\begin{aligned}\left\|\boldsymbol{\sigma}_h\right\|^2_{\boldsymbol{W_0},K} - \omega^2(\left\|p_h\right\|^2_{\rho_0,K} + 4\left\|p_h\boldsymbol{b_0}\right\|^2_{\boldsymbol{W_0},K}) \\ - 2i\omega\left(p_h, \boldsymbol{b_0}\cdot\nabla p_h\right)_K - 2i\omega\varepsilon + i\omega\left\langle \tau p_h, p_h\right\rangle_{\partial K} = 0\end{aligned} \tag{31}$$

Step 2: Boundary condition
Taking the imaginary part (31) leads to

$$\begin{aligned}2\omega\mathfrak{Re}\left(p_h\boldsymbol{b_0}, \nabla p_h\right)_K + \omega\left\langle \tau p_h, p_h\right\rangle_{\partial K} + 2\omega\mathfrak{Re}\varepsilon = 0, \\ \text{(by LEMMA 4.1)} \quad \left\langle(\tau + \boldsymbol{b_0}\cdot\boldsymbol{n})p_h, p_h\right\rangle_{\partial K} = -2\mathfrak{Re}\varepsilon.\end{aligned}$$

As we have chosen τ such that there is $0 < \tau_0 \leqslant \tau + \boldsymbol{b_0} \cdot \boldsymbol{n}$ where τ_0 does not depend on h_K, we have

$$\|p_h\|_{\partial K}^2 \lesssim |\varepsilon|. \tag{32}$$

As we do not have shown that $p|_{\partial K} = 0$, we cannot use Poincaré's inequality, however according to [EG04, Lemma B.63 & Example B.64] we have

$$\|p_h\|_K \leqslant \|p_h\|_{1,K} \leqslant C_K \|\nabla p_h\|_K + \frac{C_K}{\text{meas}(\partial K)} \|p_h\|_{\partial K},$$

where C_K is the Poincaré constant of K[2]. Using standard scaling inequalities, we have

$$C_K \lesssim h_K \quad \text{and} \quad \frac{C_K}{\text{meas}(\partial K)} \lesssim h_K^{\frac{1}{2}}.$$

Using (32) to estimate the boundary term, this leads to

$$\|p_h\|_K \lesssim h_K \|\nabla p_h\|_K + h_K^{1/2} |\varepsilon|^{1/2}. \tag{33}$$

Step 3: Estimating $|\varepsilon|$ and $\|p_h\|_K$. We have

$$\begin{aligned}
|\varepsilon| \lesssim & h_K \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K} \|p_h\|_K + h_K \|p_h\|_K \|\nabla p_h\|_K + \omega h_K \|p_h\|_K^2 \\
\text{(by Young)} \quad \lesssim & h_K^2 \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K}^2 + (1 + \omega h_K) \|p_h\|_K^2 + h_K^2 \|\nabla p_h\|_K^2 \\
\lesssim & h_K^2 \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K}^2 + (1 + \omega h_K) \left(h_K^2 \|\nabla p_h\|_K^2 + h_K |\varepsilon| \right) + h_K^2 \|\nabla p_h\|_K^2
\end{aligned}$$

If h_K is small enough, the term $h_K|\varepsilon|$ in the right-hand side can be absorbed by the left-hand side leading to

$$|\varepsilon| \lesssim h_K^2 \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K}^2 + (1 + \omega h_K) h_K^2 \|\nabla p_h\|_K^2 .$$

Assuming that h_K is small enough, we can overestimate $\omega h_K \lesssim 1$, leading to

$$|\varepsilon| \lesssim h_K^2 \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K}^2 + h_K^2 \|\nabla p_h\|_K^2 . \tag{34}$$

Together with the generalized Poincaré's inequality (33) we have

$$\|p_h\|_K^2 \lesssim h_K^2 \|\nabla p_h\|_K^2 + h_K^2 \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K}^2 , \tag{35}$$

and using that $\sqrt{a^2 + b^2} \leqslant |a| + |b|$ we also have

$$\|p_h\|_K \lesssim h_K \|\nabla p_h\|_K + h_K \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K} .$$

Step 4: Estimating $\|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K}$
Taking the real part of the Garding's identity (31), we have

$$\begin{aligned}
\|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K}^2 \lesssim & \omega^2 \|p_h\|_K^2 + \omega \|p_h\|_K \|\nabla p_h\|_K + \omega |\varepsilon| \\
\text{(by (35))} \quad \lesssim & \omega^2 \left(h_K^2 \|\nabla p_h\|_K^2 + h_K^2 \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K}^2 \right) + \omega \left(h_K \|\nabla p_h\|_K + h_K \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K} \right) \|\nabla p_h\|_K + \omega |\varepsilon| \\
\text{(by Young)} \quad \lesssim & \left(\omega^2 h_K^2 + \omega h_K \right) \|\nabla p_h\|_K^2 + \omega^2 h_K^2 \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K}^2 + \omega h_K \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K}^2 + \omega |\varepsilon| \\
\text{(by (34))} \quad \lesssim & \left(\omega^2 h_K^2 + \omega h_K \right) \|\nabla p_h\|_K^2 + \left(\omega^2 h_K^2 + \omega h_K \right) \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K} \\
\lesssim & \omega h_K \|\nabla p_h\|_K^2 + \omega h_K \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K}
\end{aligned}$$

[2]The constant used by the authors of [EG04] is the inverse of the usual Poincaré constant.

We obtained the last line by assuming that $\omega^2 h_K^2 \lesssim \omega h_K$ which is true if h_K is small enough. If h_K is small enough the term $\omega h_K \|\boldsymbol{\sigma}_h\|^2_{\boldsymbol{W_0},K}$ in the right-hand side can be absorbed by the left-hand side, leading to

$$\|\boldsymbol{\sigma}_h\|^2_{\boldsymbol{W_0},K} \lesssim \omega h_K \|\nabla p_h\|^2_K . \tag{36}$$

Step 5: Estimating $\|\nabla p_h\|_K$
Taking $\boldsymbol{r}_h = \nabla p_h$ in (30a) and reverting the integration by parts, we have

$$\begin{aligned}
\|\nabla p_h\|^2_K &= |(\boldsymbol{W_0}\boldsymbol{\sigma}_h, \nabla p_h)_K + 2i\omega\,(p_h \boldsymbol{W_0}\boldsymbol{b_0}, \nabla p_h)_K + \langle p_h, \nabla p_h \cdot \boldsymbol{n}\rangle_{\partial K}| \\
&\lesssim \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K} \|\nabla p_h\|_K + \omega \|p_h\|_K \|\nabla p_h\|_K + \|p_h\|_{\partial K} \|\nabla p_h\|_{\partial K} \\
\text{(by Lemma 3.1)} \quad &\lesssim \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K} \|\nabla p_h\|_K + \omega \|p_h\|_K \|\nabla p_h\|_K + h_K^{-1/2} \|p_h\|_{\partial K} \|\nabla p_h\|_K
\end{aligned}$$

So we have

$$\begin{aligned}
\|\nabla p_h\|_K &\lesssim \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K} + \omega \|p_h\|_K + h_K^{-1/2} \|p_h\|_{\partial K} \\
&\lesssim \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K} + \omega \|p_h\|_K + (1+\omega h_K) h_K \|\nabla p_h\|_K \\
&\lesssim \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K} + \omega h_K \|\nabla p_h\|_K + \omega h_K^{3/2} \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K} + (1+\omega h_K) h_K \|\nabla p_h\|_K
\end{aligned}$$

Finally we have

$$\|\nabla p_h\|_K \lesssim \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K} \lesssim \|\boldsymbol{\sigma}_h\|_K . \tag{37}$$

Step 6: Contradiction
Combining (36) and (37), we have

$$\|\boldsymbol{\sigma}_h\|^2_K \lesssim \omega h_K \|\boldsymbol{\sigma}_h\|^2_K$$

as we assumed that $\boldsymbol{\sigma}_h \neq 0$, we can divide by $\|\boldsymbol{\sigma}_h\|_K$, leading to

$$1 \lesssim \omega h_K,$$

which does not hold if ωh_K is small enough.
This is the desired contradiction, and we can therefore conclude that $(\boldsymbol{\sigma}_h, p_h) = (\boldsymbol{0}, 0)$ is the only solution of the system (30a)–(30b). ∎

3.4 Error analysis

The error analysis can be carried out by following the projection analysis for the Helmholtz equation given in [DS19, Sec. 3.5.1 & 3.5.2] with some minor changes.
This error analysis relies on the tailored HDG projection that fits the structure of the numerical trace. This projection $(\boldsymbol{\Pi}, \Pi)$, with

$$(\boldsymbol{\Pi}, \Pi) : \boldsymbol{H}_{\mathrm{div}}(\mathcal{O}) \times H^1(\mathcal{O}) \longrightarrow \boldsymbol{V}_h \times W_h := \boldsymbol{\mathcal{P}}_k(\mathcal{T}_h) \times \mathcal{P}_k(\mathcal{T}_h)$$

is defined by the following equations

$$\begin{aligned}
(\boldsymbol{\Pi}\boldsymbol{\sigma}, \boldsymbol{r}_h)_K &= (\boldsymbol{\sigma}, \boldsymbol{r}_h)_K , && \forall \boldsymbol{r}_h \in \boldsymbol{\mathcal{P}}_{k-1}(K), \\
(\Pi p, w_h)_K &= (p, w_h)_K , && \forall w_h \in \mathcal{P}_{k-1}(K), \\
\langle \boldsymbol{\Pi}\boldsymbol{\sigma} \cdot \boldsymbol{n} + i\omega\tau \Pi p, \mu_h\rangle_{\partial K} &= \langle \boldsymbol{\sigma} \cdot \boldsymbol{n} + i\omega\tau p, \mu_h\rangle_{\partial K} , && \forall \mu_h \in \mathcal{R}_k(\partial K).
\end{aligned}$$

Notice that denoting the image of $(\boldsymbol{\sigma}, p)$ under $(\boldsymbol{\Pi}, \Pi)$ by $(\boldsymbol{\Pi}\boldsymbol{\sigma}, \Pi p)$ is a slight abuse of notation as both components depend on $\boldsymbol{\sigma}$ and p. However it is very convenient and often found in the literature.

We define the following error quantities

$$\boldsymbol{\delta}_h^{\boldsymbol{\sigma}} := \boldsymbol{\Pi}\boldsymbol{\sigma} - \boldsymbol{\sigma} \quad ; \quad \delta_h^p := \Pi p - p \quad ; \quad \widehat{\delta}_h^p := p - P_M p$$

and

$$\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}} := \boldsymbol{\Pi}\boldsymbol{\sigma} - \boldsymbol{\sigma}_h \in \boldsymbol{V}_h \quad ; \quad \varepsilon_h^p := \Pi p - p_h \in W_h \quad ; \quad \widehat{\varepsilon}_h^p := P_M p - \widehat{p_h} \in M_h.$$

We will split the errors as

$$\begin{aligned} \|\boldsymbol{\sigma}_h - \boldsymbol{\sigma}\|_{\boldsymbol{W_0},\mathcal{T}_h} &\leqslant \|\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}\|_{\boldsymbol{W_0},\mathcal{T}_h} + \|\boldsymbol{\delta}_h^{\boldsymbol{\sigma}}\|_{\boldsymbol{W_0},\mathcal{T}_h}, \\ \|p_h - p\|_{\rho_0,\mathcal{T}_h} &\leqslant \|\varepsilon_h^p\|_{\rho_0,\mathcal{T}_h} + \|\delta_h^p\|_{\rho_0,\mathcal{T}_h}, \end{aligned}$$

Notice that the following estimates hold

$$\begin{aligned} \|\delta_h^p\|_K &\lesssim h_K^{k+1}\left(|p|_{k+1,K} + \tau_{\max}^{-1}|\operatorname{div}\boldsymbol{\sigma}|_{k,K}\right), \\ \|\boldsymbol{\delta}_h^{\boldsymbol{\sigma}}\|_K &\lesssim h_K^{k+1}\left(|\boldsymbol{\sigma}|_{k+1,K} + \tau^\star|p|_{k+1,K}\right), \end{aligned}$$

for some constants $\tau_{\max}$ and $\tau^\star$. So we only need to prove estimates for $\|\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}\|_{\boldsymbol{W_0},\mathcal{T}_h}$, $\|\varepsilon_h^p\|_{\rho_0,\mathcal{T}_h}$.
The error analysis can be summarized as follows

1. we derive an estimate for $\|\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}\|_{\boldsymbol{W_0},\mathcal{T}_h}$ using the energy-like inequality,
2. we use a *dual problem* to estimate $\|\varepsilon_h^p\|_{\rho_0,\mathcal{T}_h}$ in terms of $\|\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}\|_{\boldsymbol{W_0},\mathcal{T}_h}$,
3. those estimates are combined through a *bootstrapping process*.

This analysis is therefore strongly related to the *Aubin-Nitsche method* and only works for regular solutions.
We will now give the main changes needed to adapt the error analysis from [DS19].
The error equations (3.30) become

$$\begin{aligned} (\boldsymbol{W_0}\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}, \boldsymbol{r}_h)_{\mathcal{T}_h} - (\varepsilon_h^p, \operatorname{div}(\boldsymbol{r}_h))_{\mathcal{T}_h} + 2i\omega\,(\varepsilon_h^p \boldsymbol{W_0}\boldsymbol{b_0}, \boldsymbol{r}_h)_{\mathcal{T}_h} + \langle\widehat{\varepsilon}_h^p, \boldsymbol{r}_h\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= (\boldsymbol{W_0}\boldsymbol{\delta}_h^{\boldsymbol{\sigma}}, \boldsymbol{r}_h)_{\mathcal{T}_h} + 2i\omega\,(\delta_h^p \boldsymbol{W_0}\boldsymbol{b_0}, \boldsymbol{r}_h)_{\mathcal{T}_h} \\ -\omega^2(\rho_0\varepsilon_h^p, w_h)_{\mathcal{T}_h} + (\operatorname{div}(\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}), w_h)_{\mathcal{T}_h} + i\omega\langle\tau(\varepsilon_h^p - \widehat{\varepsilon}_h^p), w_h\rangle_{\partial\mathcal{T}_h} &= -\omega^2(\rho_0\delta_h^p, w_h)_{\mathcal{T}_h} \\ -\langle\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}\cdot\boldsymbol{n} + i\omega\tau(\varepsilon_h^p - \widehat{\varepsilon}_h^p), \mu_h\rangle_{\partial\mathcal{T}_h} &= 0. \end{aligned}$$

The energy-like identity of Prop. 3.7 becomes

$$\begin{aligned} \|\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}\|^2_{\boldsymbol{W_0},\mathcal{T}_h} - \omega^2\|\varepsilon_h^p\|^2_{\rho_0,\mathcal{T}_h} - 2i\omega\,(\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}, \varepsilon_h^p\boldsymbol{W_0}\boldsymbol{b_0})_{\mathcal{T}_h} + i\omega\langle\tau(\varepsilon_h^p - \widehat{\varepsilon}_h^p), \varepsilon_h^p - \widehat{\varepsilon}_h^p\rangle_{\partial\mathcal{T}_h} &= (\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}, \boldsymbol{W_0}\boldsymbol{\delta}_h^{\boldsymbol{\sigma}})_{\mathcal{T}_h} \\ &\quad - 2i\omega\,(\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}, \delta_h^p\boldsymbol{W_0}\boldsymbol{b_0})_{\mathcal{T}_h} \\ &\quad - \omega^2(\rho_0\delta_h^p, \varepsilon_h^p)_{\mathcal{T}_h}, \end{aligned}$$

leading to the following estimate

$$\begin{aligned} \left|\|\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}\|^2_{\boldsymbol{W_0},\mathcal{T}_h} + i\omega\langle\tau(\varepsilon_h^p - \widehat{\varepsilon}_h^p), \varepsilon_h^p - \widehat{\varepsilon}_h^p\rangle_{\partial\mathcal{T}_h}\right| \lesssim{}& \omega^2\|\varepsilon_h^p\|^2_{\rho_0,\mathcal{T}_h} + \omega\|\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}\|_{\boldsymbol{W_0},\mathcal{T}_h}\|\delta_h^p\|_{\rho_0,\mathcal{T}_h} + \|\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}\|_{\boldsymbol{W_0},\mathcal{T}_h}\|\boldsymbol{\delta}_h^{\boldsymbol{\sigma}}\|_{\boldsymbol{W_0},\mathcal{T}_h} \\ &+ \omega\|\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}\|_{\boldsymbol{W_0},\mathcal{T}_h}\|\delta_h^p\|_{\rho_0,\mathcal{T}_h} + \omega^2\|\varepsilon_h^p\|_{\rho_0,\mathcal{T}_h}\|\delta_h^p\|_{\rho_0,\mathcal{T}_h}. \end{aligned}$$

The adjoint problem (3.31) becomes

$$\begin{aligned} \boldsymbol{W_0}\boldsymbol{\xi} - \nabla\theta - 2i\omega\theta\boldsymbol{W_0}\boldsymbol{b_0} &= 0 \\ -\rho_0\omega^2\theta - \operatorname{div}(\boldsymbol{\xi}) &= \varepsilon_h^p. \end{aligned}$$

We state the *elliptic regularity* assumption which is a key ingredient in the error analysis

$$\|\theta\|_{2,\mathcal{O}} + \|\boldsymbol{\xi}\|_{1,\mathcal{O}} \leqslant C_{\text{reg}} \|\varepsilon_h^p\|_{\mathcal{O}}$$

The identity of Prop. 3.8 becomes

$$\begin{aligned}\|\varepsilon_h^p\|^2_{\rho_0,\mathcal{T}_h} =& \omega^2 \left(\rho_0(\Pi\theta - \theta), \varepsilon_h^p - \delta_h^p\right)_{\mathcal{T}_h} - \omega^2 \left(\rho_0\theta - \{\rho_0\theta\}, \delta_h^p\right)_{\mathcal{T}_h} \\ & - \left(\boldsymbol{W_0}(\boldsymbol{\Pi\xi} - \boldsymbol{\xi}), \boldsymbol{\varepsilon_h^\sigma} - \boldsymbol{\delta_h^\sigma}\right)_{\mathcal{T}_h} + \left(\boldsymbol{W_0\xi} - \{\boldsymbol{W_0\xi}\}, \boldsymbol{\delta_h^\sigma}\right)_{\mathcal{T}_h} \\ & - 2i\omega \left(\theta \boldsymbol{b_0}, \boldsymbol{W_0\varepsilon_h^\sigma}\right)_{\mathcal{T}_h} + 2i\omega \left(\boldsymbol{W_0\xi}, (\varepsilon_h^p - \delta_h^p)\boldsymbol{b_0}\right)_{\mathcal{T}_h} \\ & + 2i\omega \left(\boldsymbol{W_0}(\boldsymbol{\Pi\xi} - \boldsymbol{\xi}), (\varepsilon_h^p - \delta_h^p)\boldsymbol{b_0}\right)_{\mathcal{T}_h},\end{aligned}$$

leading to the following estimate

$$\begin{aligned}\|\varepsilon_h^p\|_{\rho_0,\mathcal{T}_h} \lesssim & \left(\omega^3 + \omega^2 + \omega\right) h \left(\|\varepsilon_h^p\|_{\rho_0,\mathcal{T}_h} + \|\delta_h^p\|_{\rho_0,\mathcal{T}_h}\right) + (1+\omega)h \left(\|\boldsymbol{\varepsilon_h^\sigma}\|_{\boldsymbol{W_0},\mathcal{T}_h} + \|\boldsymbol{\delta_h^\sigma}\|_{\boldsymbol{W_0},\mathcal{T}_h}\right) \\ & + \|\boldsymbol{\varepsilon_h^\sigma}\|_{\boldsymbol{W_0},K} + \omega \|\varepsilon_h^p\|_{\rho_0,K} + \omega \|\delta_h^p\|_{\rho_0,K}.\end{aligned}$$

Notice that in contrast to the HDG method for the standard Helmholtz equation, both p_h and $\boldsymbol{\sigma}_h$ have the *same* convergence rate.
It is now straightforward to follow the bootstrapping argument of Sec. 3.5.2 to obtain the

Theorem 2 : *Convergence of the HDG method with total flux*

Assuming that ωh is small enough and under the elliptic regularity assumption, we have

$$\|p_h - \Pi p\|_{\mathcal{T}_h} = \mathcal{O}(h^{k+1}) \quad ; \quad \|\boldsymbol{\sigma}_h - \boldsymbol{\Pi\sigma}\|_{\mathcal{T}_h} = \mathcal{O}(h^{k+1}).$$

Remark 3.5: Some HDG methods are known to achieve *super-convergence*, *ie* taking $p_h \in \mathcal{P}_k$ leads to the following error estimate

$$\|\Pi p - p_h\|_{\mathcal{T}_h} = \mathcal{O}(h^{k+2}).$$

Superconvergence is an attractive property for a numerical scheme, indeed using a postprocessing scheme it is possible to use the solution $(\boldsymbol{\sigma}_h, p_h)$ to construct a new approximation $\widetilde{p_h}$ which converges with order $\mathcal{O}(h^{k+2})$, see [Ste91], [CGS10, Sec. 5] for more details.
Here we were only able to prove *optimal convergence* and the use of post-processing schemes to improve the convergence rate is therefore not possible.

3.5 Global solvability

The analysis that we have carried out in the previous subsection works for any solution $(\boldsymbol{\sigma}_h, p_h, \widehat{p_h})$ of the discrete system (15a)–(15b)–(15c) provided that such solution exists. We already discussed the well-posedness of the local problems in Theorem 1, but we have not yet proved that the global problem (16) for $\widehat{p_h}$ was well-posed.
To do that we can either directly show the well-posedness of the global problem (16). Or we can choose take advantage of the error estimates of Theorem 2 as we will describe below[3].

Resonant frequencies: We recall that the convected Helmholtz equation is a problem of Fredholm type. It is therefore uniquely solvable except on a set of *resonant frequencies*. For those frequencies, there exists non-zero solutions to the homogenous equation and unique solvability cannot be guaranteed.

[3] In [DS19] this idea is attributed to B. Cockburn.

Main result: We can now state and prove the main result of this section.

Theorem 3 : *Global solvability*

Under the assumptions of THEOREM 1 and THEOREM 2 and if ω is not a resonant frequency of the convected Helmholtz equation (1) then the global problem is well-posed, *ie* $\widehat{p_h}$ is uniquely defined by (16).

Proof: First we recall that (15a)–(15b)–(15c), or equivalently (16), is a square system of linear equations, we therefore only need to show the uniqueness of the solution of the homogenous system (when $g_N = g_D = s = 0$).
Assuming that ω is not a resonant frequency of (1), the exact solution is $p = 0$ and $\boldsymbol{\sigma} = \mathbf{0}$, and therefore

$$\|p\|_{s,\mathcal{O}} = 0 \quad \text{and} \quad \|\boldsymbol{\sigma}\|_{t,\mathcal{O}} = 0$$

and

$$\varepsilon_h^p = -p_h \;\; ; \;\; \boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}} = -\boldsymbol{\sigma}_h \;\; ; \;\; \widehat{\varepsilon}_h^p = -\widehat{p_h}.$$

The aim of the error analysis was to prove the following inequalities when h is small enough :

$$\|\varepsilon_h^p\|_{\mathcal{T}_h} \lesssim \|p\|_{s,\mathcal{O}} + \|\boldsymbol{\sigma}_h\|_{t,\mathcal{O}} = 0 \tag{38a}$$

$$\|\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}\|_{\mathcal{T}_h} \lesssim \|p\|_{s,\mathcal{O}} + \|\boldsymbol{\sigma}_h\|_{t,\mathcal{O}} = 0 \tag{38b}$$

Notice that we have hidden the powers of h in $\lesssim$ as they do not play an important part here. Therefore using (38a) and (38b) we have shown that

$$p_h \equiv 0 \quad \text{and} \quad \boldsymbol{\sigma}_h \equiv \mathbf{0}$$

when h is small enough.
For all $K \in \mathcal{T}_h$, we can now rewrite (12a) as

$$\langle \widehat{p_h}, \boldsymbol{r}_h \cdot \boldsymbol{n} \rangle_{\partial K} = 0, \quad \forall \boldsymbol{r}_h \in \boldsymbol{V}_h(K),$$

which leads to

$$\widehat{p_h} \equiv 0.$$

■

4 HDG(+) methods for the diffusive flux formulation

In this section, we will construct HDG methods based on the *diffusive flux* formulation, where $\boldsymbol{q}$ is used instead of $\boldsymbol{\sigma}$. We will mostly describe the HDG+ method where different polynomial degrees are used for the different unknowns, as it the most important novelty of this paper. Adaptation of the formulation construction and theoretical results to a more standard HDG method with the same polynomial interpolation for all the unknowns is straightforward. The main differences between the HDG and HDG+ methods are stated but the details are left out. The *global solvability* will not be included in this section as the adaptation of the result from the previous section is immediate.

4.1 Construction of the method

We recall that on element $K \in \mathcal{T}_h$, the weak formulation (7a)–(7b) reads : Seek $(\boldsymbol{q}, p) \in \boldsymbol{H}_{\mathrm{div}}(\mathcal{O}) \times H^1(\mathcal{O})$ such that

$$\int_K \boldsymbol{W_0 q} \cdot \boldsymbol{r}^* \mathrm{d}\boldsymbol{x} - \int_K p \mathrm{div}\,(\boldsymbol{r}^*)\,\mathrm{d}\boldsymbol{x} + \int_{\partial K} p \boldsymbol{r}^* \cdot \boldsymbol{n} \mathrm{d}\sigma = 0, \tag{39a}$$

$$-\omega^2 \int_K \rho_0 p w^* \mathrm{d}\boldsymbol{x} - 2i\omega \int_K \boldsymbol{b_0} \cdot \nabla p w^* \mathrm{d}\boldsymbol{x} + \int_K \mathrm{div}\,(\boldsymbol{q})\, w^* \mathrm{d}\boldsymbol{x} = \int_K s w^* \mathrm{d}\boldsymbol{x}, \tag{39b}$$

for all $(\boldsymbol{r}, w) \in \boldsymbol{H}_{\mathrm{div}}(\mathcal{O}) \times H^1(\mathcal{O})$.

Choice of approximation spaces: For the HDG+ method, the choice of approximation spaces is different from the choice made for the previous HDG method. We consider the following *local approximation spaces*

$$\begin{aligned} \boldsymbol{V}_h(K) &= \boldsymbol{\mathcal{P}}_k(K), && \text{for the flux } \boldsymbol{q}_h, \\ W_h(K) &= \mathcal{P}_{k+1}(K), && \text{for the pressure } p_h, \end{aligned}$$

where $k \geqslant 2$ is the degree of the method. The use of a higher polynomial degree for p_h is the distinctive feature of the HDG+ method.

Introduction of the hybrid unknown: As we did before, we introduce the *numerical trace* $\widehat{p_h}$ which approximates p on the skeleton $\mathcal{E}_h$ of the mesh. As before the boundary integral in (39a) will be discretized as

$$\int_{\partial K} p \boldsymbol{r}^* \cdot \boldsymbol{n} \mathrm{d}\sigma \quad \text{becomes} \quad \int_{\partial K} \widehat{p_h} \boldsymbol{r}_h^* \cdot \boldsymbol{n} \mathrm{d}\sigma.$$

For the HDG+ method, we use the following approximation space for $\widehat{p_h}$

$$M_h(e) = \mathcal{P}_k(e), \quad \forall e \in \mathcal{E}(K).$$

With this choice, p_h and $\widehat{p_h}$ do not have the same polynomial degree and we therefore have two approximations of p with different polynomial degrees on the skeleton of the mesh. We therefore need to change the *penalization term* to

$$\tau(P_M p_h - \widehat{p_h}), \tag{40}$$

where P_M is the L^2-orthogonal projection onto M_h. This is called the *reduced stabilization* and it was introduced in [Leh10]. It allows to get convergence rate of $k+2$ for p_h for the cost of a method of degree k. A large penalization parameter $\tau \sim {h_K}^{-1}$ is needed to obtain optimal convergence as it will be detailed in Subsection 4.3.

Local problem: We approximate the weak formulation (39a)–(39b) on an element $K \in \mathcal{T}_h$ leading to the so-called *local problem* : seek $(\boldsymbol{q}_h, p_h) \in \boldsymbol{V}_h(K) \times W_h(K)$ such that

$$(\boldsymbol{W_0 q}_h, \boldsymbol{r}_h)_K - (p_h, \mathrm{div}\,(\boldsymbol{r}_h))_K + \langle \widehat{p_h}, \boldsymbol{r}_h \cdot \boldsymbol{n} \rangle_{\partial K} = 0, \tag{41a}$$

$$\begin{aligned} -\omega^2 (\rho_0 p_h, w_h)_K - 2i\omega (\boldsymbol{b_0} \cdot \nabla p_h, w_h)_K + (\mathrm{div}\,(\boldsymbol{q}_h), w_h)_K \\ + 2i\omega \langle \tau(P_M p_h - \widehat{p_h}) - \tau_{\mathrm{upw}}(p_h - \widehat{p_h}), w_h \rangle_{\partial K} = (s, w_h)_K, \end{aligned} \tag{41b}$$

for all $(\boldsymbol{r}_h, w_h) \in \boldsymbol{V}_h(K) \times W_h(K)$.
Following [QS16a], we have introduced a second penalization parameter τ_{upw} defined by

$$\tau_{\mathrm{upw}} := \max(\boldsymbol{b_0} \cdot \boldsymbol{n}, 0).$$

To understand why this second parameter is required, we recall that in HDG methods the penalization serves two purposes:

1. it enforces the Dirichlet boundary condition for the local problems,
2. it controls the stability of the method.

Here as $\boldsymbol{q}_h$ does not take the convection into account, the penalization term (40) with τ only stabilizes the diffusion. We therefore need to add a second penalization to stabilize the convection. We denoted it τ_{upw} as it leads to an upwinding behavior that will be detailed in the next paragraph.

Transmission condition: Following the previous example, we introduce the *following numerical flux*

$$\widehat{\boldsymbol{q}_h} \cdot \boldsymbol{n} := \boldsymbol{q}_h \cdot \boldsymbol{n} + 2i\omega\tau \left(P_M p_h - \widehat{p_h}\right), \tag{42}$$

where $\tau = \mathcal{O}({h_K}^{-1})$. As discussed before, we need to require the normal continuity of the *total flux* on the interface between two elements, ands the quantity $\widehat{\boldsymbol{q}_h} \cdot \boldsymbol{n}$ only takes the diffusion into account. To deal with convection we add a second numerical flux

$$2i\omega\widehat{p_h\boldsymbol{b_0}} \cdot \boldsymbol{n} := 2i\omega(\boldsymbol{b_0} \cdot \boldsymbol{n})\widehat{p_h} + 2i\omega\tau_{\text{upw}}(p_h - \widehat{p_h}). \tag{43}$$

It is important to notice that this flux has an upwind behavior. Let $e = \partial K_+ \cap \partial K_-$ be an interior edge with $\boldsymbol{b_0} \cdot \boldsymbol{n}_- > 0$ on ∂K_-. We have

$$\text{On } \partial K_-: \quad \tau_{\text{upw}} := \max(\boldsymbol{b_0} \cdot \boldsymbol{n}, 0) = \boldsymbol{b_0} \cdot \boldsymbol{n}, \quad \text{so} \quad 2i\omega\widehat{p_h\boldsymbol{b_0}} \cdot \boldsymbol{n} = 2i\omega(\boldsymbol{b_0} \cdot \boldsymbol{n})p_h \tag{44a}$$

$$\text{On } \partial K_+: \quad \tau_{\text{upw}} := \max(\boldsymbol{b_0} \cdot \boldsymbol{n}, 0) = 0, \quad \text{so} \quad 2i\omega\widehat{p_h\boldsymbol{b_0}} \cdot \boldsymbol{n} = 2i\omega(\boldsymbol{b_0} \cdot \boldsymbol{n})\widehat{p_h} \tag{44b}$$

So on the outflow boundary we use the interior value p_h, whereas on the inflow boundary we use the trace value $\widehat{p_h}$.
Finally we write the *transmission condition* as

$$\left\langle \left(\widehat{\boldsymbol{q}_h} - 2i\omega\widehat{p_h\boldsymbol{b_0}}\right) \cdot \boldsymbol{n}, \mu_h \right\rangle_{\partial\mathcal{T}_h \setminus \Gamma_D} + \langle \widehat{p_h} - g_D, \mu_h \rangle_{\Gamma_D} = \langle g_N, \mu_h \rangle_{\Gamma_N}. \tag{45}$$

This formulation enforces normal continuity of the total flux between the elements and the boundary conditions on Γ_D and Γ_N.

Remark 4.1: To ensure the well-posedness of the *local problems*, the second penalization must be

$$\tau_{\text{upw}}(p_h - \widehat{p_h}),$$

and not

$$\tau_{\text{upw}}(P_M p_h - \widehat{p_h}),$$

see Theorem 4.

Remark 4.2: We would like to point out the main theoretical difficulty of this method : when the background flow is not constant, the second flux (43) leads to non-polynomial terms on the skeleton. This is usually avoided as much as possible in HDG methods.

Adaptation to a standard HDG method: With this formulation, it is also possible to consider a standard HDG method by using the same polynomial degree for p_h and $\boldsymbol{q}_h$, *ie.* by using the following local approximation spaces

$$\begin{aligned} \boldsymbol{V}_h(K) &= \boldsymbol{\mathcal{P}}_k(K), && \text{for the flux } \boldsymbol{q}_h, \\ W_h(K) &= \mathcal{P}_k(K), && \text{for the pressure } p_h. \end{aligned}$$

In this case, as W_h and M_h have the same polynomial degree, the projection term becomes simpler, indeed

$$P_M p_h = p_h.$$

For this formulation, we do not require a large penalization parameter anymore, and we only need $\tau = \mathcal{O}(1)$.

4.1.1 Compact formulation of the methods:

HDG methods are usually stated in a compact way that can be obtained by summing the local problems (41a)–(41b) over the mesh elements and by adding the transmission condition (45). This formulation reads : seek $(\boldsymbol{q}_h, p_h, \widehat{p_h}) \in \boldsymbol{V}_h \times W_h \times M_h$, such that

$$(\boldsymbol{W_0} \boldsymbol{q}_h, \boldsymbol{r}_h)_{\mathcal{T}_h} - (p_h, \operatorname{div}(\boldsymbol{r}_h))_{\mathcal{T}_h} + \langle \widehat{p_h}, \boldsymbol{r}_h \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} = 0, \tag{46a}$$

$$-\omega^2 (\rho_0 p_h, w_h)_{\mathcal{T}_h} - 2i\omega (\boldsymbol{b_0} \cdot \nabla p_h, w_h)_{\mathcal{T}_h} + (\operatorname{div}(\boldsymbol{q}_h), w_h)_{\mathcal{T}_h} \tag{46b}$$
$$+2i\omega \langle \tau(P_M p_h - \widehat{p_h}) - \tau_{\text{upw}}(p_h - \widehat{p_h}), w_h \rangle_{\partial \mathcal{T}_h} = (s, w_h)_{\mathcal{T}_h}$$

$$\left\langle \left(\widehat{\boldsymbol{q}_h} - 2i\omega \widehat{p_h \boldsymbol{b_0}} \right) \cdot \boldsymbol{n}, \mu \right\rangle_{\partial \mathcal{T}_h \setminus \Gamma_D} + \langle \widehat{p_h} - g_D, \mu_h \rangle_{\Gamma_D} = \langle g_N, \mu_h \rangle_{\Gamma_N}, \tag{46c}$$

for all $(\boldsymbol{r}_h, w_h, \mu_h) \in \boldsymbol{V}_h \times W_h \times M_h$.

4.1.2 Condensed variational formulation

The compact formulation (46a)–(46b)–(46c) cannot be used to efficiently implement the HDG method, indeed with this formulation it is not clear how the *global problem* for $\widehat{p_h}$ only can be obtained. To describe this process , we will now write a *condensed* variational formulation for $\widehat{p_h}$ only.
We introduce the so-called *local solvers*

$$\begin{aligned} \mathsf{P}^K &: (\widehat{p_h}, s) \longmapsto p_h^K, \\ \mathsf{Q}^K &: (\widehat{p_h}, s) \longmapsto \boldsymbol{q}_h^K, \\ \widehat{\mathsf{Q}}^K &: (\widehat{p_h}, s) \longmapsto \widehat{\boldsymbol{q}_h}^K, \end{aligned}$$

where $(\boldsymbol{q}_h^K, p_h^K)$ is the solution of (41a)–(41b) and $\widehat{\boldsymbol{q}_h}^K$ is defined by (42).
We can therefore rewrite the transmission condition (46c) as

$$a_h(\widehat{p_h}, \mu) = \ell_h(\mu), \tag{47}$$

where

$$\begin{aligned} a_h(\widehat{p_h}, \mu_h) := &\left\langle \mathsf{Q}^K(\widehat{p_h}, s) \cdot \boldsymbol{n} + 2i\omega\tau(P_M \mathsf{P}^K(\widehat{p_h}, s) - \widehat{p_h}) + 2i\omega\tau_{\text{upw}}(\mathsf{P}^K(\widehat{p_h}, s) - \widehat{p_h}), \mu_h \right\rangle_{\partial \mathcal{T}_h \setminus \Gamma_D} \\ &+ \langle \widehat{p_h}, \mu_h \rangle_{\Gamma_D}, \\ \ell_h(\mu_h) := &\langle g_N, \mu_h \rangle_{\Gamma_N} + \langle g_D, \mu_h \rangle_{\Gamma_D}. \end{aligned}$$

Equation (47) is the so-called *global problem* and is the main equation of the HDG method. From a computational point of view, we proceed as described in Algorithm 2.

Algorithm 2: Solving HDG+

for $K \in \mathcal{T}_h$ **do**
  Construct the local solvers P^K, Q^K, $\widehat{\mathsf{Q}}^K$
  Add local contribution to the global problem (16)
Solve (16) for $\widehat{p_h}$ `// This is the main step`
for $K \in \mathcal{T}_h$ **do**
  Reconstruct the local unknowns $p_h^K = \mathsf{P}^k(\widehat{p_h}, s)$ and $\boldsymbol{q}_h^K = \mathsf{Q}^K(\widehat{p_h}, s)$

4.2 Local solvability

It is worth remembering that HDG methods were originally developed for elliptic problems and that harmonic wave equations are only coercive. It is well-known that solving those equations with Dirichlet boundary conditions[4] leads to numerical pollution due to the resonance phenomenon. In this section we will show that the static condensation process is well-defined when the mesh is fine enough, *ie.* the local problem does not produce resonance.
Before actually showing the local solvability, we need to prove the following lemma.

Lemma 4.1:
If $p \in H^1(K)$ and $\boldsymbol{b_0} \in \boldsymbol{L}^\infty(K) \cap \boldsymbol{\mathcal{C}}(\mathcal{O})$, where $\boldsymbol{\mathcal{C}}(\mathcal{O})$ is the space of vector functions continuous in the domain $\mathcal{O}$, then the following identity holds

$$\mathfrak{Re}\left(p\boldsymbol{b_0}, \nabla p\right)_K = \frac{1}{2}\left\langle (\boldsymbol{b_0}\cdot\boldsymbol{n})p, p\right\rangle_{\partial K}.$$

Proof: We use an integration by parts to obtain a relationship between $\left(p\boldsymbol{b_0}, \nabla p\right)_K$ and its complex conjugate :

$$\begin{aligned}
2\mathfrak{Re}\left(p\boldsymbol{b_0}, \nabla p\right)_K &= \left(p\boldsymbol{b_0}, \nabla p\right)_K + \left(p\boldsymbol{b_0}, \nabla p\right)_K^* \\
&= \left(p\boldsymbol{b_0}, \nabla p\right)_K + \left(\nabla p, p\boldsymbol{b_0}\right)_K \\
&= -\left(\operatorname{div}\left(p\boldsymbol{b_0}\right), p\right)_K + \left\langle (\boldsymbol{b_0}\cdot\boldsymbol{n})p, p\right\rangle_{\partial K} + \left(\nabla p, p\boldsymbol{b_0}\right)_K \\
(\operatorname{div}(\boldsymbol{b_0}) = 0) \quad &= -\left(\nabla p, p\boldsymbol{b_0}\right)_K + \left\langle (\boldsymbol{b_0}\cdot\boldsymbol{n})p, p\right\rangle_{\partial K} + \left(\nabla p, p\boldsymbol{b_0}\right)_K \\
&= \left\langle (\boldsymbol{b_0}\cdot\boldsymbol{n})p, p\right\rangle_{\partial K}.
\end{aligned}$$

■

We can now state and prove the main result of this section.

Theorem 4 : *Local solvability for the HDG+ methods*

If

$$\forall e \in \mathcal{E}(K), \quad \tau|_e < 0 \tag{48}$$

and if

$$\omega h_K < \frac{-C_{\boldsymbol{W_0},K}\left\|\boldsymbol{b_0}\right\|_{L^\infty(K)} + \left(C^2_{\boldsymbol{W_0},K}\left\|\boldsymbol{b_0}\right\|^2_{L^\infty(K)} + \left\|\rho_0\right\|_{L^\infty(K)}\right)^{\frac{1}{2}}}{C_{\boldsymbol{W_0},K} C \left\|\rho_0\right\|_{L^\infty(K)}} \tag{49}$$

where $C > 0$ is a constant that depends only on the shape regularity of K, then the local solver

$$(\widehat{p_h}, s) \longmapsto (p_h, \boldsymbol{q}_h)$$

is well-posed.

[4] Which is what the local solver does.

Proof:
As the local problems have a finite dimension, we only need to prove uniqueness of the solution. We therefore assume that $\widehat{p_h} = s = 0$ and we need to prove that the system

$$\left(\boldsymbol{W_0} \boldsymbol{q}_h, \boldsymbol{r}_h\right)_K - \left(p_h, \operatorname{div}\left(\boldsymbol{r}_h\right)\right)_K = 0, \tag{50a}$$

$$-\omega^2 \left(\rho_0 p_h, w_h\right)_K - 2i\omega \left(\boldsymbol{b_0} \cdot \nabla p_h, w_h\right)_K + \left(\operatorname{div}\left(\boldsymbol{q}_h\right), w_h\right)_K + 2i\omega \left\langle \tau P_M p_h - \tau_{\text{upw}} p_h, w_h \right\rangle_{\partial K} = 0, \tag{50b}$$

has only one solution : $(p_h, \boldsymbol{q}_h) = (0, \boldsymbol{0})$.
We will prove the theorem by contradiction. We therefore assume that the system (50a)–(50b) has a non-zero solution $(p_h, \boldsymbol{q}_h)$.
Step 1: An energy-like system
We begin by testing (50b) with $w_h = p_h$

$$-\omega^2 \left\|p_h\right\|^2_{\rho_0,K} + 2i\omega \left(p_h \boldsymbol{b_0}, \nabla p_h\right)_K + \left(\operatorname{div}\left(\boldsymbol{q}_h\right), p_h\right)_K + \left\langle 2i\omega\tau P_M p_h - 2i\omega\tau_{\text{upw}} p_h, p_h \right\rangle_{\partial K} = 0 \tag{51}$$

Then, (50a) is tested with $\boldsymbol{r}_h = \boldsymbol{q}_h$ and conjugated :

$$\left\|\boldsymbol{q}_h\right\|^2_{\boldsymbol{W_0},K} - \left(\operatorname{div}\left(\boldsymbol{q}_h\right), p_h\right)_K = 0 \tag{52}$$

We now add (51) and (52) leading to

$$\left\|\boldsymbol{q}_h\right\|^2_{\boldsymbol{W_0},K} - \omega^2 \left\|p_h\right\|^2_{\rho_0,K} + 2i\omega \left(p_h \boldsymbol{b_0}, \nabla p_h\right)_K + \left\langle 2i\omega\tau P_M p_h - 2i\omega\tau_{\text{upw}} p_h, p_h \right\rangle_{\partial K} = 0 \tag{53}$$

We now obtain the following system by taking the real and imaginary parts of (53)

$$\mathfrak{Re} : \quad \left\|\boldsymbol{q}_h\right\|^2_{\boldsymbol{W_0},K} - \omega^2 \left\|p_h\right\|^2_{\rho_0,K} - 2\omega \mathfrak{Im}\left(p_h \boldsymbol{b_0}, \nabla p_h\right)_K = 0 \tag{54}$$

$$\mathfrak{Im} : \quad \mathfrak{Re}\left(p_h \boldsymbol{b_0}, \nabla p_h\right)_K + \left\langle \tau P_M p_h, p_h \right\rangle_{\partial K} = \left\langle \tau_{\text{upw}} p_h, p_h \right\rangle_{\partial K} \tag{55}$$

Indeed, as $P_M p_h \in M_h$ and τ is constant on each edge, one has

$$\left\langle \tau P_M p_h, p_h \right\rangle_{\partial K} = \left\langle \tau P_M p_h, P_M p_h \right\rangle_{\partial K} \in \mathbb{R}$$

Step 2: We focus on (55) to express $p_h|_{\partial K}$.
By the Lemma 4.1 we have

$$\mathfrak{Re}\left(p_h \boldsymbol{b_0}, \nabla p_h\right)_K = \frac{1}{2} \left\langle (\boldsymbol{b_0} \cdot \boldsymbol{n}) p_h, p_h \right\rangle_{\partial K}$$

and (55) becomes

$$\frac{1}{2} \left\langle (\boldsymbol{b_0} \cdot \boldsymbol{n}) p_h, p_h \right\rangle_{\partial K} + \left\langle \tau P_M p_h, p_h \right\rangle_{\partial K} = \left\langle \tau_{\text{upw}} p_h, p_h \right\rangle_{\partial K}$$

For the sake of simplicity, we assume that the sign of $\boldsymbol{b_0} \cdot \boldsymbol{n}$ is constant on each edge. It amounts to assuming that h_K is small enough. For a given edge $e \in \mathcal{E}(K)$, the three following cases are exhaustive:

- *Case 1:* $\boldsymbol{b_0} \cdot \boldsymbol{n} < 0$: therefore $\tau_{\text{upw}} := \max(\boldsymbol{b_0} \cdot \boldsymbol{n}, 0) = 0$ and

$$\underbrace{\left\langle \tau P_M p_h, P_M p_h \right\rangle_{\partial K}}_{\leqslant 0 \text{ by (48) as } \tau < 0} \overbrace{-\frac{1}{2} \left\langle |\boldsymbol{b_0} \cdot \boldsymbol{n}| p_h, p_h \right\rangle_{\partial K}}^{\leqslant 0} = 0$$

- *Case 2:* $\boldsymbol{b_0}\cdot\boldsymbol{n} > 0$: therefore $\tau_{\text{upw}} := \max(\boldsymbol{b_0}\cdot\boldsymbol{n}, 0) = \boldsymbol{b_0}\cdot\boldsymbol{n}$ and

$$\begin{aligned} &\frac{1}{2}\langle|\boldsymbol{b_0}\cdot\boldsymbol{n}|p_h, p_h\rangle_{\partial K} + \langle\tau P_M p_h, P_M p_h\rangle_{\partial K} = \langle|\boldsymbol{b_0}\cdot\boldsymbol{n}|p_h, p_h\rangle_{\partial K} \\ \iff \quad &\underbrace{\langle\tau P_M p_h, P_M p_h\rangle_{\partial K}}_{\leqslant 0 \text{ by (48) as } \tau<0} \overbrace{-\frac{1}{2}\langle|\boldsymbol{b_0}\cdot\boldsymbol{n}|p_h, p_h\rangle_{\partial K}}^{\leqslant 0} = 0 \end{aligned}$$

- *Case 3:* $\boldsymbol{b_0}\cdot\boldsymbol{n} = 0$: in this case we only have

$$\langle\tau P_M p_h, P_M p_h\rangle_{\partial K} = 0.$$

In the first two cases we have $p_h|_{\partial K} = P_M p_h = 0$ and in the third one we only have $P_M p_h = 0$. In particular, the following identity holds for all the three previous cases

$$\int_{\partial K} p_h \mathrm{d}\sigma = 0,$$

indeed as $P_M p_h$ is the L^2-orthogonal projection of p_h onto M_h, we have

$$\int_{\partial K} P_M p_h \mu_h^* \mathrm{d}\sigma = \int_{\partial K} p_h \mu_h^* \mathrm{d}\sigma, \quad \forall \mu_h \in M_h := \prod_{e\in\mathcal{E}(K)} \mathcal{P}_k(e),$$

and the previous identity is obtained by taking $\mu_h = 1$.

Step 3: Contradiction

As $p_h \in \mathcal{P}_{k+1}(K)$, we have $p_h \in H^1(K)$ and the following Poincaré-Friedrichs inequality holds[5]

$$\|p_h\|_K \leqslant C h_K \|\nabla p_h\|_K,$$

see [EG04, Lemma B.66] with $f(v) = \int_{\partial K} v \mathrm{d}\sigma$. The constant C is the same one as in (49).

Going back to (50a), integrating by parts and testing it with $\boldsymbol{r}_h = \nabla p_h$ we have

$$\begin{aligned} \|\nabla p_h\|_K^2 &= |(\boldsymbol{W_0}\boldsymbol{q}_h, \nabla p_h)_K| \\ &\leqslant C_{\boldsymbol{W_0},K} \|\boldsymbol{q}_h\|_{\boldsymbol{W_0},K} \|\nabla p_h\|_K \\ \|\nabla p_h\|_K &\leqslant C_{\boldsymbol{W_0},K} \|\boldsymbol{q}_h\|_{\boldsymbol{W_0},K} \end{aligned} \tag{56}$$

On the other hand, from (54) we see that

$$\begin{aligned} \|\boldsymbol{q}_h\|_{\boldsymbol{W_0},K}^2 &= \omega^2 \|p_h\|_{\rho_0,K}^2 + 2\omega\mathfrak{Im}\,(p_h\boldsymbol{b_0}, \nabla p_h)_K \\ &\leqslant \omega^2 \|\rho_0\|_{L^\infty(K)} \|p_h\|_K^2 + 2\omega \|\boldsymbol{b_0}\|_{L^\infty(K)} \|p_h\|_K \|\nabla p_h\|_K \\ \|\boldsymbol{q}_h\|_{\boldsymbol{W_0},K}^2 &\leqslant C^2 \|\rho_0\|_{L^\infty(K)} \omega^2 h_K^2 \|\nabla p_h\|_K^2 + 2C \|\boldsymbol{b_0}\|_{L^\infty(K)} \omega h_K \|\nabla p_h\|_K^2 \end{aligned} \tag{57}$$

Combining (56) and (57) we have

$$\|\nabla p_h\|_K^2 \leqslant C_{\boldsymbol{W_0},K}^2 \left[C^2 \|\rho_0\|_{L^\infty(K)} \omega^2 h_K^2 \|\nabla p_h\|_K^2 + 2C \|\boldsymbol{b_0}\|_{L^\infty(K)} \omega h_K \|\nabla p_h\|_K^2\right],$$

as we assumed $(\boldsymbol{q}_h, p_h) \neq (\boldsymbol{0}, 0)$ we can divide by $\|\nabla p_h\|_K$ to obtain

$$1 \leqslant C_{\boldsymbol{W_0},K}^2 \left[C^2 \|\rho_0\|_{L^\infty(K)} \omega^2 h_K^2 + 2C \|\boldsymbol{b_0}\|_{L^\infty(K)} \omega h_K\right]. \tag{58}$$

[5] When $\boldsymbol{b_0}\cdot\boldsymbol{n} \neq 0$ we can use the standard Poincaré inequality instead.-

We now define the function

$$f : \alpha \longmapsto C^2_{\boldsymbol{W_0},K} C^2 \left\|\rho_0\right\|_{L^\infty(K)} \alpha^2 + 2C^2_{\boldsymbol{W_0},K} C \left\|\boldsymbol{b_0}\right\|_{L^\infty(K)} \alpha - 1$$

Rewriting (58) in terms of f gives

$$f(\omega h_K) \geqslant 0.$$

We notice that f is a second-order polynomial whose roots are

$$\alpha_\pm = \frac{-C_{\boldsymbol{W_0},K} \left\|\boldsymbol{b_0}\right\|_{L^\infty(K)} \pm \left(C^2_{\boldsymbol{W_0},K} \left\|\boldsymbol{b_0}\right\|^2_{L^\infty(K)} + \left\|\rho_0\right\|_{L^\infty(K)}\right)^{\frac{1}{2}}}{C_{\boldsymbol{W_0},K} C \left\|\rho_0\right\|_{L^\infty(K)}}$$

As the leading coefficient of f is positive, we know that

$$\forall \alpha \in (\alpha_-, \alpha_+), \quad f(\alpha) < 0$$

and it is obvious that $\alpha_- < 0$ and $\alpha_+ > 0$.
Finally, we can see that the assumption on ωh_K (49) is exactly

$$0 < \omega h_K < \alpha_+,$$

which means

$$f(\omega h_K) < 0.$$

This is the desired contradiction and concludes the proof, as we necessarily have $p_h \equiv 0$ and $\boldsymbol{q}_h \equiv \boldsymbol{0}$. ∎

Remark 4.3: For triangular elements, the constant C satisfies

$$C < \frac{1}{\pi}.$$

Remark 4.4: when $\boldsymbol{b_0} = \boldsymbol{0}$, the solvability assumption (48) becomes

$$\omega h_K < \frac{1}{C C_{\boldsymbol{W_0},K} \left\|\rho_0\right\|^{\frac{1}{2}}_{L^\infty(K)}}$$

which is similar to the ones given in [DS19, Prop. 3.9] and [Hun19, Prop. 3.4.2].

Remark 4.5: this proof is written for the HDG+ method, for the more standard HDG method only minor changes are needed : in *Step 2*, $P_M p$ should be replaced by p. Assumption (48) can therefore be replaced with

$$\forall e \in \mathcal{E}(K), \quad \tau|_e < 0 \quad \text{or} \quad \tau|_e > \max_e \left(\frac{1}{2}|\boldsymbol{b_0} \cdot \boldsymbol{n}|\right).$$

4.3 Error analysis of the HDG+ method

In this section we will carry out a detailed error analysis of the HDG+ method. The adaptation of this process to the HDG method is straightforward, see Subsection 4.4.
We chose to use the orthogonal L^2 projections instead of the tailored HDG(+) projections that fit the numerical trace. As we study problems involving convection, the design of a new projection would be required as using the standard HDG(+) projection would not lead to cleaner error system. The design of such a projection seems very difficult when $\boldsymbol{b_0}$ is not constant.

We denote by $\boldsymbol{\pi}_{\boldsymbol{V}}$, π_W and P_M the L^2-orthogonal projections onto $\boldsymbol{V}_h$, W_h and M_h respectively. We recall the following estimates due to standard approximation theory for polynomials and trace inequalities which will be useful for our analysis, see *eg.* [EG04, Prop. 1.135] :

$$\|p - \pi_W p\|_{\mathcal{O}} \lesssim h^s \|p\|_{s,\mathcal{O}}, \qquad 0 \leqslant s \leqslant k+2, \tag{59a}$$
$$\|\boldsymbol{q} - \boldsymbol{\pi}_{\boldsymbol{V}} \boldsymbol{q}\|_{\mathcal{O}} \lesssim h^t \|\boldsymbol{q}\|_{t,\mathcal{O}}, \qquad 0 \leqslant t \leqslant k+1, \tag{59b}$$
$$\|p - P_M p\|_{\partial\mathcal{T}_h} \lesssim h^{s-\frac{1}{2}} \|p\|_{s,\mathcal{O}}, \qquad 1 \leqslant s \leqslant k+1, \tag{59c}$$
$$\|p - \pi_W p\|_{\partial K} \lesssim h^{s-\frac{1}{2}} \|p\|_{s,K}, \qquad 1 \leqslant s \leqslant k+2, \tag{59d}$$
$$\|\boldsymbol{q}\cdot\boldsymbol{n} - \boldsymbol{\pi}_{\boldsymbol{V}}\boldsymbol{q}\cdot\boldsymbol{n}\|_{\partial K} \lesssim h^{t-\frac{1}{2}} \|\boldsymbol{q}\|_{t,K}, \qquad 1 \leqslant t \leqslant k+1, \tag{59e}$$

where $a \lesssim b$ means that there exists a constant $C > 0$ independent of the mesh size and frequency such that $a \leqslant Cb$.
We will also frequently use the following inverse inequality

$$\|w\|_{\partial K} \lesssim h^{-\frac{1}{2}} \|w\|_K, \qquad \forall w \in W_h. \tag{60}$$

4.3.1 Error equations

Let $(p, \boldsymbol{q})$ be the solution of the original problem (7a)–(7b). We define the projection errors

$$\boldsymbol{\delta}_h^{\boldsymbol{q}} := \boldsymbol{\pi}_{\boldsymbol{V}}\boldsymbol{q} - \boldsymbol{q} \quad ; \quad \delta_h^p := \pi_W p - p \quad ; \quad \widehat{\delta}_h^p := p - P_M p$$

and

$$\boldsymbol{\varepsilon}_h^{\boldsymbol{q}} := \boldsymbol{\pi}_{\boldsymbol{V}}\boldsymbol{q} - \boldsymbol{q}_h \in \boldsymbol{V}_h \quad ; \quad \varepsilon_h^p := \pi_W p - p_h \in W_h \quad ; \quad \widehat{\varepsilon}_h^p := P_M p - \widehat{p_h} \in M_h$$

Lemma 4.2:
The error quantities $(\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}, \varepsilon_h^p, \widehat{\varepsilon}_h^p)$ satisfy the following error equations:

$$(\boldsymbol{W_0}\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}, \boldsymbol{r}_h)_{\mathcal{T}_h} - (\varepsilon_h^p, \operatorname{div}(\boldsymbol{r}_h))_{\mathcal{T}_h} + \langle \widehat{\varepsilon}_h^p, \boldsymbol{r}_h \cdot \boldsymbol{n} \rangle_{\partial\mathcal{T}_h} = (\boldsymbol{W_0}\boldsymbol{\delta}_h^{\boldsymbol{q}}, \boldsymbol{r}_h)_{\mathcal{T}_h} \tag{61a}$$

$$-\omega^2 (\rho_0 \varepsilon_h^p, w_h)_{\mathcal{T}_h} + 2i\omega (\varepsilon_h^p \boldsymbol{b_0}, \nabla w_h)_{\mathcal{T}_h} - (\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}, \nabla w_h)_{\mathcal{T}_h} + \left\langle \widehat{\mathcal{Q}} \cdot \boldsymbol{n} - \widehat{\mathcal{Q}_h} \cdot \boldsymbol{n}, w_h \right\rangle_{\partial\mathcal{T}_h} =$$
$$-\omega^2 (\rho_0 \delta_h^p, w_h)_{\mathcal{T}_h} + 2i\omega (\delta_h^p \boldsymbol{b_0}, w_h)_{\mathcal{T}_h} \tag{61b}$$

$$\left\langle \widehat{\mathcal{Q}} \cdot \boldsymbol{n} - \widehat{\mathcal{Q}_h} \cdot \boldsymbol{n}, \mu_h \right\rangle_{\partial\mathcal{T}_h} = 0 \tag{61c}$$

where

$$\widehat{\mathcal{Q}} \cdot \boldsymbol{n} = \boldsymbol{q} \cdot \boldsymbol{n} - 2i\omega(\boldsymbol{b_0} \cdot \boldsymbol{n})p \quad \text{on } \partial\mathcal{T}_h$$

and

$$\widehat{\mathcal{Q}} \cdot \boldsymbol{n} - \widehat{\mathcal{Q}_h} \cdot \boldsymbol{n} = \boldsymbol{\varepsilon}_h^{\boldsymbol{q}} \cdot \boldsymbol{n} - 2i\omega(\boldsymbol{b_0} \cdot \boldsymbol{n})\widehat{\varepsilon}_h^p + 2i\omega\tau (P_M \varepsilon_h^p - \widehat{\varepsilon}_h^p) - 2i\omega\tau_{\text{upw}} (\varepsilon_h^p - \widehat{\varepsilon}_h^p)$$
$$-\boldsymbol{\delta}_h^{\boldsymbol{q}} \cdot \boldsymbol{n} - 2i\omega(\boldsymbol{b_0} \cdot \boldsymbol{n})\widehat{\delta}_h^p - 2i\omega\tau P_M \delta_h^p + 2i\omega\tau_{\text{upw}} \left(\delta_h^p - \widehat{\delta}_h^p\right) \tag{62}$$

Proof: Notice that $(p, \boldsymbol{q})$ satisfy the equations (46a)–(46b)–(46c), introduce the projections wherever possible and subtract the actual discrete equations. ∎

A useful estimate: We will need to use the following estimate for $\|\nabla\varepsilon_h^p\|_{\partial\mathcal{T}_h}$ to carry out our analysis

Lemma 4.3:
The following estimate holds

$$\|\nabla\varepsilon_h^p\|_{\mathcal{T}_h} \leqslant C_{\boldsymbol{W_0},\mathcal{T}_h} \left(\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} + \|\boldsymbol{\delta}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} \right) + C \left\| |\tau|^{\frac{1}{2}} (P_M \varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\|_{\partial\mathcal{T}_h}$$

Proof: Going back to (61a), testing with $\boldsymbol{r}_h = \nabla\varepsilon_h^p$ and integrating by parts leads to

$$(\boldsymbol{W_0}\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}, \nabla\varepsilon_h^p)_{\mathcal{T}_h} + \|\nabla\varepsilon_h^p\|_{\mathcal{T}_h}^2 + \langle \widehat{\varepsilon}_h^p - \varepsilon_h^p, \nabla\varepsilon_h^p \cdot \boldsymbol{n} \rangle_{\partial\mathcal{T}_h} = (\boldsymbol{W_0}\boldsymbol{\delta}_h^{\boldsymbol{q}}, \nabla\varepsilon_h^p)_{\mathcal{T}_h}$$

As $\varepsilon_h^p \in W_h$, $\nabla\varepsilon_h^p \cdot \boldsymbol{n} \in \mathcal{P}_k$ and we can use the following property of the projection P_M :

$$\langle P_M \varepsilon_h^p, \nabla\varepsilon_h^p \cdot \boldsymbol{n} \rangle_{\partial\mathcal{T}_h} = \langle \varepsilon_h^p, \nabla\varepsilon_h^p \cdot \boldsymbol{n} \rangle_{\partial\mathcal{T}_h}$$

Using the Cauchy-Schwartz inequality we get

$$\left| \langle \widehat{\varepsilon}_h^p - \varepsilon_h^p, \nabla\varepsilon_h^p \cdot \boldsymbol{n} \rangle_{\partial\mathcal{T}_h} \right| = \left| \langle \widehat{\varepsilon}_h^p - P_M \varepsilon_h^p, \nabla\varepsilon_h^p \cdot \boldsymbol{n} \rangle_{\partial\mathcal{T}_h} \right|$$
$$\leqslant C \|P_M \varepsilon_h^p - \widehat{\varepsilon}_h^p\|_{\partial\mathcal{T}_h} \|\nabla\varepsilon_h^p\|_{\partial\mathcal{T}_h}$$

for some constant $C > 0$.
Using the following trace inequality (60)

$$\forall w \in W_h, \quad \|w\|_{\partial K} \leqslant C h_K^{-\frac{1}{2}} \|w\|_K$$

we have

$$\|\nabla\varepsilon_h^p\|_{\mathcal{T}_h}^2 \leqslant C_{\boldsymbol{W_0},K} \|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},K} \|\nabla\varepsilon_h^p\|_{\mathcal{T}_h} + \|P_M \varepsilon_h^p - \widehat{\varepsilon}_h^p\|_{\partial\mathcal{T}_h} C h_K^{-\frac{1}{2}} \|\nabla\varepsilon_h^p\|_{\mathcal{T}_h} + C_{\boldsymbol{W_0},K} \|\boldsymbol{\delta}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},K} \|\nabla\varepsilon_h^p\|_{\mathcal{T}_h}$$

which is the desired estimate as $\tau|_K = \mathcal{O}(h_K^{-1})$. ∎

Using the Poincaré-Wirtinger inequality: We denote by $\{u\}$ the L^2-projection of u on $\mathcal{P}_0$, *ie*

$$\forall K \in \mathcal{T}_h, \quad \{u\}\,|_K = \frac{1}{|K|}\int_K u \mathrm{d}\boldsymbol{x}$$

We will need to subtract $\{u\}$ from the equations to apply the Poincaré-Wirtinger inequality :

$$\|u - \{u\}\|_{\mathcal{T}_h} \leqslant Ch\,\|\nabla u\|_{\mathcal{T}_h}\,. \tag{63}$$

We can do that thanks to the following property of the projections π_W and $\boldsymbol{\pi_V}$, indeed we have for $\boldsymbol{\xi}$ and $\boldsymbol{q}$

$$(\boldsymbol{\pi_V q}, \{\boldsymbol{W_0 \xi}\})_{\mathcal{T}_h} = (\boldsymbol{q}, \{\boldsymbol{W_0 \xi}\})_{\mathcal{T}_h} \quad \text{as} \quad \{\boldsymbol{W_0 \xi}\} \in \boldsymbol{\mathcal{P}}_0 \subset \boldsymbol{\mathcal{P}}_k$$

therefore

$$(\boldsymbol{\delta}_h^{\boldsymbol{q}}, \boldsymbol{W_0 \xi})_{\mathcal{T}_h} = (\boldsymbol{q} - \boldsymbol{\pi_V q}, \boldsymbol{W_0 \xi})_{\mathcal{T}_h} = (\boldsymbol{q} - \boldsymbol{\pi_V q}, \boldsymbol{W_0 \xi} - \{\boldsymbol{W_0 \xi}\})_{\mathcal{T}_h}\,. \tag{64}$$

Similar results can be obtained in the same way for the other quantities.

Best approximation property of P_M: During the analysis, we will often need to compare quantities like $\|u - P_M u\|_{\partial K}$ and $\|u - \{u\}\|_{\partial K}$.

Lemma 4.4:
For $u \in \mathcal{P}_{k+1}(\mathcal{T}_h)$, the following inequality holds

$$\|u - P_M u\|_{\partial K} \lesssim \|u - \{u\}\|_{\partial K}$$

Proof:
We recall that $M_h := \prod_{e \in \mathcal{E}_h} \mathcal{P}_k(e)$ is a finite-dimensional vector subspace of $L^2(\partial\mathcal{T}_h)$. We recalled that functions in M_h are bi-valued piecewise polynomials of degree up to k on the skeleton of the mesh.
On an internal edge $e = \partial K_- \cap \partial K_+$, we define

$$\{u\}_e := \begin{cases} \{u^-\} \text{ on } K_- \\ \{u^+\} \text{ on } K_+ \end{cases}, \quad \text{where } \; u^\pm = u|_{K^\pm}.$$

With this definition $\{u\}_e$ is a bi-valued piecewise constant on the skeleton of the mesh, and therefore $\{u\}_e \in M_h$.
As P_M is the orthogonal projection onto M_h, we can use the Hilbert projection theorem to obtain

$$\|u - P_M u\|_{\partial\mathcal{T}_h} \leqslant \inf_{v \in M_h} \|u - v\|_{\partial\mathcal{T}_h}\,.$$

We can therefore conclude that

$$\|u - P_M u\|_{\partial\mathcal{T}_h} \leqslant \|u - \{u\}_e\|_{\partial\mathcal{T}_h}\,.$$

When no confusions are possible, we will denote $\{u\}_e$ by $\{u\}$. ∎
This property will often be referred to as the *best approximation property* of P_M.

Discrete energy-like equality: We will now establish a discrete energy-like equality which will be one of the key ingredients to study the convergence of our method.

Lemma 4.5:
The following discrete energy-like equality holds

$$\begin{aligned}\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|^2_{\boldsymbol{W_0},\mathcal{T}_h} - \omega^2\|\varepsilon_h^p\|^2_{\rho_0,\mathcal{T}_h} - 2i\omega\left\||\tau|^{\frac12}(P_M\varepsilon_h^p-\widehat{\varepsilon}_h^p)\right\|^2_{\partial\mathcal{T}_h} - 2i\omega\left\|\left(\frac12|\boldsymbol{b_0}\cdot\boldsymbol{n}|\right)^{\frac12}(\varepsilon_h^p-\widehat{\varepsilon}_h^p)\right\|^2_{\partial\mathcal{T}_h}\\ -2\omega\mathfrak{Im}\left(\varepsilon_h^p\boldsymbol{b_0},\nabla\varepsilon_h^p\right)_K\\ = -\omega^2\left(\rho_0\delta_h^p,\varepsilon_h^p\right)_K + 2i\omega\left(\delta_h^p\boldsymbol{b_0},\nabla\varepsilon_h^p\right)_K + \left(\boldsymbol{\varepsilon}_h^{\boldsymbol{q}},\boldsymbol{W_0}\boldsymbol{\delta}_h^{\boldsymbol{q}}\right)_K\\ +\left\langle\boldsymbol{\delta}_h^{\boldsymbol{q}}\cdot\boldsymbol{n}+2i\omega(\boldsymbol{b_0}\cdot\boldsymbol{n})\widehat{\delta}_h^p+2i\omega\tau P_M\delta_h^p-2i\omega\tau_{\text{upw}}\left(\delta_h^p-\widehat{\delta}_h^p\right),\varepsilon_h^p-\widehat{\varepsilon}_h^p\right\rangle_{\partial\mathcal{T}_h}\end{aligned}\tag{65}$$

Furthermore if $p\in H^s(\mathcal{O})$ and $\boldsymbol{q}\in\boldsymbol{H}^{\boldsymbol{t}}(\mathcal{O})$ where $s\in[1,k+2]$ and $t\in[1,k+1]$ then the following estimate holds

$$\begin{aligned}&\left|\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|^2_{\boldsymbol{W_0},\mathcal{T}_h} - 2i\omega\left(\left\||\tau|^{\frac12}(P_M\varepsilon_h^p-\widehat{\varepsilon}_h^p)\right\|^2_{\partial\mathcal{T}_h}+\left\|\left(\frac12|\boldsymbol{b_0}\cdot\boldsymbol{n}|\right)^{\frac12}(\varepsilon_h^p-\widehat{\varepsilon}_h^p)\right\|^2_{\partial\mathcal{T}_h}\right)\right|\\ \lesssim&\omega^2\|\varepsilon_h^p\|^2_{\mathcal{T}_h}+\omega\|\varepsilon_h^p\|_{\mathcal{T}_h}\left(\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h}+h^t\|\boldsymbol{q}\|_{t,\mathcal{O}}+\left\||\tau|^{\frac12}(P_M\varepsilon_h^p-\widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h}+\omega h^{s-1}\|p\|_{s,\mathcal{O}}\right)\\ &+\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h}\left(h^t\|\boldsymbol{q}\|_{t,\mathcal{O}}+\omega h^s\|p\|_{s,\mathcal{O}}\right)+h^{2t}\|\boldsymbol{q}\|^2_{t,\mathcal{O}}+\omega h^{s-1}\|p\|_{s,\mathcal{O}}h^t\|\boldsymbol{q}\|_{t,\mathcal{O}}\\ &+\left\||\tau|^{\frac12}(P_M\varepsilon_h^p-\widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h}\left(\omega h^{s-1}\|p\|_{s,\mathcal{O}}+h^t\|\boldsymbol{q}\|_{t,\mathcal{O}}\right)\end{aligned}\tag{66}$$

where

$$h:=\max_{K\in\mathcal{T}_h}h_K.$$

Proof: Test (61a)–(61b)–(61c) with $(\boldsymbol{\varepsilon}_h^{\boldsymbol{q}},\varepsilon_h^p,\widehat{\varepsilon}_h^p)$ and sum the resulting equations to obtain

$$\begin{aligned}\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|^2_{\boldsymbol{W_0},K}-\omega^2\|\varepsilon_h^p\|^2_{\rho_0,K}+2i\omega\left(\varepsilon_h^p\boldsymbol{b_0},\nabla\varepsilon_h^p\right)_K+\left\langle\widehat{\mathcal{Q}}\cdot\boldsymbol{n}-\widehat{\mathcal{Q}_h}\cdot\boldsymbol{n}-\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\cdot\boldsymbol{n},\varepsilon_h^p-\widehat{\varepsilon}_h^p\right\rangle_{\partial K}=\\ -\omega^2\left(\rho_0\delta_h^p,\varepsilon_h^p\right)_K+2i\omega\left(\delta_h^p\boldsymbol{b_0},\nabla\varepsilon_h^p\right)_K+\left(\boldsymbol{\varepsilon}_h^{\boldsymbol{q}},\boldsymbol{W_0}\boldsymbol{\delta}_h^{\boldsymbol{q}}\right)_K\end{aligned}$$

We will now compute the boundary terms using (62).
Boundary terms involving P_M:
Notice that, as $P_M\varepsilon_h^p-\widehat{\varepsilon}_h^p\in M_h(K)$ and $\tau\in\mathcal{R}_0$, we have

$$\langle\tau(P_M\varepsilon_h^p-\widehat{\varepsilon}_h^p),P_M\varepsilon_h^p\rangle_{\partial K}=\langle\tau(P_M\varepsilon_h^p-\widehat{\varepsilon}_h^p),\varepsilon_h^p\rangle_{\partial K}$$

and therefore

$$\langle\tau(P_M\varepsilon_h^p-\widehat{\varepsilon}_h^p),\varepsilon_h^p-\widehat{\varepsilon}_h^p\rangle_{\partial K}=\langle\tau(P_M\varepsilon_h^p-\widehat{\varepsilon}_h^p),P_M\varepsilon_h^p-\widehat{\varepsilon}_h^p\rangle_{\partial K}=\left\|\tau^{\frac12}(P_M\varepsilon_h^p-\widehat{\varepsilon}_h^p)\right\|^2_{\partial K}$$

.
We also have the following estimate

$$\begin{aligned}2\omega\left|\langle\tau P_M\delta_h^p,\varepsilon_h^p-\widehat{\varepsilon}_h^p\rangle_{\partial\mathcal{T}_h}\right|&\leqslant 2\omega\left|\langle|\tau|P_M\delta_h^p,P_M\varepsilon_h^p-\widehat{\varepsilon}_h^p\rangle_{\partial\mathcal{T}_h}\right|\\ &\leqslant 2\omega\left|\left\langle|\tau|^{\frac12}\delta_h^p,|\tau|^{\frac12}(P_M\varepsilon_h^p-\widehat{\varepsilon}_h^p)\right\rangle_{\partial\mathcal{T}_h}\right|\\ &\lesssim\omega\left\||\tau|^{\frac12}(P_M\varepsilon_h^p-\widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h}\||\tau|^{\frac12}\|\delta_h^p\|_{\partial\mathcal{T}_h}\\ (\tau=\mathcal{O}(h^{-1})\ \text{ and by (59d))}\quad&\lesssim\omega h^{s-1}\left\||\tau|^{\frac12}(P_M\varepsilon_h^p-\widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h}\|p\|_{s,\mathcal{O}}\end{aligned}$$

Boundary terms involving convection:
As in *Step 2* of the proof of THEOREM 4, we will separate the volumetric term involving $\boldsymbol{b_0}$ into its real and imaginary parts. By the LEMMA 4.1 we have

$$\mathfrak{Re}\left(\varepsilon_h^p \boldsymbol{b_0}, \nabla \varepsilon_h^p\right)_K = \frac{1}{2}\left\langle (\boldsymbol{b_0}\cdot\boldsymbol{n})\varepsilon_h^p, \varepsilon_h^p\right\rangle_{\partial K}$$

and we can now obtain the second boundary norm

$$\begin{aligned}
&\langle (\boldsymbol{b_0}\cdot\boldsymbol{n})\widehat{\varepsilon}_h^p, \varepsilon_h^p - \widehat{\varepsilon}_h^p\rangle_{\partial\mathcal{T}_h} - \frac{1}{2}\left\langle (\boldsymbol{b_0}\cdot\boldsymbol{n})\varepsilon_h^p, \varepsilon_h^p\right\rangle_{\partial\mathcal{T}_h} + \langle \tau_{\mathrm{upw}}(\varepsilon_h^p - \widehat{\varepsilon}_h^p), \varepsilon_h^p - \widehat{\varepsilon}_h^p\rangle_{\partial\mathcal{T}_h}\\
&= \left\langle -\frac{1}{2}(\boldsymbol{b_0}\cdot\boldsymbol{n})(\varepsilon_h^p - \widehat{\varepsilon}_h^p), \varepsilon_h^p - \widehat{\varepsilon}_h^p\right\rangle_{\partial\mathcal{T}_h} - \frac{1}{2}\left\langle (\boldsymbol{b_0}\cdot\boldsymbol{n})\widehat{\varepsilon}_h^p, \widehat{\varepsilon}_h^p\right\rangle_{\partial\mathcal{T}_h} + \langle \tau_{\mathrm{upw}}(\varepsilon_h^p - \widehat{\varepsilon}_h^p), \varepsilon_h^p - \widehat{\varepsilon}_h^p\rangle_{\partial\mathcal{T}_h}\\
&= \left\langle \left(\tau_{\mathrm{upw}} - \frac{1}{2}(\boldsymbol{b_0}\cdot\boldsymbol{n})\right)(\varepsilon_h^p - \widehat{\varepsilon}_h^p), \varepsilon_h^p - \widehat{\varepsilon}_h^p\right\rangle_{\partial\mathcal{T}_h}\\
&= \left\| \left(\frac{1}{2}|\boldsymbol{b_0}\cdot\boldsymbol{n}|\right)^{\frac{1}{2}}(\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h}^2
\end{aligned}$$

indeed as $\widehat{\varepsilon}_h^p$ is single-valued across the skeleton of the mesh we have

$$\langle (\boldsymbol{b_0}\cdot\boldsymbol{n})\widehat{\varepsilon}_h^p, \widehat{\varepsilon}_h^p\rangle_{\partial\mathcal{T}_h} = \langle [\![\widehat{\varepsilon}_h^p \boldsymbol{b_0}]\!], \widehat{\varepsilon}_h^p\rangle_{\partial\mathcal{T}_h} = 0$$

and we also use that

$$\tau_{\mathrm{upw}} - \frac{1}{2}(\boldsymbol{b_0}\cdot\boldsymbol{n}) = \frac{1}{2}|\boldsymbol{b_0}\cdot\boldsymbol{n}| \geqslant 0 \tag{67}$$

to use the square root.
We will now eliminate the terms involving τ_{upw} from the right-hand side.

$$\begin{aligned}
\left\| \left(\frac{1}{2}|\boldsymbol{b_0}\cdot\boldsymbol{n}|\right)^{\frac{1}{2}}(\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h} &\lesssim \|\varepsilon_h^p - \widehat{\varepsilon}_h^p\|_{\partial\mathcal{T}_h}\\
&\lesssim \|\varepsilon_h^p - P_M\varepsilon_h^p\|_{\partial\mathcal{T}_h} + \|P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p\|_{\partial\mathcal{T}_h}\\
\text{(by LEMMA 4.4 and } \tau = \mathcal{O}(h^{-1})) \quad &\lesssim \|\varepsilon_h^p - \{\varepsilon_h^p\}\|_{\partial\mathcal{T}_h} + h^{\frac{1}{2}}\left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h}\\
\text{(by (60))} \quad &\lesssim h^{-\frac{1}{2}}\|\varepsilon_h^p - \{\varepsilon_h^p\}\|_{\mathcal{T}_h} + h^{\frac{1}{2}}\left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h}\\
\text{(by (63))} \quad &\lesssim h^{\frac{1}{2}}\|\nabla\varepsilon_h^p\|_{\mathcal{T}_h} + h^{\frac{1}{2}}\left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h}\\
\text{(by LEMMA 4.3)} \quad &\lesssim h^{\frac{1}{2}}\left(\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} + \|\boldsymbol{\delta}_h^{\boldsymbol{q}}\|_{\mathcal{T}_h} + \left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h}\right)\\
\text{(by (59e))} \quad &\lesssim h^{\frac{1}{2}}\left(\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} + h^t\|\boldsymbol{q}\|_{t,\mathcal{O}} + \left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h}\right)
\end{aligned} \tag{68}$$

Using the following inequalities that can be derived from (67)

$$\begin{aligned}
\tau_{\mathrm{upw}} &\lesssim \frac{1}{2}|\boldsymbol{b_0}\cdot\boldsymbol{n}|\\
\boldsymbol{b_0}\cdot\boldsymbol{n} &\lesssim \frac{1}{2}|\boldsymbol{b_0}\cdot\boldsymbol{n}|
\end{aligned}$$

and using (59d) with $s-1$ instead of s to keep $s \leqslant k+2$, we deduce that

$$\begin{aligned}
2\omega \left\langle (\boldsymbol{b_0}\cdot\boldsymbol{n})\widehat{\delta}_h^p, \varepsilon_h^p - \widehat{\varepsilon}_h^p \right\rangle_{\partial\mathcal{T}_h} &= 2\omega \left\langle (\boldsymbol{b_0}\cdot\boldsymbol{n})^{\frac{1}{2}}\delta_h^p, (\boldsymbol{b_0}\cdot\boldsymbol{n})^{\frac{1}{2}}(\varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\rangle_{\partial\mathcal{T}_h} \\
&\lesssim 2\omega \left\langle (\boldsymbol{b_0}\cdot\boldsymbol{n})^{\frac{1}{2}}\delta_h^p, (\frac{1}{2}|\boldsymbol{b_0}\cdot\boldsymbol{n}|)^{\frac{1}{2}}(\varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\rangle_{\partial\mathcal{T}_h} \\
&\lesssim \omega \left\|\widehat{\delta}_h^p\right\|_{\partial\mathcal{T}_h} \left\| \left(\frac{1}{2}|\boldsymbol{b_0}\cdot\boldsymbol{n}|\right)^{\frac{1}{2}} (\varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\|_{\partial\mathcal{T}_h} \\
&\lesssim \omega h^{s-\frac{3}{2}} \left\| \left(\frac{1}{2}|\boldsymbol{b_0}\cdot\boldsymbol{n}|\right)^{\frac{1}{2}} (\varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\|_{\partial\mathcal{T}_h} \|p\|_{s,\mathcal{O}} \\
&\lesssim \omega h^{s-1} \left(\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} + h^t \|\boldsymbol{q}\|_{t,\mathcal{O}} + \left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h} \right) \|p\|_{s,\mathcal{O}}
\end{aligned}$$

and

$$\begin{aligned}
2\omega \left\langle \tau_{\mathrm{upw}}(\delta_h^p - \widehat{\delta}_h^p), \varepsilon_h^p - \widehat{\varepsilon}_h^p \right\rangle_{\partial\mathcal{T}_h} &\lesssim \omega \left\|\delta_h^p - \widehat{\delta}_h^p\right\|_{\partial\mathcal{T}_h} \left\| \left(\tau_{\mathrm{upw}} - \frac{1}{2}\boldsymbol{b_0}\cdot\boldsymbol{n}\right)^{\frac{1}{2}} (\varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\|_{\partial\mathcal{T}_h} \\
&\lesssim \omega h^{s-\frac{3}{2}} \left\| \left(\frac{1}{2}|\boldsymbol{b_0}\cdot\boldsymbol{n}|\right)^{\frac{1}{2}} (\varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\|_{\partial\mathcal{T}_h} \|p\|_{s,\mathcal{O}} \\
&\lesssim \omega h^{s-1} \left(\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} + h^t \|\boldsymbol{q}\|_{t,\mathcal{O}} + \left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h} \right) \|p\|_{s,\mathcal{O}}
\end{aligned}$$

Boundary term involving $\boldsymbol{\delta}_h^{\boldsymbol{q}}$:

$$\begin{aligned}
\langle \boldsymbol{\delta}_h^{\boldsymbol{q}}\cdot\boldsymbol{n}, \varepsilon_h^p - \widehat{\varepsilon}_h^p\rangle_{\partial\mathcal{T}_h} &= \langle \boldsymbol{\delta}_h^{\boldsymbol{q}}\cdot\boldsymbol{n}, \varepsilon_h^p - P_M\varepsilon_h^p + P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p\rangle_{\partial\mathcal{T}_h} \\
&= \underbrace{\langle \boldsymbol{\delta}_h^{\boldsymbol{q}}\cdot\boldsymbol{n}, \varepsilon_h^p - P_M\varepsilon_h^p\rangle_{\partial\mathcal{T}_h}}_{=:T_1} + \underbrace{\langle \boldsymbol{\delta}_h^{\boldsymbol{q}}\cdot\boldsymbol{n}, P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p\rangle_{\partial\mathcal{T}_h}}_{=:T_2}
\end{aligned}$$

Using a weighted Cauchy-Schwartz inequality and recalling that $\tau = \mathcal{O}(h^{-1})$ we have

$$\begin{aligned}
T_2 &\lesssim \left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h} h^{\frac{1}{2}} \|\boldsymbol{\delta}_h^{\boldsymbol{q}}\|_{\partial\mathcal{T}_h} \\
\text{(by (59e))} \quad &\lesssim h^t \left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h} \|\boldsymbol{q}\|_{t,\mathcal{O}}
\end{aligned}$$

$$\begin{aligned}
T_1 &= \langle \boldsymbol{\delta}_h^{\boldsymbol{q}}\cdot\boldsymbol{n}, \varepsilon_h^p - P_M\varepsilon_h^p\rangle_{\partial\mathcal{T}_h} \\
&\lesssim \|\boldsymbol{\delta}_h^{\boldsymbol{q}}\|_{\partial\mathcal{T}_h} \|\varepsilon_h^p - P_M\varepsilon_h^p\|_{\partial\mathcal{T}_h} \\
\text{(by Lemma 4.4)} \quad &\lesssim \|\boldsymbol{\delta}_h^{\boldsymbol{q}}\|_{\partial\mathcal{T}_h} \|\varepsilon_h^p - \{\varepsilon_h^p\}\|_{\partial\mathcal{T}_h} \\
\text{(by (59e) and (59d))} \quad &\lesssim h^{t-\frac{1}{2}} \|\boldsymbol{q}\|_{t,\mathcal{O}}\, h^{-\frac{1}{2}} \|\varepsilon_h^p - \{\varepsilon_h^p\}\|_{\mathcal{T}_h} \\
\text{(by (63))} \quad &\lesssim h^t \|\boldsymbol{q}\|_{t,\mathcal{O}} \|\nabla\varepsilon_h^p\|_{\mathcal{T}_h} \\
\text{(by Lemma 4.3)} \quad &\lesssim h^t \|\boldsymbol{q}\|_{t,\mathcal{O}} \left(\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} + \|\boldsymbol{\delta}_h^{\boldsymbol{q}}\|_{\mathcal{T}_h} + \left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h} \right) \\
\text{by (59b))} \quad &\lesssim h^t \|\boldsymbol{q}\|_{t,\mathcal{O}} \left(\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} + \left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h} \right) + h^{2t} \|\boldsymbol{q}\|_{t,\mathcal{O}}^2
\end{aligned}$$

Therefore

$$\langle \boldsymbol{\delta}_h^{\boldsymbol{q}}\cdot\boldsymbol{n}, \varepsilon_h^p - \widehat{\varepsilon}_h^p\rangle_{\partial\mathcal{T}_h} \lesssim h^t \|\boldsymbol{q}\|_{t,\mathcal{O}} \left(\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} + \left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h} \right) + h^{2t} \|\boldsymbol{q}\|_{t,\mathcal{O}}^2$$

Volumetric terms: By similar computations using the Cauchy-Schwartz inequality, the projection estimates in (59) and LEMMA 4.3 we can show that

$$\begin{aligned}
\omega^2 \left(\rho_0 \delta_h^p, \varepsilon_h^p\right)_{\mathcal{T}_h} &\lesssim \omega^2 h^s \left\|p\right\|_{s,\mathcal{O}} \left\|\varepsilon_h^p\right\|_{\mathcal{T}_h} \\
2\omega \left(\delta_h^p \boldsymbol{b_0}, \nabla \varepsilon_h^p\right)_{\mathcal{T}_h} &\lesssim \omega h^s \left\|p\right\|_{s,\mathcal{O}} \left(\left\|\boldsymbol{\varepsilon_h^q}\right\|_{\boldsymbol{W_0},\mathcal{T}_h} + h^t \left\|\boldsymbol{q}\right\|_{t,\mathcal{O}} + \left\||\tau|^{\frac{1}{2}}(P_M \varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h}\right) \\
\left(\boldsymbol{W_0} \boldsymbol{\varepsilon_h^q}, \boldsymbol{\delta_h^q}\right)_{\mathcal{T}_h} &\lesssim h^t \left\|\boldsymbol{\varepsilon_h^q}\right\|_{\boldsymbol{W_0},\mathcal{T}_h} \left\|\boldsymbol{q}\right\|_{t,\mathcal{O}} \\
2\omega \left(\varepsilon_h^p \boldsymbol{b_0}, \nabla \varepsilon_h^p\right)_{\mathcal{T}_h} &\lesssim \left\|\varepsilon_h^p\right\|_{\mathcal{T}_h} \left(\left\|\boldsymbol{\varepsilon_h^q}\right\|_{\boldsymbol{W_0},\mathcal{T}_h} + h^t \left\|\boldsymbol{q}\right\|_{t,\mathcal{O}} + \left\||\tau|^{\frac{1}{2}}(P_M \varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h}\right)
\end{aligned}$$

■

We can rewrite estimate (66) in a more readable form :

Corollary 4.1:
The following estimate holds

$$\begin{aligned}
&\left| \left\|\boldsymbol{\varepsilon_h^q}\right\|^2_{\boldsymbol{W_0},\mathcal{T}_h} - 2i\omega \left(\left\||\tau|^{\frac{1}{2}}(P_M \varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|^2_{\partial\mathcal{T}_h} + \left\| \left(\frac{1}{2}|\boldsymbol{b_0}\cdot\boldsymbol{n}|\right)^{\frac{1}{2}} (\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|^2_{\partial\mathcal{T}_h} \right) \right| \\
\lesssim &\varepsilon \left(\left\|\boldsymbol{\varepsilon_h^q}\right\|^2_{\boldsymbol{W_0},\mathcal{T}_h} + \left\||\tau|^{\frac{1}{2}} (P_M \varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|^2_{\partial\mathcal{T}_h} \right) + \frac{1}{\varepsilon} \left(\omega^2 \left\|\varepsilon_h^p\right\|^2_{\mathcal{T}_h} + h^{2t} \left\|\boldsymbol{q}\right\|^2_{t,\mathcal{O}} + \omega^2 h^{2s-2} \left\|p\right\|^2_{s,\mathcal{O}} \right)
\end{aligned}$$

Proof: apply the weighted Young's inequality to the right-hand side of (66). The value of ε will be discussed later. ■

4.3.2 Adjoint problem

As the identity (65) does not allow us to directly obtain any error estimate, we need to use a duality argument. For an introduction to the *Aubin-Nitsche method* we refer to [EG04, Sec. 2.3.4], similar processes for HDG(+) methods in the context of wave equations have been carried out in [QSS16], [Hun19, Sec. 3.5] and [DS19, Sec. 3.5.2] and for coercive problems with convection in [QS16a].
The adjoint problem is

$$\begin{aligned}
\boldsymbol{W_0}\boldsymbol{\xi} - \nabla\theta &= 0 && \text{in } \mathcal{O} \\
-\rho_0\omega^2\theta - 2i\omega\boldsymbol{b_0}\cdot\nabla\theta - \operatorname{div}(\boldsymbol{\xi}) &= \varepsilon_h^p && \text{in } \mathcal{O} \\
\theta &= 0 && \text{on } \Gamma_D \\
\boldsymbol{\xi}\cdot\boldsymbol{n} - 2i\omega(\boldsymbol{b_0}\cdot\boldsymbol{n})\theta &= 0 && \text{on } \Gamma_N
\end{aligned}$$

and $(\boldsymbol{\xi}, \theta) \in \boldsymbol{H^1}(\mathcal{O}) \times H^2(\mathcal{O})$ satisfy the following discrete problem for all $(\boldsymbol{r}_h, w_h, \mu_h) \in \boldsymbol{V}_h \times W_h \times M_h$

$$\left(\boldsymbol{W_0}\boldsymbol{\xi}, \boldsymbol{r}_h\right)_{\mathcal{T}_h} + \left(\theta, \operatorname{div}(\boldsymbol{r}_h)\right)_{\mathcal{T}_h} + \left\langle\theta, \boldsymbol{r}_h\cdot\boldsymbol{n}\right\rangle_{\partial\mathcal{T}_h} = 0 \tag{70a}$$

$$-\omega^2\left(\rho_0\theta, w_h\right)_{\mathcal{T}_h} + \left(2i\omega\theta\boldsymbol{b_0} + \boldsymbol{\xi}, \nabla w_h\right)_{\mathcal{T}_h} + \left\langle\boldsymbol{\xi}\cdot\boldsymbol{n} - 2i\omega(\boldsymbol{b_0}\cdot\boldsymbol{n})\theta, w_h\right\rangle_{\partial\mathcal{T}_h} = \left(\varepsilon_h^p, w_h\right)_{\mathcal{T}_h} \tag{70b}$$

$$\left\langle\boldsymbol{\xi}\cdot\boldsymbol{n}, \mu_h\right\rangle_{\partial\mathcal{T}_h\setminus\Gamma_D} = 0 \tag{70c}$$

The last equation (70c) translates the continuity of $\boldsymbol{\xi}\cdot\boldsymbol{n}$ between the elements and should be interpreted as a jump term. Indeed by the same argument as when we discussed weak continuity of $\boldsymbol{q}_h\cdot\boldsymbol{n}$ in SUBSECTION 4.1, we can show that

$$\sum_{K\in\mathcal{T}_h}\int_{\partial K}\boldsymbol{\xi}\cdot\boldsymbol{n}\mu_h^*\mathrm{d}\sigma = \sum_{e\in\mathcal{E}_h^i}\int_e[\![\boldsymbol{\xi}]\!]\mu_h^*\mathrm{d}\sigma - \int_{\Gamma_D}\boldsymbol{\xi}\cdot\boldsymbol{n}\mu_h^*\mathrm{d}\sigma. \tag{71}$$

Remark 4.6: The functional framework for (70c) is a bit complicated. Indeed the interior integrals should formally be interpreted as duality brackets between $H^{-\frac{1}{2}}$ and $H^{\frac{1}{2}}$ and the restriction of those distributions to a segment is not defined. Notice however that the right-hand side of (71) is well-defined, therefore giving meaning to the left-hand side. Moreover, as we assume additional regularity $(\boldsymbol{\xi},\theta)\in \boldsymbol{H^1}(\mathcal{O})\times H^2(\mathcal{O})$ for the solution of the adjoint problem and as we will work with polynomial quantities at the discrete level, this is not problematic.

In our analysis we will need to use the following elliptic regularity estimate for the dual problem

$$\|\theta\|_{2,\mathcal{O}}+\|\boldsymbol{\xi}\|_{1,\mathcal{O}}\leqslant C_{\mathrm{reg}}\,\|\varepsilon_h^p\|_{\mathcal{O}}\,. \tag{72}$$

This estimate holds when $\boldsymbol{\xi}$ and θ are regular enough, which amounts to requiring enough regularity on the background quantities ρ_0, c_0 and $\boldsymbol{b_0}$, and the convexity of the domain $\mathcal{O}$.

Lemma 4.6:
We have the following dual identity :

$$\begin{aligned}\|\varepsilon_h^p\|_{\mathcal{T}_h}^2 =&-\left(\boldsymbol{W_0}\boldsymbol{\varepsilon}_h^{\boldsymbol{q}},\boldsymbol{\pi_V}\boldsymbol{\xi}-\boldsymbol{\xi}\right)_{\mathcal{T}_h}+\omega^2\left(\rho_0\varepsilon_h^p,\pi_W\theta-\theta\right)_{\mathcal{T}_h}+2i\omega\left(\nabla\varepsilon_h^p,(\pi_W\theta-\theta)\boldsymbol{b_0}\right)_{\mathcal{T}_h}\\
&-2i\omega\left\langle(\boldsymbol{b_0}\cdot\boldsymbol{n})\varepsilon_h^p,\pi_W\theta-\theta\right\rangle_{\partial\mathcal{T}_h}\\
&+\left(\boldsymbol{W_0}\boldsymbol{\delta}_h^{\boldsymbol{q}},\boldsymbol{\pi_V}\boldsymbol{\xi}\right)_{\mathcal{T}_h}-\omega^2\left(\rho_0\delta_h^p,\pi_W\theta\right)_{\mathcal{T}_h}+2i\omega\left(\delta_h^p\boldsymbol{b_0},\nabla(\pi_W\theta)\right)_{\mathcal{T}_h}\\
&+2i\omega\left\langle(\boldsymbol{b_0}\cdot\boldsymbol{n})\widehat{\varepsilon}_h^p-\tau\left(P_M\varepsilon_h^p-\widehat{\varepsilon}_h^p\right)+\tau_{\mathrm{upw}}\left(\varepsilon_h^p-\widehat{\varepsilon}_h^p\right),\pi_W\theta-P_M\theta\right\rangle_{\partial\mathcal{T}_h}\\
&-2i\omega\left\langle(\boldsymbol{b_0}\cdot\boldsymbol{n})\widehat{\delta}_h^p-\tau P_M\delta_h^p+\tau_{\mathrm{upw}}\left(\delta_h^p-\widehat{\delta}_h^p\right),\pi_W\theta-P_M\theta\right\rangle_{\partial\mathcal{T}_h}+\left\langle\boldsymbol{\delta}_h^{\boldsymbol{q}}\cdot\boldsymbol{n},\pi_W\theta-P_M\theta\right\rangle_{\partial\mathcal{T}_h}\end{aligned}$$

Proof: Introducing the projections in (70a)–(70b)–(70c) and testing with $(\boldsymbol{\varepsilon}_h^{\boldsymbol{q}},\varepsilon_h^p,\widehat{\varepsilon}_h^p)$

$$\left(\boldsymbol{W_0}\boldsymbol{\pi_V}\boldsymbol{\xi},\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\right)_{\mathcal{T}_h}+\left(\pi_W\theta,\mathrm{div}\left(\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\right)\right)_{\mathcal{T}_h}-\left\langle P_M\theta,\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\cdot\boldsymbol{n}\right\rangle_{\partial\mathcal{T}_h}=\left(\boldsymbol{W_0}\left(\boldsymbol{\pi_V}\boldsymbol{\xi}-\boldsymbol{\xi}\right),\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\right)_{\mathcal{T}_h} \tag{73a}$$

$$\begin{aligned}-\omega^2\left(\rho_0\pi_W\theta,\varepsilon_h^p\right)_{\mathcal{T}_h}+2i\omega\left((\pi_W\theta)\boldsymbol{b_0},\nabla\varepsilon_h^p\right)_{\mathcal{T}_h}-\left(\mathrm{div}\left(\boldsymbol{\pi_V}\boldsymbol{\xi}\right),\varepsilon_h^p\right)_{\mathcal{T}_h}-2i\omega\left\langle(\boldsymbol{b_0}\cdot\boldsymbol{n})\pi_W\theta,\varepsilon_h^p\right\rangle_{\partial\mathcal{T}_h}=\\
\left(\varepsilon_h^p,\varepsilon_h^p\right)_{\mathcal{T}_h}-\omega^2\left(\rho_0(\pi_W\theta-\theta),\varepsilon_h^p\right)_{\mathcal{T}_h}+2i\omega\left((\pi_W\theta-\theta)\boldsymbol{b_0},\nabla\varepsilon_h^p\right)_{\mathcal{T}_h}-2i\omega\left\langle(\boldsymbol{b_0}\cdot\boldsymbol{n})(\pi_W\theta-\theta),\varepsilon_h^p\right\rangle_{\partial\mathcal{T}_h}\end{aligned} \tag{73b}$$

$$\left\langle\boldsymbol{\pi_V}\boldsymbol{\xi}\cdot\boldsymbol{n},\widehat{\varepsilon}_h^p\right\rangle_{\partial\mathcal{T}_h}=0 \tag{73c}$$

Now conjugate and sum those equations and compare with the sum of (61a)–(61b)–(61c) tested with $(\boldsymbol{\pi_V}\boldsymbol{\xi},\pi_W\theta,P_M\theta)$.

$$\left(\boldsymbol{W_0}\boldsymbol{\varepsilon}_h^{\boldsymbol{q}},\boldsymbol{\pi_V}\boldsymbol{\xi}\right)_{\mathcal{T}_h}-\left(\varepsilon_h^p,\mathrm{div}\left(\boldsymbol{\pi_V}\boldsymbol{\xi}\right)\right)_{\mathcal{T}_h}+\left\langle\widehat{\varepsilon}_h^p,\boldsymbol{\pi_V}\boldsymbol{\xi}\cdot\boldsymbol{n}\right\rangle_{\partial\mathcal{T}_h}=\ell_1(\boldsymbol{\pi_V}\boldsymbol{\xi}) \tag{74a}$$

$$-\omega^2\left(\rho_0\varepsilon_h^p,\pi_W\theta\right)_{\mathcal{T}_h}-2i\omega\left(\nabla\varepsilon_h^p,(\pi_W\theta)\boldsymbol{b_0}\right)_{\mathcal{T}_h} \tag{74b}$$

$$+\left(\mathrm{div}\left(\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\right),\pi_W\theta\right)_{\mathcal{T}_h}+2i\omega\left\langle(\boldsymbol{b_0}\cdot\boldsymbol{n})\varepsilon_h^p,\pi_W\theta\right\rangle_{\partial\mathcal{T}_h}=\ell_2(\pi_W\theta)$$

$$-\left\langle\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\cdot\boldsymbol{n},P_M\theta\right\rangle_{\partial\mathcal{T}_h}=\ell_3(P_M\theta) \tag{74c}$$

where

$$\begin{aligned}
\ell_1(\boldsymbol{\pi_V \xi}) &:= (\boldsymbol{W_0 \delta_h^q}, \boldsymbol{\pi_V \xi})_{\mathcal{T}_h} \\
\ell_2(\pi_W \theta) &:= -\omega^2 \left(\rho_0 \delta_h^p, \pi_W \theta\right)_{\mathcal{T}_h} + 2i\omega \left(\delta_h^p \boldsymbol{b_0}, \nabla \pi_W \theta\right)_{\mathcal{T}_h} \\
&\quad + 2i\omega \left\langle (\boldsymbol{b_0} \cdot \boldsymbol{n}) \widehat{\varepsilon}_h^p - \tau \left(P_M \varepsilon_h^p - \widehat{\varepsilon}_h^p\right) + \tau_{\text{upw}} \left(\varepsilon_h^p - \widehat{\varepsilon}_h^p\right), \pi_W \theta \right\rangle_{\partial \mathcal{T}_h} \\
&\quad + \left\langle \boldsymbol{\delta_h^q} \cdot \boldsymbol{n} + 2i\omega (\boldsymbol{b_0} \cdot \boldsymbol{n}) \widehat{\delta}_h^p + 2i\omega\tau P_M \delta_h^p - 2i\omega\tau_{\text{upw}} (\delta_h^p - \widehat{\delta}_h^p), \pi_W \theta \right\rangle_{\partial \mathcal{T}_h} \\
\ell_3(P_M \theta) &:= -2i\omega \left\langle (\boldsymbol{b_0} \cdot \boldsymbol{n}) \widehat{\varepsilon}_h^p - \tau \left(P_M \varepsilon_h^p - \widehat{\varepsilon}_h^p\right) + \tau_{\text{upw}} \left(\varepsilon_h^p - \widehat{\varepsilon}_h^p\right), P_M \theta \right\rangle_{\partial \mathcal{T}_h} \\
&\quad - \left\langle \boldsymbol{\delta_h^q} \cdot \boldsymbol{n} + 2i\omega (\boldsymbol{b_0} \cdot \boldsymbol{n}) \widehat{\delta}_h^p + 2i\omega\tau P_M \delta_h^p - 2i\omega\tau_{\text{upw}} \left(\delta_h^p - \widehat{\delta}_h^p\right), P_M \theta \right\rangle_{\partial \mathcal{T}_h}.
\end{aligned}$$

Notice that an integration by parts has been carried out in ℓ_2.
As $\boldsymbol{W_0}$ is real and symmetric, we have

$$(\boldsymbol{W_0 \pi_V \xi}, \boldsymbol{\varepsilon_h^q})_{\mathcal{T}_h} = (\boldsymbol{\pi_V \xi}, \boldsymbol{W_0 \varepsilon_h^q})_{\mathcal{T}_h}$$

and we can therefore notice that the left-hand sides of (73a)–(73b)–(73c) and (74a)–(74b)–(74c) (after being conjugated) are the same, leading to

$$\begin{aligned}
&(\boldsymbol{W_0 \varepsilon_h^q}, \boldsymbol{\pi_V \xi} - \boldsymbol{\xi})_{\mathcal{T}_h} + (\varepsilon_h^p, \varepsilon_h^p)_{\mathcal{T}_h} - \omega^2 \left(\rho_0 \varepsilon_h^p, \pi_W \theta - \theta\right)_{\mathcal{T}_h} \\
&-2i\omega \left(\nabla \varepsilon_h^p, (\pi_W \theta - \theta) \boldsymbol{b_0}\right)_{\mathcal{T}_h} + 2i\omega \left\langle (\boldsymbol{b_0} \cdot \boldsymbol{n}) \varepsilon_h^p, \pi_W \theta - \theta \right\rangle_{\partial \mathcal{T}_h} \\
&\qquad = \ell_1(\boldsymbol{\pi_V \xi}) + \ell_2(\pi_W \theta) + \ell_3(P_M \theta)
\end{aligned}$$

And the identity is obtained by a reorganisation of the different terms. ∎

Remark 4.7: conjugating gives the good sign for the convection term, indeed : $[iz]^* = -iz^*$, therefore

$$\left[2i\omega \left((\pi_W \theta) \boldsymbol{b_0}, \nabla \varepsilon_h^p\right)_{\mathcal{T}_h}\right]^* = -2i\omega \left(\nabla \varepsilon_h^p, (\pi_W \theta) \boldsymbol{b_0}\right)_{\mathcal{T}_h}$$

Lemma 4.7:
Assuming that the regularity assumption (72) holds and that $\omega^2 h^2 \left\|\rho_0\right\|_\infty C_{\text{reg}} C$ (where C is the constant of Theorem 4) is small enough, if $p \in H^s(\mathcal{O})$ and $\boldsymbol{q} \in \boldsymbol{H^t}(\mathcal{O})$ where $s \in [1, k+2]$ and $t \in [1, k+1]$ then

$$\left\|\varepsilon_h^p\right\|_{\mathcal{T}_h} \lesssim h^{t+1}(1+\omega) \left\|\boldsymbol{q}\right\|_{t,\mathcal{O}} + h^s(1+\omega+\omega^2) \left\|p\right\|_{s,\mathcal{O}} + \omega h \left\|\tau^{\frac{1}{2}} (P_M \varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial \mathcal{T}_h} + h(1+\omega) \left\|\boldsymbol{\varepsilon_h^q}\right\|_{\boldsymbol{W_0}, \mathcal{T}_h}$$

where

$$h := \max_{K \in \mathcal{T}_h} h_K$$

Proof: we are going to estimate the terms in the right hand side of the Lemma 4.6.
Volumetric terms involving $\boldsymbol{\varepsilon_h^q}$:

$$(\boldsymbol{W_0 \varepsilon_h^q}, \boldsymbol{\xi} - \boldsymbol{\pi_V \xi})_{\mathcal{T}_h} \lesssim h \left\|\boldsymbol{\varepsilon_h^q}\right\|_{\boldsymbol{W_0}, K} \left\|\boldsymbol{\xi}\right\|_{1,\mathcal{O}} \lesssim h \left\|\boldsymbol{\varepsilon_h^q}\right\|_{\boldsymbol{W_0}, \mathcal{T}_h} \left\|\varepsilon_h^p\right\|_{\mathcal{T}_h} \tag{75a}$$

and

$$\begin{aligned}
\left|(\boldsymbol{W_0 \delta_h^q}, \boldsymbol{\pi_V \xi})_{\mathcal{T}_h}\right| &= \left|(\boldsymbol{W_0 \delta_h^q}, \boldsymbol{\xi})_{\mathcal{T}_h} - (\boldsymbol{W_0 \delta_h^q}, \boldsymbol{\xi} - \boldsymbol{\pi_V \xi})_{\mathcal{T}_h}\right| \\
\text{(by (64))} \quad &= \left|(\boldsymbol{\delta_h^q}, \boldsymbol{W_0 \xi} - \{\boldsymbol{W_0 \xi}\})_{\mathcal{T}_h} - (\boldsymbol{W_0 \delta_h^q}, \boldsymbol{\xi} - \boldsymbol{\pi_V \xi})_{\mathcal{T}_h}\right| \\
&\lesssim \left\|\boldsymbol{q} - \boldsymbol{\pi_V q}\right\|_{\mathcal{T}_h} \left\|\boldsymbol{W_0 \xi} - \{\boldsymbol{W_0 \xi}\}\right\|_{\mathcal{T}_h} + \left\|\boldsymbol{q} - \boldsymbol{\pi_V q}\right\|_{\mathcal{T}_h} \left\|\boldsymbol{\xi} - \boldsymbol{\pi_V \xi}\right\|_{\mathcal{T}_h} \\
\text{(by (63) and (59b))} \quad &\lesssim h^t \left\|\boldsymbol{q}\right\|_{t,\mathcal{O}} h |\boldsymbol{\xi}|_{1,\mathcal{O}} + h^t \left\|\boldsymbol{q}\right\|_{t,\mathcal{O}} h \left\|\boldsymbol{\xi}\right\|_{1,\mathcal{O}} \\
\text{(by regularity (72))} \quad &\lesssim h^{t+1} \left\|\boldsymbol{q}\right\|_{t,\mathcal{O}} \left\|\varepsilon_h^p\right\|_{\mathcal{T}_h}
\end{aligned} \tag{75b}$$

Volumetric term involving ε_h^p :

$$\omega^2 \left(\rho_0 \varepsilon_h^p, \pi_W\theta - \theta\right)_{\mathcal{T}_h} \lesssim \omega^2 h^2 \left\|\varepsilon_h^p\right\|_{\rho_0,K} \left\|\theta\right\|_{2,\mathcal{O}} \lesssim \omega^2 h^2 \left\|\varepsilon_h^p\right\|_{\mathcal{O}} \tag{75c}$$

and

$$\begin{aligned}
\left|\omega^2 \left(\rho_0 \delta_h^p, \pi_W\theta\right)_{\mathcal{T}_h}\right| &= \left|\left(\rho_0 \delta_h^p, \theta\right)_{\mathcal{T}_h} - \left(\rho_0 \delta_h^p, \theta - \pi_W\theta\right)_{\mathcal{T}_h}\right| \\
\text{(by (64))} \quad &= \omega^2 \left|\left(\delta_h^p, \rho_0\theta - \{\rho_0\theta\}\right)_{\mathcal{T}_h} - \left(\rho_0 \delta_h^p, \theta - \pi_W\theta\right)_{\mathcal{T}_h}\right| \\
&\lesssim \omega^2 \left\|p - \pi_W p\right\|_{\mathcal{T}_h} \left\|\rho_0\theta - \{\rho_0\theta\}\right\|_{\mathcal{T}_h} + \omega^2 \left\|p - \pi_W p\right\|_{\mathcal{T}_h} \left\|\theta - \pi_W\theta\right\|_{\mathcal{T}_h} \\
\text{(by (63) and (59a))} \quad &\lesssim \omega^2 h^s \left\|p\right\|_{s,\mathcal{O}} h|\theta|_{1,\mathcal{O}} + \omega^2 h^s \left\|p\right\|_{s,\mathcal{O}} h \left\|\theta\right\|_{1,\mathcal{O}} \\
\text{(by regularity (72))} \quad &\lesssim \omega^2 h^{s+1} \left\|p\right\|_{s,\mathcal{O}} \left\|\varepsilon_h^p\right\|_{\mathcal{T}_h}
\end{aligned} \tag{75d}$$

Volumetric convection term :

$$\begin{aligned}
2\omega \left(\nabla\varepsilon_h^p, (\pi_W\theta - \theta)\boldsymbol{b_0}\right)_{\mathcal{T}_h} &\lesssim \omega \left\|\nabla\varepsilon_h^p\right\|_{\mathcal{T}_h} \left\|\pi_W\theta - \theta\right\|_{\mathcal{T}_h} \\
\text{(by Lemma 4.3)} \quad &\lesssim \omega \left(\left\|\boldsymbol{\varepsilon_h^q}\right\|_{\boldsymbol{W_0},\mathcal{T}_h} + \left\|\boldsymbol{\delta_h^q}\right\|_{\mathcal{T}_h} + \left\|\tau^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h}\right) \left\|\pi_W\theta - \theta\right\|_{\mathcal{T}_h} \\
\text{(by (59b), (59a), (63) and (72))} \quad &\lesssim \omega \left(\left\|\boldsymbol{\varepsilon_h^q}\right\|_{\boldsymbol{W_0},\mathcal{T}_h} + h^t \left\|\boldsymbol{q}\right\|_{t,\mathcal{O}} + \left\|\tau^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h}\right) h \left\|\varepsilon_h^p\right\|_{\mathcal{T}_h} \\
&\lesssim \omega \left(h \left\|\boldsymbol{\varepsilon_h^q}\right\|_{\boldsymbol{W_0},\mathcal{T}_h} + h^{t+1} \left\|\boldsymbol{q}\right\|_{t,\mathcal{O}} + h \left\|\tau^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h}\right) \left\|\varepsilon_h^p\right\|_{\mathcal{T}_h}
\end{aligned} \tag{75e}$$

and

$$\begin{aligned}
2\omega \left|\left(\delta_h^p \boldsymbol{b_0}, \nabla(\pi_W\theta)\right)_{\mathcal{T}_h}\right| &= 2\omega \left|\left(\delta_h^p \boldsymbol{b_0}, \nabla\theta\right)_{\mathcal{T}_h} - \left(\delta_h^p \boldsymbol{b_0}, \nabla(\theta - \pi_W\theta)\right)_{\mathcal{T}_h}\right| \\
\text{(by (59a))} \quad &\lesssim \omega \left\|p - \pi_W p\right\|_{\mathcal{T}_h} \left\|\theta\right\|_{1,\mathcal{O}} \\
\text{(by (59a) and (72))} \quad &\lesssim \omega h^s \left\|p\right\|_{s,\mathcal{O}} \left\|\varepsilon_h^p\right\|_{\mathcal{T}_h}
\end{aligned} \tag{75f}$$

Boundary terms involving $\boldsymbol{b_0}\cdot\boldsymbol{n}$: As we want to keep $s \leqslant k+2$, we will use (59c) with $s-1$ instead of s.
As

$$\left\langle(\boldsymbol{b_0}\cdot\boldsymbol{n})\widehat{\varepsilon}_h^p, \pi_W\theta - P_M\theta\right\rangle_{\partial\mathcal{T}_h} = \left\langle(\boldsymbol{b_0}\cdot\boldsymbol{n})\widehat{\varepsilon}_h^p, \pi_W\theta - \theta\right\rangle_{\partial\mathcal{T}_h}$$

we focus on

$$\begin{aligned}
2\omega \left\langle(\boldsymbol{b_0}\cdot\boldsymbol{n})(\varepsilon_h^p - \widehat{\varepsilon}_h^p), \pi_W\theta - \theta\right\rangle_{\partial\mathcal{T}_h} &\lesssim \omega \left\|\varepsilon_h^p - \widehat{\varepsilon}_h^p\right\|_{\partial\mathcal{T}_h} \left\|\pi_W\theta - \theta\right\|_{\partial\mathcal{T}_h} \\
\text{(by (68))} \quad &\lesssim \omega h^2 \left\|\varepsilon_h^p\right\|_{\mathcal{T}_h} \left[\left\|\boldsymbol{\varepsilon_h^q}\right\|_{\boldsymbol{W_0},\mathcal{T}_h} + h^t \left\|\boldsymbol{q}\right\|_{t,\mathcal{O}} + \left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h}\right]
\end{aligned} \tag{75g}$$

and

$$\begin{aligned}
2\omega \left\langle(\boldsymbol{b_0}\cdot\boldsymbol{n})\widehat{\delta}_h^p, \pi_W\theta - P_M\theta\right\rangle_{\partial\mathcal{T}_h} &\lesssim \omega \left\|\widehat{\delta}_h^p\right\|_{\partial\mathcal{T}_h} \left\|\pi_W\theta - P_M\theta\right\|_{\partial\mathcal{T}_h} \\
\text{(by (59d) and (59c))} \quad &\lesssim \omega h^{s-\frac{1}{2}} \left\|p\right\|_{s,\mathcal{O}} h^{\frac{3}{2}} \left\|\theta\right\|_{2,\mathcal{O}} \\
\text{(by (72))} \quad &\lesssim \omega h^s \left\|p\right\|_{s,\mathcal{O}} \left\|\varepsilon_h^p\right\|_{\mathcal{T}_h}
\end{aligned} \tag{75h}$$

Boundary terms involving τ:

$$\begin{aligned} 2\omega \left\langle \tau \left(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p\right), \pi_W\theta - P_M\theta\right\rangle_{\partial\mathcal{T}_h} &\lesssim \omega \left\|\tau(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h} \left\|\pi_W\theta - P_M\theta\right\|_{\partial\mathcal{T}_h} \\ (\tau = \mathcal{O}(h^{-1})) \quad &\lesssim \omega h^{-\frac{1}{2}} \left\|\tau^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h} h^{\frac{3}{2}} \left\|\theta\right\|_{2,\mathcal{O}} \\ &\lesssim \omega h \left\|\tau^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h} \left\|\varepsilon_h^p\right\|_{\mathcal{T}_h} \end{aligned} \tag{75i}$$

and

$$\begin{aligned} 2\omega \left\langle \tau P_M\delta_h^p, \pi_W\theta - P_M\theta\right\rangle_{\partial\mathcal{T}_h} &= 2\omega \left\langle \tau P_M\delta_h^p, P_M(\pi_W\theta - \theta)\right\rangle_{\partial\mathcal{T}_h} \\ &\lesssim \omega \left\|\tau P_M\delta_h^p\right\|_{\partial\mathcal{T}_h} \left\|P_M(\theta - \pi_W\theta)\right\|_{\partial\mathcal{T}_h} \\ (\tau = \mathcal{O}(h^{-1})) \quad &\lesssim \omega h^{-1} \left\|p - \pi_W p\right\|_{\partial\mathcal{T}_h} \left\|\theta - \pi_W\theta\right\|_{\partial\mathcal{T}_h} \\ (\text{by (59d)}) \quad &\lesssim \omega h^{-1} h^{s-\frac{1}{2}} \left\|p\right\|_{s,\mathcal{O}} h^{2-\frac{1}{2}} \left\|\theta\right\|_{2,\mathcal{O}} \\ &\lesssim \omega h^s \left\|p\right\|_{s,\mathcal{O}} \left\|\varepsilon_h^p\right\|_{\mathcal{T}_h} \end{aligned} \tag{75j}$$

Boundary terms involving τ_{upw}: As we want to keep $s \leqslant k+2$, we will use (59c) with $s-1$ instead of s.

$$\begin{aligned} 2\omega \left\langle \tau_{\text{upw}} \left(\varepsilon_h^p - \widehat{\varepsilon}_h^p\right), \pi_W\theta - P_M\theta\right\rangle_{\partial\mathcal{T}_h} &\lesssim \omega \left\|\varepsilon_h^p - \widehat{\varepsilon}_h^p\right\|_{\partial\mathcal{T}_h} \left\|\pi_W\theta - P_M\theta\right\|_{\partial\mathcal{T}_h} \\ &\lesssim \omega h^{s-\frac{3}{2}} \left\|p\right\|_{s,\mathcal{O}} h^{\frac{3}{2}} \left\|\theta\right\|_{2,\mathcal{O}} \\ &\lesssim \omega h^s \left\|p\right\|_{s,\mathcal{O}} \left\|\varepsilon_h^p\right\|_{\mathcal{T}_h} \end{aligned} \tag{75k}$$

and

$$\begin{aligned} 2\omega \left\langle \tau_{\text{upw}} \left(\delta_h^p - \widehat{\delta}_h^p\right), \pi_W\theta - P_M\theta\right\rangle_{\partial\mathcal{T}_h} &\lesssim \omega h^s \left\|p\right\|_{s,\mathcal{O}} \left\|\theta\right\|_{2,\mathcal{O}} \\ &\lesssim \omega h^s \left\|p\right\|_{s,\mathcal{O}} \left\|\varepsilon_h^p\right\|_{\mathcal{T}_h} \end{aligned} \tag{75l}$$

Boundary term involving $\boldsymbol{\delta}_h^{\boldsymbol{q}}$:

$$\begin{aligned} \left\langle \boldsymbol{\delta}_h^{\boldsymbol{q}} \cdot \boldsymbol{n}, \pi_W\theta - P_M\theta\right\rangle_{\partial\mathcal{T}_h} &\lesssim h^{t-\frac{1}{2}} \left\|\boldsymbol{q}\right\|_{t,\mathcal{O}} h^{\frac{3}{2}} \left\|\theta\right\|_{2,\mathcal{O}} \quad (\text{by (59e) and (59d)}) \\ &\lesssim h^{t+1} \left\|\boldsymbol{q}\right\|_{t,\mathcal{O}} \left\|\varepsilon_h^p\right\|_{\mathcal{T}_h} \end{aligned} \tag{75m}$$

The desired estimate can now be obtained by collecting the estimates (75a), (75b), (75c), (75d), (75e), (75f), (75g), (75h), (75i), (75j), (75k), (75l) and (75m), and using that $h \longrightarrow 0$. ∎

4.3.3 Bootstraping process

We will now combine the results of Corollary 4.1 of Lemma 4.5 and Lemma 4.7 through a bootstrapping process to obtain a convergence result.

Theorem 5 : *Convergence of the HDG+ method*

Assuming that the regularity assumption (72) holds and that $\omega^2 h^2 \|\rho_0\|_\infty C_{\text{reg}} C$ (where C is the constant of THEOREM 4) is small enough, if $p \in H^s(\mathcal{O})$ and $\boldsymbol{q} \in \boldsymbol{H}^{\boldsymbol{t}}(\mathcal{O})$ where $s \in [1, k+2]$ and $t \in [1, k+1]$ then

$$\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} + \sqrt{2\omega}\left(\left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h} + \left\|\left(\frac{1}{2}|\boldsymbol{b_0}\cdot\boldsymbol{n}|\right)^{\frac{1}{2}}(\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h}\right) \lesssim h^t \|\boldsymbol{q}\|_{t,\mathcal{O}} + h^{s-1}\|p\|_{s,\mathcal{O}}$$

and

$$\|\varepsilon_h^p\|_{\mathcal{T}_h} \lesssim (1+\omega)\left(h^{t+1}\|\boldsymbol{q}\|_{t,\mathcal{O}} + h^s\|p\|_{s,\mathcal{O}}\right)$$

Optimal error estimates are

$$\|\pi_W p - p_h\|_{\mathcal{T}_h} = \mathcal{O}(h^{k+2}) \quad \text{and} \quad \|\boldsymbol{\pi_V q} - \boldsymbol{q}_h\|_{\mathcal{T}_h} = \mathcal{O}(h^{k+1}),$$

and

$$\|p - p_h\|_{\mathcal{T}_h} = \mathcal{O}(h^{k+2}) \quad \text{and} \quad \|\boldsymbol{q} - \boldsymbol{q}_h\|_{\mathcal{T}_h} = \mathcal{O}(h^{k+1}).$$

Proof: to make the computations easier we introduce the following notations

$$P := \omega\|\varepsilon_h^p\|_{\mathcal{T}_h} \quad ; \quad Q := \|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} \quad ; \quad T := \left\|\tau^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h}$$

and

$$P_h := \omega h^{s-1}\|p\|_{s,\mathcal{O}} \quad ; \quad Q_h := h^t\|\boldsymbol{q}\|_{t,\mathcal{O}}$$

$$B := \left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h} + \left\|\left(\frac{1}{2}|\boldsymbol{b_0}\cdot\boldsymbol{n}|\right)^{\frac{1}{2}}(\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h}$$

$$\alpha := (\omega+\omega^2)h \quad ; \quad \beta := \left(1+\omega+\omega^2\right)h \quad ; \quad \gamma := \omega^2 h$$

With those shorthands the estimate of LEMMA 4.7 can be rewritten as

$$\begin{aligned} P &\lesssim (\omega+\omega^2)hQ_h + (1+\omega+\omega^2)hP_h + \omega^2 hT + \left(\omega+\omega^2\right)hQ \\ &\lesssim \alpha Q_h + \beta P_h + \gamma B + \alpha Q \end{aligned} \tag{76}$$

and the estimate of COROLLARY 4.1 can be rewritten as

$$\begin{aligned} Q^2 + 2\omega B^2 \lesssim \left|Q^2 - 2i\omega B^2\right| &\lesssim \varepsilon\left(Q^2+T^2\right) + \frac{1}{\varepsilon}\left(P^2+Q_h^2+P_h^2\right) \\ &\lesssim \varepsilon\left(Q^2+B^2\right) + \frac{1}{\varepsilon}\left(P^2+Q_h^2+P_h^2\right) \end{aligned} \tag{77}$$

Taking the square of (76) and using Young's inequality leads to

$$P^2 \lesssim \alpha^2 Q_h^2 + \beta^2 P_h^2 + \gamma^2 B^2 + \alpha^2 Q^2. \tag{78}$$

Using (78) in (77) gives

$$Q^2 + 2\omega B^2 \lesssim \left(1+\frac{1}{\varepsilon}\right)\left[\left(1+\alpha^2\right)Q_h^2 + \left(1+\beta^2\right)P_h^2\right] + \left(\varepsilon + \gamma^2\left(1+\frac{1}{\varepsilon}\right)\right)B^2 + \left(\varepsilon + \alpha^2\left(1+\frac{1}{\varepsilon}\right)\right)Q^2.$$

Let C_1 denote the constant hidden in $\lesssim$. Choosing ε so that $C_1\varepsilon < 1$ and assuming that h is small enough the last two terms of the right hand side can be absorbed by the left hand side, leading to

$$Q^2 + 2\omega B^2 \lesssim \left(1 + \frac{1}{\varepsilon}\right)\left[\left(1+\alpha^2\right)Q_h^2 + \left(1+\beta^2\right)P_h^2\right].$$

As ε does not depend on ω and h, we can hide the first factor of the right-hand side into $\lesssim$ leading to

$$Q^2 + 2\omega B^2 \lesssim \left(1+\alpha^2\right)Q_h^2 + \left(1+\beta^2\right)P_h^2.$$

As $\alpha, \beta = \mathcal{O}(h)$, we can overestimate $\alpha, \beta \lesssim 1$ leading to

$$Q^2 + 2\omega B^2 \lesssim Q_h^2 + P_h^2.$$

And finally we have

$$Q + \sqrt{2\omega}B \lesssim P_h + Q_h. \tag{79}$$

Now by taking $s = k+2$ and $t = k+1$ in (59b) and (59d), we can see that

$$Q = \mathcal{O}(h^{k+1}) \quad \text{and} \quad B = \mathcal{O}(h^{k+1}) \tag{80}$$

and finally by using (79) in (76), we have

$$P = \mathcal{O}(h^{k+2}). \tag{81}$$

■

It is also possible to obtain a convergence result for the trace $\widehat{p_h}$ which is the main unknown of the method :

Corollary 4.2:
Under the assumptions of Theorem 5, the following error estimates for $\widehat{p_h}$ hold

$$\|\widehat{\varepsilon}_h^p\|_{\partial\mathcal{T}_h} = \mathcal{O}(h^{k+\frac{3}{2}}) \quad \text{and} \quad \|p - \widehat{p_h}\|_{\partial\mathcal{T}_h} = \mathcal{O}(h^{k+\frac{1}{2}}).$$

Proof: We have

$$\begin{aligned}
\|\varepsilon_h^p\|_{\partial\mathcal{T}_h} &\leqslant \|P_M\varepsilon_h^p\|_{\partial\mathcal{T}_h} + \|P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p\|_{\partial\mathcal{T}_h} \\
(\tau = \mathcal{O}(h^{-1})) \quad &\lesssim \|P_M\varepsilon_h^p\|_{\partial\mathcal{T}_h} + h^{\frac{1}{2}}\left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h} \\
(P_M \text{ continuous}) \quad &\lesssim \|\varepsilon_h^p\|_{\partial\mathcal{T}_h} + h^{\frac{1}{2}}\left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h} \\
(\text{by (60)}) \quad &\lesssim h^{-\frac{1}{2}}\|\varepsilon_h^p\|_{\mathcal{T}_h} + h^{\frac{1}{2}}\left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h} \\
(\text{by (80) and (81)}) \quad &= \mathcal{O}(h^{-\frac{1}{2}}h^{k+2} + h^{\frac{1}{2}}h^{k+1}) \\
&= \mathcal{O}(h^{k+\frac{3}{2}}),
\end{aligned}$$

which is the first estimate. The second one comes from (59c) with $s = k+1$. ■

4.4 Error analysis of the HDG method with diffusive flux

As the error analysis for the HDG method is very similar to the one for the HDG+ method we only state the main theorem. The intermediate results (error equations, Garding's equality, dual estimate, ...) are stated without proof in Appendix A.

Theorem 6 : *Convergence of the HDG method with diffusive flux*

Assuming that the regularity assumption (72) holds and that $\omega^2 h^2 \|\rho_0\|_\infty C_{\text{reg}} C$ (where C is the constant of THEOREM 4) is small enough, if $p \in H^s(\mathcal{O})$ and $\boldsymbol{q} \in \boldsymbol{H}^{\boldsymbol{t}}(\mathcal{O})$ where $s, t \in [1, k+1]$ then

$$\|\pi_W p - p_h\|_{\mathcal{T}_h} = \mathcal{O}(h^{k+\frac{3}{2}}) \;\; ; \;\; \|P_M p - \widehat{p_h}\|_{\partial\mathcal{T}_h} = \mathcal{O}(h^{k+\frac{1}{2}}) \;\; ; \;\; \text{and} \;\; \|\boldsymbol{\pi_V q} - \boldsymbol{q}_h\|_{\mathcal{T}_h} = \mathcal{O}(h^{k+\frac{1}{2}}),$$

and

$$\|p - p_h\|_{\mathcal{T}_h} = \mathcal{O}(h^{k+1}) \;\; ; \;\; \|p - \widehat{p_h}\|_{\partial\mathcal{T}_h} = \mathcal{O}(h^{k+\frac{1}{2}}) \;\; ; \;\; \text{and} \;\; \|\boldsymbol{q} - \boldsymbol{q}_h\|_{\mathcal{T}_h} = \mathcal{O}(h^{k+\frac{1}{2}}).$$

Remark 4.8: Some HDG methods are known to achieve superconvergence, *ie* taking $p_h \in \mathcal{P}_k$ leads to the following error estimate

$$\|\pi_W p - p_h\|_{\mathcal{T}_h} = \mathcal{O}(h^{k+2}),$$

this is not possible for this method because of the convection term

$$2\omega \left| (\delta_h^p \boldsymbol{b_0}, \nabla(\pi_W \theta))_{\mathcal{T}_h} \right| \lesssim \omega h^s \|p\|_{s,\mathcal{O}} \|\varepsilon_h^p\|_{\mathcal{T}_h}$$

in the dual estimate which locks the convergence rate to $\mathcal{O}(h^{k+1})$ for the scalar variable p_h. Superconvergence is an attractive property for a numerical scheme, indeed using a postprocessing scheme it is possible to use the solution $(p_h, \boldsymbol{q}_h)$ to construct a new approximation $\widetilde{p_h}$ which converges with order $\mathcal{O}(h^{k+2})$, see [Ste91], [CGS10, Sec. 5] for more details.

5 Implementation

In this section, we will give details on how the HDG(+) methods can be implemented. For the sake of simplicity, we will assume that the physical coefficients are constant on each element. In this case, the integrals may be computed using

- an analytic integration procedure that relies on the decomposition of the basis functions in the monomial basis,
- high-order quadrature rules.

Taking into account the variations of the physical parameters inside the elements only requires a straightforward generalization of the material presented here. However this implementation is only possible if quadrature rules are used to evaluate the integrals.

To make the notations lighter, we will drop the subscript h in this section. The quantities $(\boldsymbol{q}, p, \widehat{p})$ will denote the solutions of (46a)–(46b)–(46c) and $(\boldsymbol{q}^K, p^K, \widehat{p}^e)$ their restrictions to an element $K \in \mathcal{T}_h$ and an edge $e \in \mathcal{E}_h$ respectively.
In TABLE 2, we recall the main differences between the three variants of the HDG methods considered in this work.

Variable	Space	HDG $(p_h, \boldsymbol{\sigma}_h)$	HDG $(p_h, \boldsymbol{q}_h)$	HDG+ $(p_h, \boldsymbol{q}_h)$
Pressure p_h	$W_h(K)$	$\mathcal{P}_k(K)$		$\mathcal{P}_{k+1}(K)$
Flux $\boldsymbol{q}_h$ or $\boldsymbol{\sigma}_h$	$\boldsymbol{V}_h(K)$	$\boldsymbol{\mathcal{P}}_k(K)$		
Trace $\widehat{p_h}$	$M_h(e)$	$\mathcal{P}_k(e)$		
Penalization $\tau\|_K$		$\mathcal{O}(1)$		$\mathcal{O}(h_K^{-1})$

Table 2: Choice of local spaces and penalization parameter for the different methods

5.1 Framework and notations

In this section we present the notations used in [BCDL15, FS20] which are very close to the ones used in `hawen`[6]. We will then give details on how the matrices are assembled.

Local problem: The discretization of the local problem (46a)–(46b) leads to the following system

$$\mathbb{A}^K \underline{W^K} + \mathbb{C}^K \underline{\Lambda^K} = \mathcal{S}^K, \tag{82}$$

where

$$\underline{W^K} := \begin{bmatrix} \underline{p^K} & \underline{q_x^K} & \underline{q_y^K} & \underline{q_z^K} \end{bmatrix}^T \quad \text{and } \underline{\Lambda^K} := \begin{bmatrix} \underline{\widehat{p}^{g(K,1)}} & \underline{\widehat{p}^{g(K,2)}} & \underline{\widehat{p}^{g(K,3)}} & \underline{\widehat{p}^{g(K,4)}} \end{bmatrix}^T,$$

and $\underline{p^K}$ denotes the vector of the coefficients of p_h^K in the basis of $W_h(K)$.

Global problem: The discretization of the transmission condition (46c) leads to

$$\sum_{K \in \mathcal{T}_h} \mathbb{B}^K \underline{W^K} + \mathbb{L}^K \underline{\Lambda^K} = 0 \tag{83}$$

We denote by $\underline{\Lambda}$ the global trace approximation and for each element $K \in \mathcal{T}_h$ we introduce the *connectivity map* which is the operator $\mathcal{A}_K$ such that[7]

$$\mathcal{A}_K \underline{\Lambda} = \underline{\Lambda^K}.$$

Using (82), we can express $\underline{W^K}$ in terms of $\underline{\Lambda}$

$$\underline{W^K} = \left(\mathbb{A}^K\right)^{-1} \mathcal{S}^K - \left(\mathbb{A}^K\right)^{-1} \mathbb{C}^K \mathcal{A}_K \underline{\Lambda}$$

and we can finally construct the global problem using (83)

$$\sum_{K \in \mathcal{T}_h} \mathcal{A}_K^T \left[-\mathbb{B}^K \left(\mathbb{A}^K\right)^{-1} \mathbb{C}^K + \mathbb{L}^K \right] \mathcal{A}_K \underline{\Lambda} = - \sum_{K \in \mathcal{T}_h} \mathcal{A}_K^T \mathbb{B}^K \left(\mathbb{A}^K\right)^{-1} \mathcal{S}^K. \tag{84}$$

For conciseness, we will denote

$$\mathbb{K}^K := -\mathbb{B}^K \left(\mathbb{A}^K\right)^{-1} \mathbb{C}^K + \mathbb{L}^K, \qquad \mathbb{K} := \sum_{K \in \mathcal{T}_h} \mathcal{A}_K^T \mathbb{K}^K \mathcal{A}_K \quad \text{and } \mathcal{S} := - \sum_{K \in \mathcal{T}_h} \mathcal{A}_K^T \mathbb{B}^K \left(\mathbb{A}^K\right)^{-1} \mathcal{S}^K.$$

Remark 5.1: the invertibility of matrix $\mathbb{A}^K$ was the object of Theorem 1 (HDG-$\boldsymbol{\sigma}_h$) and Theorem 4 (HDG+ and HDG-$\boldsymbol{q}_h$) and the invertibility of the global matrix $\mathbb{K}$ was the object of Theorem 3.

Remark 5.2: to efficiently implement the HDG(+) methods, we will need to compute the *inverse* of $\mathbb{A}^K$. As this matrix is not too large, it is possible to perform this operation using `lapack`. However the discretization of the global problem will lead to a large sparse system which will be solved using `mumps`.

[6] See https://ffaucher.gitlab.io/hawen-website/.

[7] The aim of operator $\mathcal{A}_K$ is to copy the global informations to the local solver, depending on the element it may be a simple copy or also involve reordering.

Solving HDG: In Algorithm 3, we describe how the HDG method can be efficiently implemented. Details on the construction of the local matrices $\mathbb{A}^K$, $\mathbb{C}^K$, $\mathbb{B}^K$ and $\mathbb{L}^K$ and on the implementation of boundary conditions will be given in the next sections.

Algorithm 3: Solving HDG

`/* Step 1: Construction of the local and global problems */`
for $K \in \mathcal{T}_h$ **do**
  Construct the local matrices $\mathbb{A}^K$, $\mathbb{C}^K$, $\mathbb{B}^K$ and $\mathbb{L}^K$
  Compute $(\mathbb{A}^K)^{-1}$ using `lapack`
  Compute $\mathbb{K}^K = \mathbb{B}^K\left(\mathbb{A}^K\right)^{-1}\mathbb{C}^K + \mathbb{L}^K$
  Modify $\mathbb{K}^K$ to enforce the boundary conditions
  Use the connectivity operator $\mathcal{A}_K$ to add the local contribution to the global matrix $\mathbb{K}$
`/* Step 2: Construction of the source term */`
Localize the source
for $K \in \mathcal{T}_h$ where $s \neq 0$ **do**
  Construct the local source term $\mathcal{S}^K$
  Compute $-\mathbb{B}^K(\mathbb{A}^K)^{-1}\mathcal{S}^K$
  Use the connectivity operator $\mathcal{A}_K$ to add the local contribution to the global source $\mathcal{S}$
`/* Step 3: Resolution of the global system */`
Solve $\mathbb{K}\underline{\Lambda} = \mathcal{S}$ with `mumps`
`/* Step 4: Reconstruction of the solution */`
for $K \in \mathcal{T}_h$ **do**
  Compute the local unknowns $\underline{W^K} = \left(\mathbb{A}^K\right)^{-1}\mathcal{S}^K - \left(\mathbb{A}^K\right)^{-1}\mathbb{C}^K\mathcal{A}_K\underline{\Lambda}$

Convention for the indices: we will always use the following convention for the indices :

- index of basis functions : i (test), j (trial), r (other)
- component of a vector : $u, v \in \{x, y, z\}$
- global number of an edge : m
- local number of an edge : ℓ
- polynomial degree : k

The global number of the ℓ-th face of element K will be denoted by $g(K, \ell)$.

Dimension of polynomial spaces: We recall that the dimension of the space of polynomials of n variables and of degree up to k is given by

$$\dim\left(\mathcal{P}_k(K)\right) = \binom{k+n}{n} =: d_n(k), \qquad \text{for } K \subset \mathbb{R}^n.$$

5.2 Implementation of the diffusive flux HDG method

In this section, we focus on the implementation of the HDG-$\boldsymbol{q}_h$ method.

5.2.1 Definition of the elementary matrices

We introduce the basis $(\Phi_j^K)_{j=1}^{d_n(k)}$ of $\mathcal{P}_k(K)$ and $(\Psi_j^e)_{j=1}^{d_{n-1}(k)}$ of $\mathcal{P}_k(e)$ and we decompose the unknowns as

$$p^K = \sum_{j=1}^{d_n(k)} p_j^K \Phi_j^K; \quad q_u^K = \sum_{j=1}^{d_n(k)} q_{u,j}^K \Phi_j^K; \quad \widehat{p}^e(x) = \sum_{j=1}^{d_{n-1}(k)} \widehat{p}_j^e \Psi_j^e$$

for $u \in \{x, y, z\}$, $K \in \mathcal{T}_h$ and $e \in \mathcal{E}_h$.
We introduce the following matrices

$$\begin{aligned} \mathbb{M}^K_{ij} &= \int_K \Phi^K_i \Phi^K_j \mathrm{d}\boldsymbol{x} \qquad \mathbb{D}^K_{u,ij} = \int_K \Phi^K_j \partial_u \Phi^K_i \mathrm{d}\boldsymbol{x} \\ \mathbb{E}^K_{\ell,ij} &= \int_{\partial K^\ell} \Phi^K_i \Phi^K_j \mathrm{d}\sigma \qquad \mathbb{F}^K_{\ell,ij} = \int_{\partial K^\ell} \Psi^{g(K,\ell)}_j \Phi^K_i \mathrm{d}\sigma \\ \mathbb{G}^m_{ij} &= \int_{e_m} \Psi^m_i \Psi^m_j \mathrm{d}\sigma \end{aligned} \tag{85}$$

where $g(K,\ell)$ is the global number of the ℓ-th face of element K and m is also the global number of the edge e_m.
Notice that all those matrices except $\mathbb{F}^K_\ell$ (which has dimension $d_n(k) \times d_{n-1}(k)$) are square matrices and that all the elementary matrices are real.

5.2.2 Local problem

We introduce the following shorthand to make the algebra easier

$$\alpha_\ell = \min_{\partial K^\ell}(\boldsymbol{b_0} \cdot \boldsymbol{n}, 0) \quad \text{and} \quad t_{\mathrm{upw},\ell} = \max_{\partial K^\ell}(\boldsymbol{b_0} \cdot \boldsymbol{n}, 0)$$

it corresponds to using the expression (44a)–(44b) of the numerical flux. From an implementation point of view, this can be evaluated using array slicing or a ternary operator.
Using the elementary matrices introduced in (85) we can write (46a) as

$$\sum_u W_{0,vu} \mathbb{M}^K \underline{q^K_u} - \mathbb{D}^K_v \underline{p^K} + \sum_\ell n^{K,\ell}_v \mathbb{F}^K_\ell \underline{\widehat{p}^{g(K,\ell)}} = 0, \qquad \forall v \in \{x, y, z\} \tag{86}$$

and (46b) as

$$\begin{aligned} -\omega^2 \rho_0 \mathbb{M}^K \underline{p^K} + 2i\omega \sum_u b_{0,u} \mathbb{D}^K_u \underline{p^K} + \sum_u (\mathbb{D}^K_u)^T \underline{q_u} \\ +2i\omega \sum_\ell \left[-\alpha_\ell \mathbb{F}^K_\ell \underline{\widehat{p}^{g(K,\ell)}} - t_{\mathrm{upw},\ell} \mathbb{E}^K_\ell \underline{p^K} + \tau_\ell (\mathbb{E}^K_\ell \underline{p^K} - \mathbb{F}^K_\ell \underline{\widehat{p}^{g(K,\ell)}}) \right] = \mathbb{S}^K \end{aligned} \tag{87}$$

Matrix form: To construct the global problem (84) we need to construct the matrices $\mathbb{A}^K$ and $\mathbb{C}^K$ of the local problem (82) using (86) and (87).
Matrix $\mathbb{A}^K$:

$\mathbb{A}^K$	$\underline{p^K}$	$\underline{q^K_x}$	$\underline{q^K_y}$	$\underline{q^K_z}$
$\underline{p^K}$	$-\rho_0\omega^2\mathbb{M}^K + 2i\omega \sum_u b_{0,u}\mathbb{D}^K_u$ $+2i\omega \sum_\ell (\tau_\ell - t_{\mathrm{upw},\ell})\mathbb{E}^K_\ell$	$(\mathbb{D}^K_x)^T$ $-\mathbb{D}^K_x + \sum_\ell n^{K,\ell}_x \mathbb{E}^K_\ell$	$(\mathbb{D}^K_y)^T$ $-\mathbb{D}^K_y + \sum_\ell n^{K,\ell}_y \mathbb{E}^K_\ell$	$(\mathbb{D}^K_z)^T$ $-\mathbb{D}^K_z + \sum_\ell n^{K,\ell}_y \mathbb{E}^K_\ell$
$\underline{q^K_x}$	$-\mathbb{D}^K_x$	$W^K_{0,11}\mathbb{M}^K$	$W^K_{0,12}\mathbb{M}^K$	$W^K_{0,13}\mathbb{M}^K$
$\underline{q^K_y}$	$-\mathbb{D}^K_y$	$W^K_{0,21}\mathbb{M}^K$	$W^K_{0,22}\mathbb{M}^K$	$W^K_{0,23}\mathbb{M}^K$
$\underline{q^K_z}$	$-\mathbb{D}^K_z$	$W^K_{0,31}\mathbb{M}^K$	$W^K_{0,32}\mathbb{M}^K$	$W^K_{0,33}\mathbb{M}^K$

The red terms can be used instead of the black ones, depending on wether or not an integration by parts is performed.
Matrix $\mathbb{C}^K$:

$\mathbb{C}^K$	$\underline{\widehat{p}^{g(K,1)}}$	$\underline{\widehat{p}^{g(K,2)}}$	$\underline{\widehat{p}^{g(K,3)}}$	$\underline{\widehat{p}^{g(K,4)}}$
$\underline{p^K}$	$-2i\omega(\tau_1+\alpha_1)\mathbb{F}^K_1$	$-2i\omega(\tau_2+\alpha_2)\mathbb{F}^K_2$	$-2i\omega(\tau_3+\alpha_3)\mathbb{F}^K_3$	$-2i\omega(\tau_4+\alpha4)\mathbb{F}^K_4$
$\underline{q^K_x}$	$n^{K,1}_x\mathbb{F}^K_1$	$n^{K,2}_x\mathbb{F}^K_2$	$n^{K,3}_x\mathbb{F}^K_3$	$n^{K,4}_x\mathbb{F}^K_4$
$\underline{q^K_y}$	$n^{K,1}_y\mathbb{F}^K_1$	$n^{K,2}_y\mathbb{F}^K_2$	$n^{K,3}_y\mathbb{F}^K_3$	$n^{K,4}_y\mathbb{F}^K_4$
$\underline{q^K_z}$	$n^{K,1}_z\mathbb{F}^K_1$	$n^{K,2}_z\mathbb{F}^K_2$	$n^{K,3}_z\mathbb{F}^K_3$	$n^{K,4}_z\mathbb{F}^K_4$

Remark 5.3: We chose to write the matrix form of the method with the equation for p^K first which is the opposite of the system (46a)–(46b)–(46c). This convention is interesting because the matrices for the method in dimension n are submatrices of the ones for dimension $n+1$.

5.2.3 Global problem

The transmission condition (46c) can be written as

$$\sum_{K,\ell}\left(\sum_u n_u^{K,\ell}(\mathbb{F}_\ell^K)^T\underline{q_u^K}+2i\omega\left[-\left(\alpha_\ell\mathbb{G}^{g(K,\ell)}\underline{\widehat{p}}^{g(K,\ell)}+t_{\mathrm{upw},\ell}(\mathbb{F}_\ell^K)^T\underline{p^K}\right)+\tau\left((\mathbb{F}_\ell^K)^T\underline{p^K}-\mathbb{G}^{g(K,\ell)}\underline{\widehat{p}}^{g(K,\ell)}\right)\right]\right)=0 \tag{88}$$

Matrix form: We will now use (88) to construct the matrices $\mathbb{B}^K$ and $\mathbb{L}^K$ of (83).
Matrix $\mathbb{B}^K$:

$\mathbb{B}^K$	$\underline{p^K}$	$\underline{q_x^K}$	$\underline{q_y^K}$	$\underline{q_z^K}$
e_1	$2i\omega(\tau_1-t_{\mathrm{upw},1})(\mathbb{F}_1^K)^T$	$n_x^{K,1}(\mathbb{F}_1^K)^T$	$n_y^{K,1}(\mathbb{F}_1^K)^T$	$n_z^{K,1}(\mathbb{F}_1^K)^T$
e_2	$2i\omega(\tau_2-t_{\mathrm{upw},2})(\mathbb{F}_2^K)^T$	$n_x^{K,2}(\mathbb{F}_2^K)^T$	$n_y^{K,2}(\mathbb{F}_2^K)^T$	$n_z^{K,2}(\mathbb{F}_2^K)^T$
e_3	$2i\omega(\tau_3-t_{\mathrm{upw},3})(\mathbb{F}_3^K)^T$	$n_x^{K,3}(\mathbb{F}_3^K)^T$	$n_y^{K,3}(\mathbb{F}_3^K)^T$	$n_z^{K,3}(\mathbb{F}_3^K)^T$
e_4	$2i\omega(\tau_4-t_{\mathrm{upw},4})(\mathbb{F}_4^K)^T$	$n_x^{K,4}(\mathbb{F}_4^K)^T$	$n_y^{K,4}(\mathbb{F}_4^K)^T$	$n_z^{K,4}(\mathbb{F}_4^K)^T$

Matrix $\mathbb{L}^K$:

$\mathbb{L}^K$	$\underline{\widehat{p}^{g(K,1)}}$	$\underline{\widehat{p}^{g(K,2)}}$	$\underline{\widehat{p}^{g(K,3)}}$	$\underline{\widehat{p}^{g(K,4)}}$
e_1	$-2i\omega(\tau_1+\alpha_1)\mathbb{G}^{g(K,1)}$			
e_2		$-2i\omega(\tau_2+\alpha_2)\mathbb{G}^{g(K,2)}$		
e_3			$-2i\omega(\tau_3+\alpha_3)\mathbb{G}^{g(K,3)}$	
e_4				$-2i\omega(\tau_4+\alpha_4)\mathbb{G}^{g(K,4}$

5.2.4 Boundary conditions:

The transmission condition (46c) was also used to weakly enforce the boundary conditions

$$\begin{aligned}\boldsymbol{q}\cdot\boldsymbol{n}-2i\omega(\boldsymbol{b_0}\cdot\boldsymbol{n})p&=g_N && \text{on } \Gamma_N\\ p&=g_D && \text{on } \Gamma_D\end{aligned}$$

Neumann boundary condition: If we want to enforce the Neumann boundary condition on edge ℓ of element K we only need to add a right-hand side to (88) :

$$\mathbb{G}^{g(K,\ell)}\underline{g_N}$$

where we assumed that

$$g_N=\sum_i g_{N,i}\Psi_i.$$

Weak Dirichlet boundary conditions: However if we want to enforce the Dirichlet boundary condition on edge ℓ of element K we need to change the matrices $\mathbb{B}^K$ and $\mathbb{L}^K$. This first method of implementation corresponds to weakly enforcing the boundary conditions, *ie* using the expression

$$\langle\widehat{p_h}-g_D,\mu\rangle_{\Gamma_D}=0.$$

This method is used for a *modal* implementation of the HDG method. It can be implemented as follows described in Algorithm 4.

Algorithm 4: Implementation of weak Dirichlet BC

1 The corresponding row of $\mathbb{B}^K$ is set to 0 : $\mathbb{B}^K[e_\ell, :] = 0$
2 The corresponding entry of $\mathbb{L}^K$ is changed : $\mathbb{L}^K[e_\ell, \underline{\widehat{p}^{g(K,\ell)}}] = \mathbb{G}^{g(K,\ell)}$
3 The following right-hand side is added : $\mathbb{G}^{g(K,\ell)}\underline{g_D}$, where $g_D = \sum_i g_{D,i}\Psi_i$.

Strong Dirichlet boundary conditions: For a nodal implementation of the HDG method, it is also possible to strongly enforce the boundary condition on Γ_D :

$$\widehat{p_h} = g_D,$$

on edge ℓ of element K. We need to change the local contribution

$$\mathbb{K}^K = \mathbb{L}^K - \mathbb{B}^K\left(\mathbb{A}^K\right)^{-1}\mathbb{C}^K,$$

to the global problem as described in Algorithm 5. We recall that $\mathbb{K}^K$ has the same shape as $\mathbb{L}^K$.

Algorithm 5: Implementation of strong Dirichlet BC

1 The corresponding entries of $\mathbb{K}^K$ are replaced with an identity block : $\mathbb{K}^K[e_\ell, \underline{\widehat{p}^{g(K,\ell)}}] = \mathbf{Id}$
2 The following right-hand side is added : $[g_D(\boldsymbol{x}_r)]_r^T$, where $\boldsymbol{x}_r \in \text{dof}\left(e^{g(K,\ell)}\right)$

5.3 Implementation of the total flux HDG method

Using the same notations as in the previous section, we can write the discrete form of the *total flux* HDG method (15a)–(15b)–(15c).

5.3.1 Local problem

The local problem (12a)–(12b) can be written as

$$\sum_u W_{0,vu}\underline{\sigma_u^K} - \mathbb{D}_v^K\underline{p^K} + 2i\omega\left(\sum_u W_{0,vu}b_{0,u}\right)\mathbb{M}^K\underline{p^K} + \sum_\ell n_v^\ell\mathbb{F}_\ell^K\underline{\widehat{p}^{g(K,\ell)}} = 0, \quad \forall v \in \{x, y, e\}$$
$$-\omega^2\rho_0\mathbb{M}^K\underline{p^K} + \sum_u(\mathbb{D}_u^K)^T\underline{\sigma_u^K} + i\omega\sum_\ell\tau^\ell\left(\mathbb{E}_\ell^K\underline{p^K} - \mathbb{F}_\ell^K\underline{\widehat{p}^{g(K,\ell)}}\right) = \mathbb{S}^K$$

Matrix form: Matrix $\mathbb{A}^K$: We introduce $\beta_v := \sum_u W_{0,vu}b_{0,u}$.

$\mathbb{A}^K$	$\underline{p^K}$	$\underline{\sigma_x^K}$	$\underline{\sigma_y^K}$	$\underline{\sigma_z^K}$
$\underline{p^K}$	$-\rho_0\omega^2\mathbb{M}^K + i\omega\sum_\ell\tau_\ell\mathbb{E}_\ell^K$	$(\mathbb{D}_x^K)^T$	$(\mathbb{D}_y^K)^T$	$(\mathbb{D}_z^K)^T$
$\underline{\sigma_x^K}$	$-\mathbb{D}_x^K + 2i\omega\beta_x\mathbb{M}^K$	$W_{0,11}^K\mathbb{M}^K$	$W_{0,12}^K\mathbb{M}^K$	$W_{0,13}^K\mathbb{M}^K$
$\underline{\sigma_y^K}$	$-\mathbb{D}_y^K + 2i\omega\beta_y\mathbb{M}^K$	$W_{0,21}^K\mathbb{M}^K$	$W_{0,22}^K\mathbb{M}^K$	$W_{0,23}^K\mathbb{M}^K$
$\underline{\sigma_z^K}$	$-\mathbb{D}_z^K + 2i\omega\beta_z\mathbb{M}^K$	$W_{0,31}^K\mathbb{M}^K$	$W_{0,32}^K\mathbb{M}^K$	$W_{0,33}^K\mathbb{M}^K$

Matrix $\mathbb{C}^K$:

$\mathbb{C}^K$	$\underline{\widehat{p}^{g(K,1)}}$	$\underline{\widehat{p}^{g(K,2)}}$	$\underline{\widehat{p}^{g(K,3)}}$	$\underline{\widehat{p}^{g(K,4)}}$
$\underline{p^K}$	$-i\omega\tau_1\mathbb{F}_1^K$	$-i\omega\tau_2\mathbb{F}_2^K$	$-i\omega\tau_3\mathbb{F}_3^K$	$-i\omega\tau_4\mathbb{F}_4^K$
$\underline{\sigma_x^K}$	$n_x^{K,1}\mathbb{F}_1^K$	$n_x^{K,2}\mathbb{F}_2^K$	$n_x^{K,3}\mathbb{F}_3^K$	$n_x^{K,4}\mathbb{F}_4^K$
$\underline{\sigma_y^K}$	$n_y^{K,1}\mathbb{F}_1^K$	$n_y^{K,2}\mathbb{F}_2^K$	$n_y^{K,3}\mathbb{F}_3^K$	$n_y^{K,4}\mathbb{F}_4^K$
$\underline{\sigma_z^K}$	$n_z^{K,1}\mathbb{F}_1^K$	$n_z^{K,2}\mathbb{F}_2^K$	$n_z^{K,3}\mathbb{F}_3^K$	$n_z^{K,4}\mathbb{F}_4^K$

5.3.2 Global problem

The discrete transmission condition (14) can be written

$$\sum_{K,\ell}\left[n_u^{K,\ell}(\mathbb{F}_\ell^K)^T\underline{\sigma_u^K}+i\omega\tau^\ell\left((\mathbb{F}_\ell^K)\underline{p^K}-\mathbb{G}^{g(K,\ell)}\underline{\widehat{p}^{g(K,\ell)}}\right)\right]=0$$

Matrix form: Matrix $\mathbb{B}^K$:

$\mathbb{B}^K$	$\underline{p^K}$	$\underline{\sigma_x^K}$	$\underline{\sigma_y^K}$	$\underline{\sigma_z^K}$
e_1	$i\omega\tau_1(\mathbb{F}_1^K)^T$	$n_x^{K,1}(\mathbb{F}_1^K)^T$	$n_y^{K,1}(\mathbb{F}_1^K)^T$	$n_z^{K,1}(\mathbb{F}_1^K)^T$
e_2	$i\omega\tau_2(\mathbb{F}_2^K)^T$	$n_x^{K,2}(\mathbb{F}_2^K)^T$	$n_y^{K,2}(\mathbb{F}_2^K)^T$	$n_z^{K,2}(\mathbb{F}_2^K)^T$
e_3	$i\omega\tau_3(\mathbb{F}_3^K)^T$	$n_x^{K,3}(\mathbb{F}_3^K)^T$	$n_y^{K,3}(\mathbb{F}_3^K)^T$	$n_z^{K,3}(\mathbb{F}_3^K)^T$
e_4	$i\omega\tau_4(\mathbb{F}_4^K)^T$	$n_x^{K,4}(\mathbb{F}_4^K)^T$	$n_y^{K,4}(\mathbb{F}_4^K)^T$	$n_z^{K,4}(\mathbb{F}_4^K)^T$

Matrix $\mathbb{L}^K$:

$\mathbb{L}^K$	$\underline{\widehat{p}^{g(K,1)}}$	$\underline{\widehat{p}^{g(K,2)}}$	$\underline{\widehat{p}^{g(K,3)}}$	$\underline{\widehat{p}^{g(K,4)}}$
e_1	$-i\omega\tau_1\mathbb{G}^{g(K,1)}$			
e_2		$-i\omega\tau_2\mathbb{G}^{g(K,2)}$		
e_3			$-i\omega\tau_3\mathbb{G}^{g(K,3)}$	
e_4				$-i\omega\tau_4\mathbb{G}^{g(K,4}$

5.4 Implementation of the HDG+ method

As $p^K\in\mathcal{P}_{k+1}(K)$, $\boldsymbol{q}^K\in\boldsymbol{\mathcal{P}}_k(K)$ and $\widehat{p}^e\in\mathcal{P}_k(e)$, one should be very careful while writing the matrix form of the system, indeed we have

$$\underline{p^K}\in\mathbb{C}^{d_n(k+1)}\quad\text{and}\quad\underline{\boldsymbol{q}^K}\in\mathbb{C}^{n\times d_n(k)}\ \ (\Longleftrightarrow\underline{q_u^K}\in\mathbb{C}^{d_n(k)},u\in\{x,y,z\})$$

so some of the elementary matrices will be rectangular.

5.4.1 Definition of the elementary matrices

Let $(\Phi_j^{K,k})_{j=1}^{d_n(k)}$ be the basis for $\mathcal{P}_k(K)$ and $(\Psi_j^{e,k})_{j=1}^{d_{n-1}(k)}$ be the basis for $\mathcal{P}_k(e)$.
The unknowns are therefore decomposed in the following way

$$p^K=\sum_{j=1}^{d_n(k+1)}p_j^K\Phi_j^{K,k+1}\ \ ;\ \ q_u^K=\sum_{j=1}^{d_n(k)}q_{u,j}^K\Phi_j^{K,k}\ \ ;\ \ \widehat{p}^e=\sum_{j=1}^{d_{n-1}(k)}\widehat{p}_j^e\Psi_j^{e,k}$$

for $u\in\{x,y,z\}$.
We define the following elementary matrices :

$$\begin{aligned}
\mathbb{M}_{ij}^{K,k}&=\int_K\Phi_i^{K,k}\Phi_j^{K,k}\mathrm{d}\boldsymbol{x}\\
\mathbb{D}_{u,ij}^{K,k}&=\int_K\Phi_j^{K,k}\partial_u\Phi_i^{K,k}\mathrm{d}\boldsymbol{x}\qquad&\mathbb{D}_{u,ij}^{K,k,k+1}&=\int_K\Phi_j^{K,k+1}\partial_u\Phi_i^{K,k}\mathrm{d}\boldsymbol{x}\\
\mathbb{F}_{\ell,ij}^{K,k}&=\int_{\partial K^\ell}\Phi_i^{K,k}\Psi_j^{g(K,\ell),k}\mathrm{d}\sigma\qquad&\mathbb{F}_{\ell,ij}^{K,k+1,k}&=\int_{\partial K^\ell}\Phi_i^{K,k+1}\Psi_j^{g(K,\ell),k}\mathrm{d}\sigma\\
\mathbb{E}_{\ell,ij}^{K,k}&=\int_{\partial K^\ell}\Phi_i^{K,k}\Phi_j^{K,k}\mathrm{d}\sigma\\
\mathbb{G}_{i,j}^{m}&=\int_{e_m}\Psi_i^{m,k}\Psi_j^{m,k}\mathrm{d}\sigma
\end{aligned}\tag{89}$$

where $g(K,\ell)$ is the global number of the ℓ-th face of element K and m is also the global number of the edge e_m.
Most of those matrices are not square matrices and their sizes are recalled in Table 3.

Matrix	Rows	Columns
$\mathbb{M}^{K,k+1}$	$d_n(k+1)$	$d_n(k+1)$
$\mathbb{D}_u^{K,k+1}$	$d_n(k+1)$	$d_n(k+1)$
$\mathbb{D}_u^{K,k,k+1}$	$d_n(k)$	$d_n(k+1)$
$\mathbb{F}_\ell^{K,k+1,k}$	$d_n(k+1)$	$d_{n-1}(k)$
$\mathbb{T}_\ell^{K,k+1}$	$d_n(k+1)$	$d_n(k+1)$
$\mathbb{E}_\ell^{K,k+1}$	$d_n(k+1)$	$d_n(k+1)$

(a) Matrices needed for (46b)

Matrix	Rows	Columns
$\mathbb{M}^{K,k}$	$d_n(k)$	$d_n(k)$
$\mathbb{D}_u^{K,k,k+1}$	$d_n(k)$	$d_n(k+1)$
$\mathbb{F}_\ell^{K,k}$	$d_n(k)$	$d_{n-1}(k)$

(b) Matrices needed for (46a)

Matrix	Rows	Columns
$\mathbb{F}_\ell^{K,k}$	$d_n(k)$	$d_{n-1}(k)$
$\mathbb{F}_\ell^{K,k+1,k}$	$d_n(k+1)$	$d_{n-1}(k)$
$\mathbb{G}_\ell^{m,k}$	$d_{n-1}(k)$	$d_{n-1}(k)$

(c) Matrices needed for (46c)

Table 3: Summary of the dimensions of the local matrices

The matrix $\mathbb{T}_\ell^{K,k+1}$ will be used to evaluate the projection $P_M p$ and will be defined in (93). It was added to Table 3 for completeness.

Remark 5.4: To compute the integrals involving polynomials of different degrees in a nodal framework, it is possible to directly use the expressions given in (89) or to use the following trick. Let $(\boldsymbol{x}_j^{k+1})_{j=1}^{d_n(k+1)}$ be the degrees of freedom associated with $\mathcal{P}_{k+1}(K)$. We define the *projection matrix* $\mathbb{P}$ by

$$\mathbb{P}_{ij}^{K,k,k+1} = \Phi_i^{K,k}(\boldsymbol{x}_j^{k+1}).$$

Seeing $\Phi_i^{K,k}$ as as polynomial of degree $k+1$, we can express it in the basis $(\Phi_j^{K,k+1})_{j=1}^{d_n(k+1)}$. Due to the nodal nature of the basis functions, we have

$$\Phi_i^{K,k} = \sum_{r=1}^{d_n(k+1)} \Phi_i^{K,k}(\boldsymbol{x}_j^{k+1})\Phi_r^{K,k+1} = \sum_{r=1}^{d_n(k+1)} \mathbb{P}_{ir}^{K,k,k+1}\Phi_r^{K,k+1}.$$

The matrix $\mathbb{D}^{K,k,k+1}$ is therefore given by

$$\mathbb{D}^{K,k,k+1} = \mathbb{P}^{K,k,k+1}\mathbb{D}^{K,k+1}.$$

5.4.2 Evaluating the projection:

The evaluation of the term

$$\langle \tau P_M p_h, w_h\rangle_{\partial K} \tag{90}$$

where $w_h \in \mathcal{P}_{k+1}(K)$ is a difficult part of the implementation of the HDG+ method.
Notice however that the implementation of the term

$$\langle P_M p_h, \mu_h\rangle_{\partial K}, \quad \mu \in M_h(\partial K)$$

arising in the discretization of the transmission condition (46c) is straightforward. Indeed as $\mu_h \in \mathcal{P}_k(e)$, by definition of P_M we have

$$\langle P_M p_h, \mu_h\rangle_{\partial K} = \langle p_h, \mu_h\rangle_{\partial K}.$$

In this section we present three techniques to compute (90).

Efficient implementation using a hierarchical basis: The easiest way to compute (90) is to use a hierarchical and orthogonal basis for $\mathcal{P}_{k+1}(K)$. Let $(\Phi_j^K)_{j=1}^{d_n(k+1)}$ be such a basis, then $(\Phi_j^K)_{j=1}^{d_n(k)}$ is a basis for $\mathcal{P}_k(K)$. Therefore if

$$p^K = \sum_{j=1}^{d_n(k+1)} p_j^K \Phi_j^K$$

then

$$P_M p^K = \sum_{j=1}^{d_n(k)} p_j^K \Phi_j^K.$$

A good choice for such a basis would probably be Dubiner's one which is L^2-orthogonal if the reference element is

- *in 2D:* the triangle with vertices

$$\boldsymbol{x}_1 = (-1,-1) \;\; ; \;\; \boldsymbol{x}_2 = (1,-1) \;\; ; \;\; \boldsymbol{x}_3 = (-1,1),$$

- *in 3D:* the tetrahedron with vertices

$$\boldsymbol{x}_1 = (-1,-1,-1) \;\; ; \;\; \boldsymbol{x}_2 = (1,-1,-1) \;\; ; \;\; \boldsymbol{x}_3 = (-1,1,-1) \;\; ; \;\; \boldsymbol{x}_4 = (-1,-1,1).$$

See [Dub91, Kir04] for more details.
As our solver is developed in the framework of nodal discontinuous Galerkin methods, we will not use this method for our implementation.

Efficient implementation through Gauss-Legendre quadrature: As τ is constant on each edge, we have

$$\langle \tau P_M p_h, u \rangle_{\partial K} = \langle \tau P_M p_h, P_M u \rangle_{\partial K}.$$

This integral can be efficiently computed in two dimensions by using a Gauss-Legendre quadrature. This trick is described in [Oik14, Sec. 3.4]. In two dimensions, the edges of a triangle is a one-dimensional interval I. For simplicity, we assume that $I = [-1,1]$.
Denote by $\mathcal{G}_k$ the k points Gauss-Legendre quadrature rule. For a function f, it is given by

$$\mathcal{G}_k[f] = \sum_{i=1}^{k} w_i f(a_i)$$

where $(w_i)_{i\in[\![1,k]\!]}$ are the quadrature weights and $(a_i)_{i\in[\![1,k]\!]}$ are the quadrature points. If the function f is regular enough, then $\mathcal{G}_k$ approximates the integral of f over I. In particular $\mathcal{G}_{k+1}$ is exact for polynomials of degree up to $2k+1$. The usual stabilization term with $p, u \in \mathcal{P}_{k+1}(K)$, can be exactly computed with the $k+2$ quadrature rule $\mathcal{G}_{k+2}$ (which is exact for polynomials of degree up to $2k+3$), *ie.*

$$\mathcal{G}_{k+2}[\tau p u] = \langle \tau p, u \rangle_{\partial K}.$$

If the $k+1$ points rule $\mathcal{G}_{k+1}$ is used instead, the HDG+ stabilization term is obtained

$$\mathcal{G}_{k+1}[\tau p u] = \langle \tau P_M p, P_M u \rangle_{\partial K}.$$

Lemma 5.1:

$$\forall u, v \in \mathcal{P}_{k+1}(I), \quad \mathcal{G}_{k+1}[uv] = \int_I P_M u P_M v \mathrm{d}x$$

Proof: Let φ_m be the Legendre polynomial of order $m \geqslant 0$. On I, we can write

$$u = \sum_{i=0}^{k+1} u_i \varphi_i \quad \text{and} \quad v = \sum_{i=0}^{k+1} v_i \varphi_i.$$

Due to the hierarchical nature of the Legendre polynomials, we know that

$$P_M u = \sum_{i=0}^{k} u_i \varphi_i \quad \text{and} \quad P_M v = \sum_{i=0}^{k} v_i \varphi_i.$$

We recall that $\mathcal{G}_{k+1}$ is exact for polynomials of degree up to $2k+1$ and that φ_{k+1} vanishes at the quadrature points, *ie.*

$$\forall i \in [\![1, k+1]\!], \quad \varphi_{k+1}(a_i) = 0.$$

Then we have

$$\begin{aligned}
\mathcal{G}_{k+1}[uv] &= \sum_{i=1}^{k+1} \left[w_i \sum_{\ell=0}^{k+1} u_\ell \varphi_\ell(a_i) \sum_{m=0}^{k+1} v_m \varphi_m(a_i) \right] \\
&= \sum_{i=1}^{k+1} \left[w_i \sum_{\ell=0}^{k} u_\ell \varphi_\ell(a_i) \sum_{m=0}^{k} v_m \varphi_m(a_i) \right] \\
&= \sum_{i=1}^{k+1} w_i (P_M u)(a_i)(P_M v)(a_i) \\
&= \mathcal{G}_{k+1}[P_M u P_M v] \\
&= \int_I P_M u P_M v \mathrm{d}x,
\end{aligned}$$

as $\deg(P_M u P_M v) = 2k$.

Remark 5.5: the decomposition onto the Legendre basis is needed to prove the result, but another basis (*eg.* Lagrange's nodal basis) can be used in the implementation of the method.

Remark 5.6: unfortunately this result is not yet generalized for three-dimensional problems.

Generic implementation: We will now present a way to compute (90) that works for every choice of polynomial basis in every space dimension. As this technique requires to precompute the projections for all the basis functions of order $k+1$, the nodal HDG+ method may be less attractive when p-adaptivity is needed.
We assume that

$$p^K = \sum_{j=1}^{d_n(k+1)} p_j^K \Phi_j^{K,k+1} \quad \text{and} \quad w_h = \Phi_i^{K,k+1}$$

(90) therefore becomes

$$\left\langle P_M p_h^K, w_h \right\rangle_{\partial K^\ell} = \sum_{j=1}^{d_n(k+1)} p_j^K \left\langle P_M \Phi_j^{K,k+1}, \Phi_i^{K,k+1} \right\rangle_{\partial K^\ell}. \tag{91}$$

The next step is to decompose $P_M \Phi_j^{K,k+1}$ onto the basis of $M_h(\partial K)$

$$P_M \Phi_j^{K,k+1} = \sum_{r=1}^{d_{n-1}(k)} \theta_{j,r}^{K,\ell} \Psi_r^{e,k},$$

and we can construct a linear system for $\underline{\theta_j^{K,\ell}}$ using the definition of P_M

$$\sum_{r=1}^{d_{n-1}(k)} \left\langle \Psi_r^{e,k}, \Psi_i^{e,k} \right\rangle_{\partial K^\ell} \theta_{j,r}^{K,\ell} = \left\langle \Phi_j^{K,k+1}, \Psi_i^{e,k} \right\rangle_{\partial K^\ell}, \qquad \forall i \in [\![1, d_{n-1}(k)]\!]$$

or in matrix form

$$\mathbb{G}^{g(K,\ell)} \underline{\theta_j^{K,\ell}} = \underline{s_j^{K,\ell}}$$

where

$$s_{j,i}^{K,\ell} := \left\langle \Phi_j^{K,k+1}, \Psi_i^{e,k} \right\rangle_{\partial K^\ell}.$$

Plugging this result into (91), we have

$$\left\langle P_M p_h^K, \Phi_i^{K,k+1} \right\rangle_{\partial K^\ell} = \sum_{j=1}^{d_n(k+1)} \sum_{r=1}^{d_{n-1}(k)} p_j^K \theta_{j,r}^{K,\ell} \left\langle \Psi_r^{e,k}, \Phi_i^{K,k+1} \right\rangle_{\partial K^\ell}. \tag{92}$$

We define the following matrix

$$\mathbb{T}_{\ell,ij}^{K,k+1} := \sum_{r=1}^{d_{n-1}(k)} \mathbb{F}_{\ell,ir}^{K,k+1,k} \theta_{j,r}^{K,\ell}, \tag{93}$$

where we recall that

$$\mathbb{F}_{\ell,ij}^{K,k+1,k} := \int_{\partial K^\ell} \Phi_i^{K,k+1} \Psi_j^{g(K,\ell),k} \mathrm{d}\sigma,$$

and we can rewrite (92) in matrix form

$$\left\langle P_M p_h^K, \Phi_i^{K,k+1} \right\rangle_{\partial K} = \sum_{\ell=1}^{4} \sum_{j=1}^{d_n(k+1)} \mathbb{T}_{\ell,ij}^{K,k+1} p_j^K.$$

5.4.3 Local problem

Using the elementary matrices introduced in (89) we can write (46b) as

$$-\rho_0 \omega^2 \mathbb{M}^{K,k+1} \underline{p^K} + 2i\omega \sum_u b_0^u \mathbb{D}_u^{K,k+1} \underline{p^K} + \sum_u (\mathbb{D}_u^{K,k,k+1})^T \underline{q_u^K} + 2i\omega \sum_\ell \tau_\ell \left(\mathbb{T}_\ell^{K,k+1} \underline{p^K} - \mathbb{F}^{K,k+1,k} \underline{\widehat{p}^{g(K,\ell)}} \right)$$

$$-2i\omega \sum_\ell \left[\alpha_\ell \mathbb{F}^{K,k+1,k} \underline{\widehat{p}^{g(K,\ell)}} + t_{\mathrm{upw},\ell} \mathbb{E}_\ell^{K,k+1} \underline{p^K} \right] = \mathbb{S}^K \tag{94}$$

and (46a) as

$$\sum_u W_{0,vu} \mathbb{M}^{K,k} \underline{q_u^K} - \mathbb{D}_v^{K,k,k+1} \underline{p^K} + \sum_\ell n_v^{K,\ell} \mathbb{F}_\ell^{K,k} \underline{\widehat{p}^{g(K,\ell)}} = 0 \qquad \forall v \in \{x, y, z\} \tag{95}$$

Using Table 3 it is easy (and important) to check that (94) has $d_n(k+1)$ equations and (95) has $d_n(k)$ equations.

Matrix form: We can now construct the matrices $\mathbb{A}^K$ and $\mathbb{C}^K$ for the local problem (82).
Matrix $\mathbb{A}^K$: (size : $(d_n(k+1)+3d_n(k))^2$)

$\mathbb{A}^K$	$\underline{p^K}$	$\underline{q_x^K}$	$\underline{q_y^K}$	$\underline{q_z^K}$
$\underline{p^K}$	$-\rho_0\omega^2\mathbb{M}^{K,k+1}+2i\omega\sum_u b_{0,u}\mathbb{D}_u^{K,k+1}$ $-2i\omega\sum_\ell t_{\mathrm{upw},\ell}\mathbb{E}_\ell^{K,k+1}$ $+2i\omega\sum_\ell \tau_\ell\mathbb{T}_\ell^{K,k+1}$	$(\mathbb{D}_x^{K,k,k+1})^T$	$(\mathbb{D}_y^{K,k,k+1})^T$	$(\mathbb{D}_z^{K,k,k+1})^T$
$\underline{q_x^K}$	$-\mathbb{D}_x^{K,k,k+1}$	$W_{0,11}^K\mathbb{M}^{K,k}$	$W_{0,12}^K\mathbb{M}^{K,k}$	$W_{0,13}^K\mathbb{M}^{K,k}$
$\underline{q_y^K}$	$-\mathbb{D}_y^{K,k,k+1}$	$W_{0,21}^K\mathbb{M}^{K,k}$	$W_{0,22}^K\mathbb{M}^{K,k}$	$W_{0,23}^K\mathbb{M}^{K,k}$
$\underline{q_z^K}$	$-\mathbb{D}_z^{K,k,k+1}$	$W_{0,31}^K\mathbb{M}^{K,k}$	$W_{0,32}^K\mathbb{M}^{K,k}$	$W_{0,33}^K\mathbb{M}^{K,k}$

Matrix $\mathbb{C}^K$: (size: $(d_n(k+1)+3d_n(k))\times(4d_{n-1}(k))$)

$\mathbb{C}^K$	$\underline{\widehat{p}^{g(K,1)}}$	$\underline{\widehat{p}^{g(K,2)}}$	$\underline{\widehat{p}^{g(K,3)}}$	$\underline{\widehat{p}^{g(K,4)}}$
$\underline{p^K}$	$-2i\omega(\tau_1+\alpha_1)\mathbb{F}_1^{K,k+1,k}$	$-2i\omega(\tau_2+\alpha_2)\mathbb{F}_2^{K,k+1,k}$	$-2i\omega(\tau_3+\alpha_3)\mathbb{F}_3^{K,k+1,k}$	$-2i\omega(\tau_4+\alpha_4)\mathbb{F}_4^{K,k+1,k}$
$\underline{q_x^K}$	$n_x^{K,1}\mathbb{F}_1^{K,k}$	$n_x^{K,2}\mathbb{F}_2^{K,k}$	$n_x^{K,3}\mathbb{F}_3^{K,k}$	$n_x^{K,4}\mathbb{F}_4^{K,k}$
$\underline{q_y^K}$	$n_y^{K,1}\mathbb{F}_1^{K,k}$	$n_y^{K,2}\mathbb{F}_2^{K,k}$	$n_y^{K,3}\mathbb{F}_3^{K,k}$	$n_y^{K,4}\mathbb{F}_4^{K,k}$
$\underline{q_z^K}$	$n_z^{K,1}\mathbb{F}_1^{K,k}$	$n_z^{K,2}\mathbb{F}_2^{K,k}$	$n_z^{K,3}\mathbb{F}_3^{K,k}$	$n_z^{K,4}\mathbb{F}_4^{K,k}$

5.4.4 Global problem

Using the elementary matrices introduced in (89) we can write (46c) as

$$\sum_{K,\ell}\left[\sum_u n_u^{K,\ell}(\mathbb{F}_\ell^{K,k})^T\underline{q_u^K}+2i\omega\sum_\ell\tau_\ell\left((\mathbb{F}_\ell^{K,k+1,k})^T\underline{p^K}-\mathbb{G}^{g(K,\ell),k}\underline{\widehat{p}^{g(K,\ell)}}\right)\right]$$
$$-2i\omega\sum_{K,\ell}\left[\alpha_\ell\mathbb{G}^{g(K,\ell),k}\underline{\widehat{p}^{g(K,\ell)}}+t_{\mathrm{upw},\ell}(\mathbb{F}_\ell^{K,k+1,k})^T\underline{p^K}\right]=0 \tag{96}$$

Once again, Table 3 should be used to check the dimensions of the matrices involved in (96).

Matrix form: We will now use (96) to construct the matrices $\mathbb{B}^K$ and $\mathbb{L}^K$ of (83).
Matrix $\mathbb{B}^K$: (size: $4d_{n-1}(k)\times(d_n(k+1)+3d_n(k))$)

$\mathbb{B}^K$	$\underline{p^K}$	$\underline{q_x^K}$	$\underline{q_y^K}$	$\underline{q_z^K}$
e_1	$2i\omega(\tau_1-t_{\mathrm{upw},1})(\mathbb{F}_1^{K,k+1,k})^T$	$n_x^{K,1}(\mathbb{F}_1^{K,k})^T$	$n_y^{K,1}(\mathbb{F}_1^{K,k})^T$	$n_z^{K,1}(\mathbb{F}_1^{K,k})^T$
e_2	$2i\omega(\tau_2-t_{\mathrm{upw},2})(\mathbb{F}_2^{K,k+1,k})^T$	$n_x^{K,2}(\mathbb{F}_2^{K,k})^T$	$n_y^{K,2}(\mathbb{F}_2^{K,k})^T$	$n_z^{K,2}(\mathbb{F}_2^{K,k})^T$
e_3	$2i\omega(\tau_3-t_{\mathrm{upw},3})(\mathbb{F}_3^{K,k+1,k})^T$	$n_x^{K,3}(\mathbb{F}_3^{K,k})^T$	$n_y^{K,3}(\mathbb{F}_3^{K,k})^T$	$n_z^{K,3}(\mathbb{F}_3^{K,k})^T$
e_4	$2i\omega(\tau_4-t_{\mathrm{upw},4})(\mathbb{F}_4^{K,k+1,k})^T$	$n_x^{K,4}(\mathbb{F}_4^{K,k})^T$	$n_y^{K,4}(\mathbb{F}_4^{K,k})^T$	$n_z^{K,4}(\mathbb{F}_4^{K,k})^T$

Matrix $\mathbb{L}^K$: (size: $(4d_{n-1}(k))^2$)

$\mathbb{L}^K$	$\underline{\widehat{p}^{g(K,1)}}$	$\underline{\widehat{p}^{g(K,2)}}$	$\underline{\widehat{p}^{g(K,3)}}$	$\underline{\widehat{p}^{g(K,4)}}$
e_1	$-2i\omega(\tau_1+\alpha_1)\mathbb{G}^{g(K,1),k}$			
e_2		$-2i\omega(\tau_2+\alpha_2)\mathbb{G}^{g(K,2),k}$		
e_3			$-2i\omega(\tau_3+\alpha_3)\mathbb{G}^{g(K,3),k}$	
e_4				$-2i\omega(\tau_4+\alpha_4)\mathbb{G}^{g(K,4),k}$

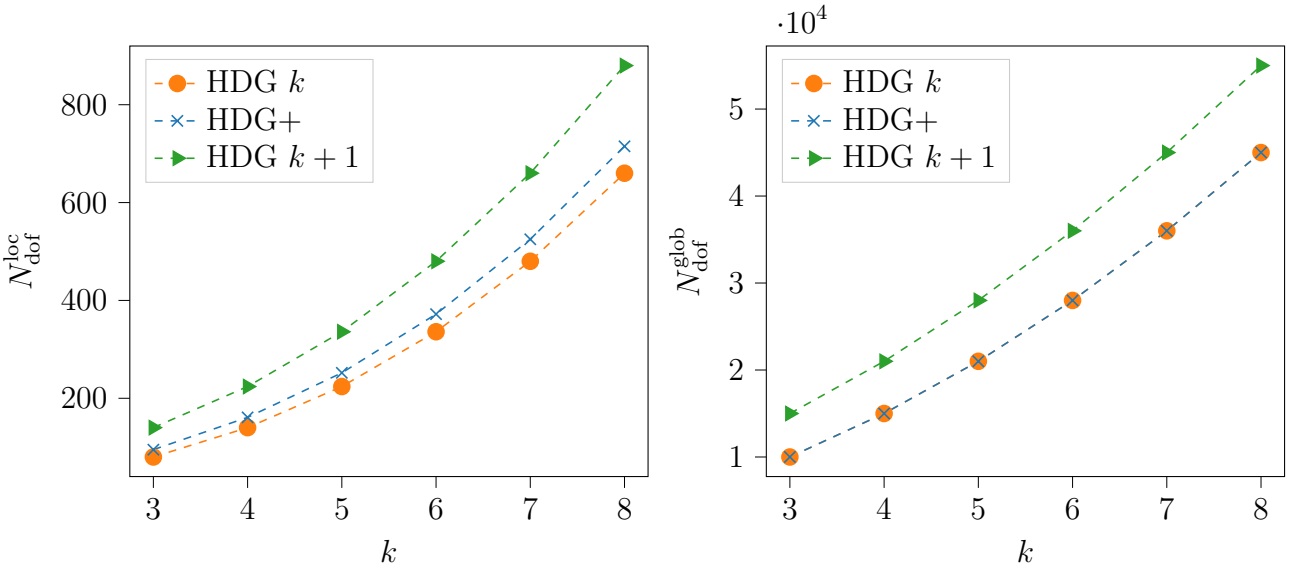
5.5 Comparison of the cost of the HDG and HDG+ methods

Now that we have written down the discrete systems for the HDG and HDG+ methods, it is interesting to compare the sizes of systems that we need to solve.
In Table 4 we have written down the sizes of the matrices for both methods. We can see that the global problem $\mathbb{K}$ of the HDG+ has the same dimension as the one of the HDG method of degree k. As the resolution of the global problem is the most expensive step in the resolution of the method, the cost of the HDG+ method is therefore similar to the cost of the HDG method of degree k, while yielding to an order of convergence of $k+2$ instead of $k+1$. We can also notice that the cost of the local problems of the HDG+ method is intermediate between the cost of the local problems of the HDG methods of degree k and $k+1$. However this has a really limited impact on the computational cost.
We denote by N_{elt} the number of elements in the mesh, *ie* $N_{\text{elt}} = \text{card}(\mathcal{T}_h)$.

Matrix	HDG k	HDG+	HDG $k+1$
$\mathbb{A}^K$	$(4d_n(k))^2$	$(d_n(k+1)+3d_n(k))^2$	$(4d_n(k+1))^2$
$\mathbb{C}^K$	$4d_n(k) \times 4d_{n-1}(k)$	$(d_n(k+1)+3d_n(k)) \times (4d_{n-1}(k))$	$4d_n(k+1) \times 4d_{n-1}(k+1)$
$\mathbb{B}^K$	$4d_{n-1}(k) \times 4d_n(k)$	$4d_{n-1}(k) \times (d_n(k+1)+3d_n(k))$	$4d_{n-1}(k+1) \times 4d_n(k+1)$
$\mathbb{L}^K$	$(4d_{n-1}(k))^2$	$(4d_{n-1}(k))^2$	$(4d_{n-1}(k+1))^2$
$\mathbb{K}$	$\sim (N_{\text{elt}} d_{n-1}(k))^2$	$\sim (N_{\text{elt}} d_{n-1}(k))^2$	$\sim (N_{\text{elt}} d_{n-1}(k+1))^2$

Table 4: Size of the matrices for the HDG and HDG+ methods in 3D

In Figure 5, we have plotted the number of degrees of freedom for local problems ($N_{\text{dof}}^{\text{loc}}$) and for the global problem ($N_{\text{dof}}^{\text{glob}}$) with $N_{\text{elt}} = 10^3$ for several polynomial degrees in 3D. Notice that $\mathbb{A}^K$ has dimension $(N_{\text{dof}}^{\text{loc}})^2$ and $\mathbb{K}$ has dimension $(N_{\text{dof}}^{\text{glob}})^2$.



(a) Local number of degrees of freedom

(b) Global number of degrees of freedom with $N_{\text{elt}} = 10^3$

Figure 5: Local and global number of degrees of freedom for the three HDG methods in 3D

To put this different costs in perspective, we can also compare them to the cost of a standard DG method. Using Table 4, we know that the number of degrees of freedom of the HDG

method of degree k is

$$N_{\text{dof}}^{\text{HDG-}k} \sim N_{\text{elt}} d_{n-1}(k) = N_{\text{elt}} \binom{k+n-1}{n-1} \sim k^{n-1} N_{\text{elt}},$$

whereas the number of degrees of freedom of a DG method of degree k is

$$N_{\text{dof}}^{\text{DG-}k} \sim N_{\text{elt}} d_n(k) = N_{\text{elt}} \binom{k+n}{n} \sim k^n N_{\text{elt}},$$

which becomes much larger as k increases. We therefore retrieved the fact that HDG in dimension n has the same cost as DG in dimension $n-1$, this is depicted in FIGURE 6.

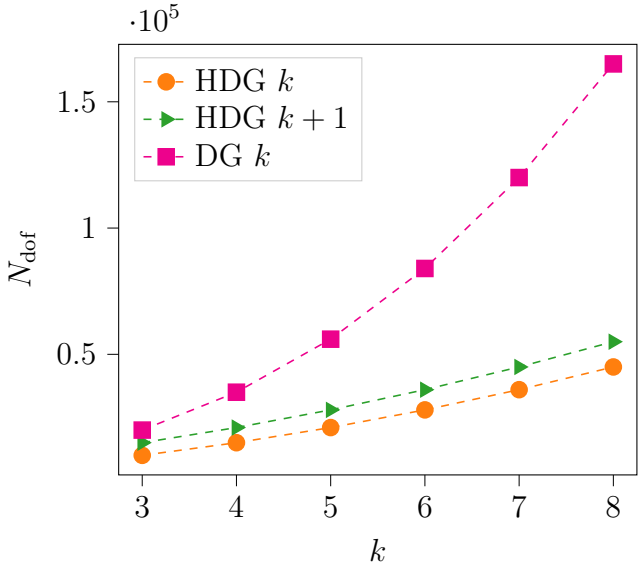


Figure 6: $N_{\text{dof}}^{\text{HDG-}k}$, $N_{\text{dof}}^{\text{HDG-}k+1}$ and $N_{\text{dof}}^{\text{DG-}k}$ in 3D with $N_{\text{elt}} = 10^3$

6 Numerical experiments

In this section we will provide some numerical experiments for the three HDG methods that were described and analysed in this paper. We will first focus on the simple case of duct modes propagating in a waveguide to obtain convergence curves and validate the theoretical results of the previous sections. We will then provide some illustrative examples to show that those methods can be used in more realistic cases.

6.1 Convergence rate

In this subsection, we will present some numerical experiments to illustrate our theoretical results. As most of the estimates obtained in our analysis involve projection errors of the form

$$\|p_h - \pi_W p\|_{\mathcal{T}_h} \quad \text{or} \quad \|\boldsymbol{q}_h - \boldsymbol{\pi}_{\boldsymbol{V}} \boldsymbol{q}\|_{\mathcal{T}_h},$$

we will need to evaluate those projections before actually computing errors. In TABLE 5 we recall the different projections used for the analysis of the three different variants of the HDG method that we considered.

Method	π_W	$\boldsymbol{\pi_V}$	P_M
HDG $\boldsymbol{q}$	L^2	L^2	L^2
HDG +	L^2	L^2	L^2
HDG $\boldsymbol{\sigma}$	HDG	HDG	L^2

Table 5: Summary of the different projections used for the analysis of the HDG methods

Geometric settings: As depicted on FIGURE 7 we consider a uniform directional flow $\boldsymbol{v_0} = Mc_0\boldsymbol{e}_x$, where M is the *Mach number*.

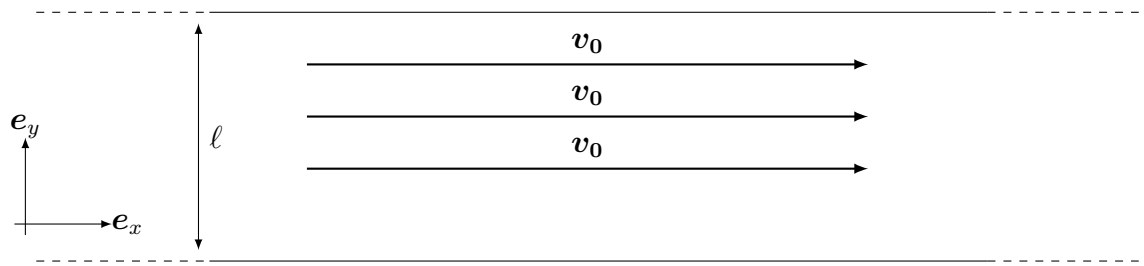


Figure 7: Sketch of the geometric configuration

Unless stated otherwise, we will always use the following parameters for the convergence tests

$$\mathcal{O} = (0,2)\times(0,1) \quad ; \quad \rho_0, c_0 \equiv 1 \quad ; \quad \omega = 5.55\pi,$$

and the choice of M will be specified for each numerical experiment.

Analytic solution: The duct modes are a family of analytic solutions of (1) in a waveguide, see [BDL03]. They are given by

$$p_n^{\pm}(x,y) = e^{i\beta_n^{\pm}x}\varphi_n(y)$$

where

$$n < N_0 : \qquad \beta_n^{\pm} = \frac{-\kappa M \pm \sqrt{\kappa^2 - \frac{n^2\pi^2}{\ell^2}(1-M^2)}}{1-M^2}$$

$$n > N_0 : \qquad \beta_n^{\pm} = \frac{-\kappa M \pm i\sqrt{\frac{n^2\pi^2}{\ell^2}(1-M^2) - \kappa^2}}{1-M^2}$$

with

$$\kappa = \frac{\omega}{c_0} \quad \text{and} \quad M = \frac{v_0}{c_0}$$

$$N_0 = \left\lfloor \frac{\kappa\ell}{\pi\sqrt{1-M^2}} \right\rfloor$$

and

$$\varphi_0(y) := \sqrt{\ell^{-1}}$$
$$\varphi_n(y) := \sqrt{2\ell^{-1}}\cos\left(\frac{n\pi y}{\ell}\right), \quad n \in \mathbb{N}^*$$

The choice of n will be specified for each numerical experiment.

Evaluating the projections: Here we give the details for evaluating the L^2 projection onto W_h, the process is very similar for the other projections.
We recall that the dimension of the polynomial spaces is given by

$$\dim \mathcal{P}_k(K) = \binom{n+k}{n} =: d_n(k), \quad \text{for } K \subset \mathbb{R}^n.$$

All the projections considered are local to an element K, so evaluating them amounts to solving a linear system on each K. Indeed, the definition of the L^2 projection onto $W_h(K) = \mathcal{P}_k(K)$ gives

$$\forall i \in [\![1, d_n(k)]\!], \quad \left(\pi_W p, \Phi_i^{K,k}\right)_K = \left(p, \Phi_i^{K,k}\right)_K.$$

As $\pi_W p \in W_h(K)$, we can write

$$\pi_W p = \sum_{j=1}^{d_n(k)} \pi_j \Phi_j^{K,k},$$

where $\underline{\pi} = (\pi_j)_j$ is the vector of the coordinates of $\pi_W p$ in the basis $(\Phi_j^K)_j$ of $\mathcal{P}_k(K)$. We therefore obtain the following system

$$\forall i \in [\![1, d_n(k)]\!], \quad \sum_{j=1}^{d_n(k)} \pi_j \left(\Phi_j^{K,k}, \Phi_i^{K,k}\right)_K = \left(p, \Phi_i^{K,k}\right)_K,$$

or in matrix form

$$\mathbb{M}^{K,k}\underline{\pi} = \underline{s}, \quad \text{where } s_i = \left(p, \Phi_i^{K,k}\right)_K.$$

The integral in the right-hand side is evaluated using a 91 points Gauss-Lobatto quadrature formula and the linear system can be solved using `lapack`.

Evaluating the L^2-error: Now that the projections can be evaluated, it remains to compute the L^2 norms. As the numerical solution and the projections are polynomial quantities, this can be done using the mass matrix. Indeed, for $u \in \mathcal{P}_k(K)$ we have

$$\begin{aligned}
\|u\|_K^2 &= \left(\sum_{j=1}^{d_n(k)} u_j \Phi_j^{K,k}, \sum_{i=1}^{d_n(k)} u_i \Phi_i^{K,k}\right)_K \\
&= \sum_{i,j=1}^{d_n(k)} u_j u_i^* \left(\Phi_j^{K,k}, \Phi_i^{K,k}\right)_K \\
&= \sum_{i=1}^{d_n(k)} u_i^* \sum_{j=1}^{d_n(k)} \mathbb{M}_{ij}^{K,k} u_j \\
&= \underline{u}^* \mathbb{M}^{K,k} \underline{u}.
\end{aligned}$$

To obtain more meaningful results, we will use relative errors instead of the standard L^2-error. This choice allows us to compare the errors computed on different meshes without any pollution coming from the different number of elements. The relative error for p_h is given by

$$\mathcal{E}_p := \frac{\|p_h - \pi_W p\|_{\mathcal{T}_h}}{\|\pi_W p\|_{\mathcal{T}_h}},$$

and similar expressions will be used for the other volumetric quantities.

Some notation: To allow the comparison between the HDG+ method and the two HDG methods, we would like to emphasize that k always denotes the polynomial degree used to approximate the *trace unknown* $\widehat{p_h}$. We have chosen to plot the relative errors against the quantity $\frac{k}{h}$ which is proportional to the number of degrees of freedom per wavelength. We would like to point out that all the plots in the next sections will use a log-log scale.

Hardware configuration: Those numerical experiments were carried out on a `miriel` node on the `plafrim` cluster[8]. This node is equipped with a 2 dodeca-core Haswell Intel Xeon E5-2680 v3 with a clock rate of 2.5 GHz and 128 Go of memory.

6.1.1 Acoustic case without flow

A first important test to validate our implementation in the *acoustic case without flow*. In this case, the convected Helmholtz equation reduces to the standard Helmholtz equation and the HDG methods, either with the diffusive or total flux, are the same when there is no convection. We should therefore be able to reproduce the *super-convergence* of the HDG method for the Helmholtz equation , *ie.*

$$\mathcal{E}_p = \mathcal{O}(h^{k+2}),$$

when an approximation of degree k is used for all the unknowns.
Here we have chosen to use the following parameters

$$n = 0, \quad \text{and} \quad M = 0,$$

which correspond to a plane-wave.
The resulting convergence curve is depicted on FIGURE 8 and we can clearly see that the super-convergence is obtained.

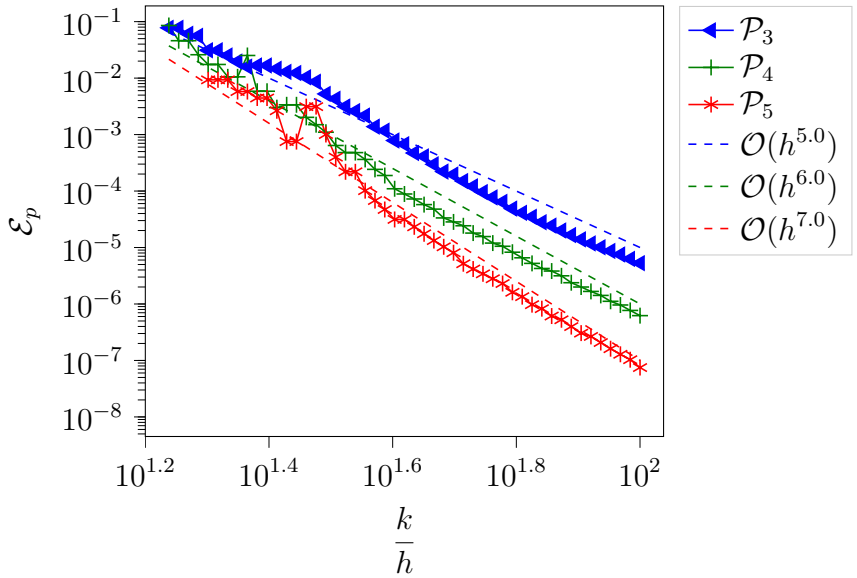


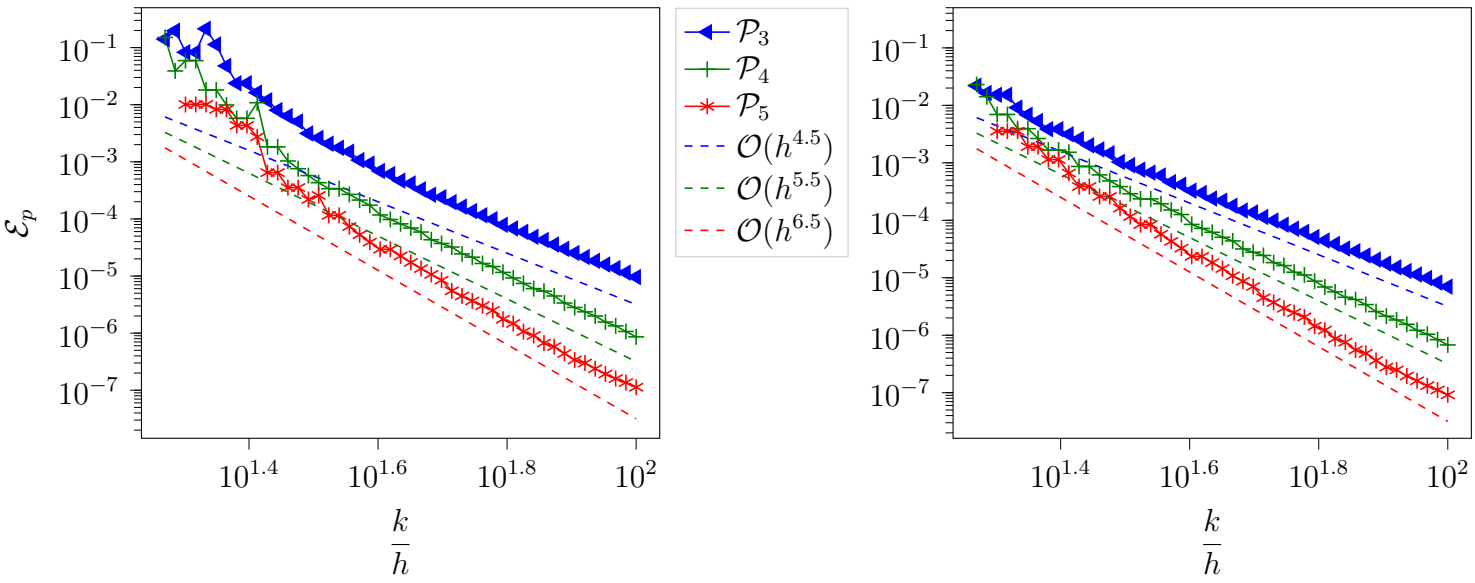
Figure 8: Convergence history for the HDG method without flow for the volumetric unknown p_h

[8] See `http://www.plafrim.fr`.

6.1.2 Low Mach

We then move to a flow with a low Mach number. In this case we have used the following parameters

$$n = 3, \quad \text{and} \quad M = 0.2.$$



(a) HDG with diffusive flux $\boldsymbol{q}_h$

(b) HDG with total flux $\boldsymbol{\sigma}_h$

Figure 9: Low Mach convergence history for the volumetric unknown p_h for both the diffusive and total flux HDG method

The convergence history for the volumetric unknown p_h is displayed on FIGURE 9 for both of the HDG methods. We can see that the *diffusive flux* formulation achieves an order of convergence of $k + 3/2$ as expected. On the other hand the *total flux* formulation also achieves an order of convergence of $k + 3/2$ which is better than the expected order $k + 1$. However for uniform flows and upon well choosing the penalization parameters both of those methods are algebraically equivalent and we therefore expect to obtain the same order of convergence for p_h. On FIGURE 10 the convergence rate for the volumetric unknown p_h for the HDG+ method is displayed. As expected the optimal convergence rate of $k + 2$ is obtained. As discussed in TABLE 6, it is clear that the use of the HDG+ method is less expensive than the use of the HDG methods.

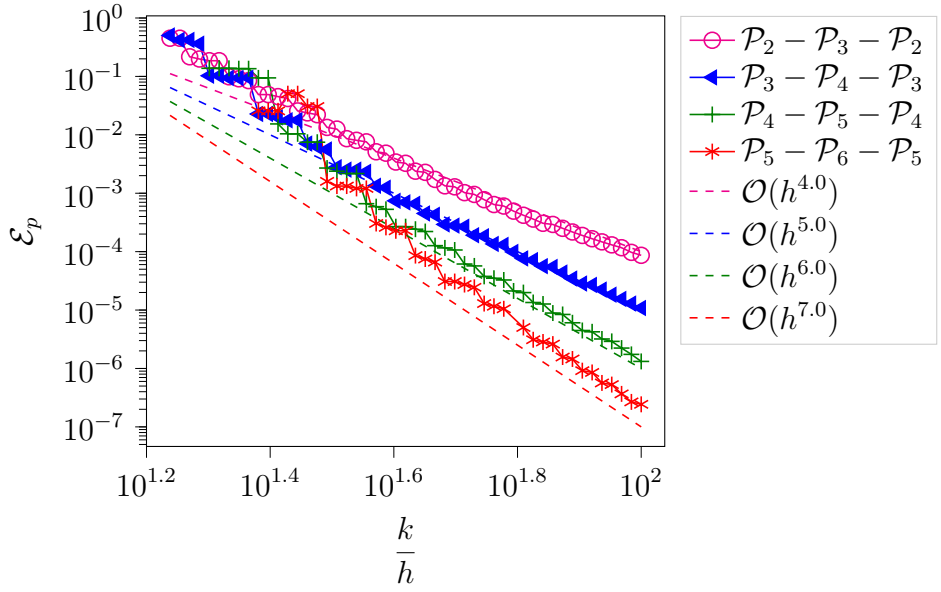


Figure 10: Low Mach convergence history for the volumetric unknown p_h for the HDG+ method

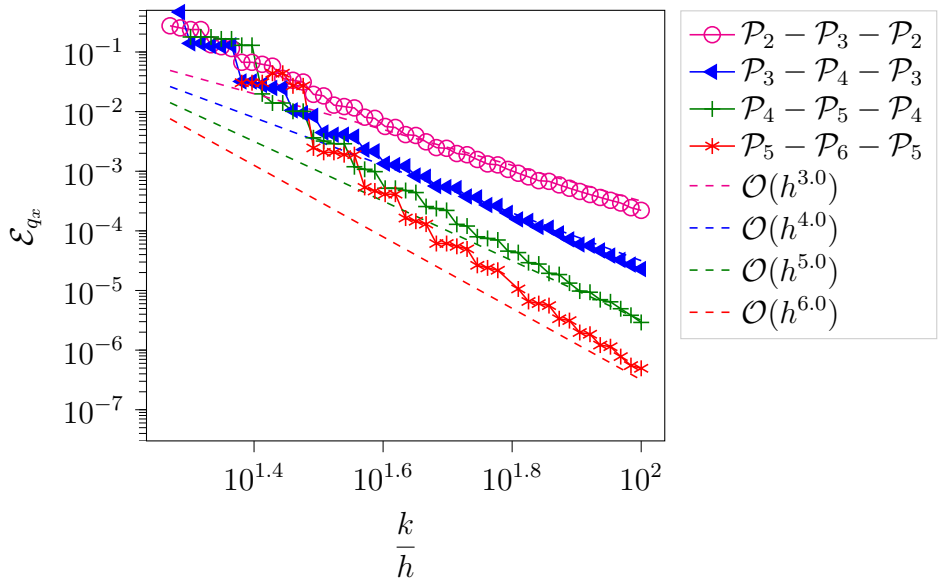
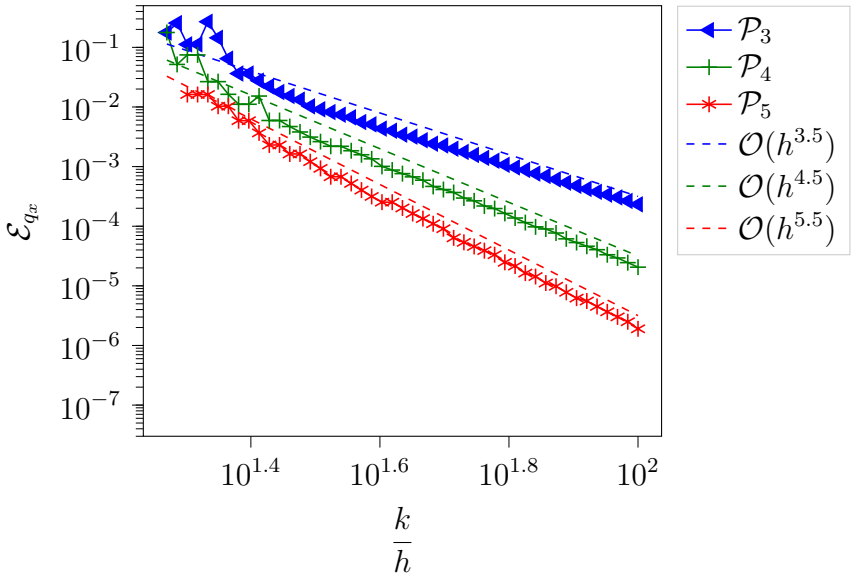
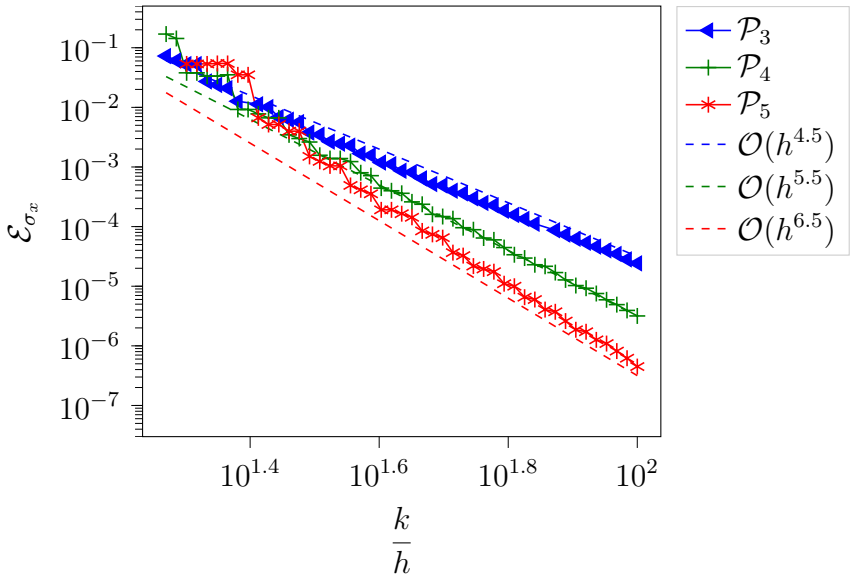


Figure 11: Low Mach convergence history for the first component of volumetric unknown $\boldsymbol{q}_h$ for the HDG+ method



(a) HDG with diffusive flux $\boldsymbol{q}_h$



(b) HDG with total flux $\boldsymbol{\sigma}_h$

Figure 12: Low Mach convergence history for the first component of the volumetric flux unknown

If we now move to the flux unknown, we can see on FIGURE 11 that $\boldsymbol{q}_h$ converges with the optimal order $k+1$ as expected. For the HDG methods, the convergence histories for $\boldsymbol{q}_h$ and $\boldsymbol{\sigma}_h$ are depicted on FIGURE 12. The *diffusive flux* formulation achieves a convergence order of $k+1/2$ as expected, whereas the *total flux* formulation outperforms the theoretical results and achieves an order of $k+3/2$. For this formulation it was however expected that both p_h and $\boldsymbol{\sigma}_h$ converge with the same order, which is actually the case. Finally we would like to point out that the choice between the formulations with $\boldsymbol{q}_h$ or $\boldsymbol{\sigma}_h$ is application dependent.
Finally information regarding the size of the global problems for a fixed number of degrees of freedom per wavelength are given in TABLE 6. We have included the HDG+ method with $(k, k-1)$ and the HDG+ method with $(k+1, k)$. When using the HDG+ method with $(k, k-1)$, a smaller global problem is solved to obtain the same convergence rate as the HDG methods. When using the HDG+ method with $(k+1, k)$, larger elements can be used to obtain the same number of degrees of freedom per wavelength because of the higher polynomial interpolation degree. We can clearly see that the HDG+ methods are computationally less expensive than

both of the HDG methods.

k	h/k		HDG-$\boldsymbol{q}_h$	HDG-$\boldsymbol{\sigma}_h$	HDG+ $(k, k-1)$	HDG+ $(k+1, k)$
3	10^{-1}	nnz	20 384	20 384	10 968	8 720
		nnz LU	36 872	36 872	15 216	9 800
		MUMPS time	$2.5 \cdot 10^{-2}$	$1.6 \cdot 10^{-2}$	$1.8 \cdot 10^{-2}$	$1.7 \cdot 10^{-2}$
3	10^{-2}	nnz	872 216	872 216	481 528	467 280
		nnz LU	3 572 752	3 572 752	1 835 148	1 673 560
		MUMPS time	0.27	0.23	0.15	0.10
4	10^{-1}	nnz	17 130	17 130	8 720	4 980
		nnz LU	24 125	24 125	9 800	4 644
		MUMPS time	$1.9 \cdot 10^{-2}$	$2.0 \cdot 10^{-2}$	$1.7 \cdot 10^{-2}$	$9.5 \cdot 10^{-3}$
4	10^{-2}	nnz	752 330	752 330	467 280	423 600
		nnz LU	2 874 075	2 874 075	1 673 560	1 397 418
		MUMPS time	0.19	0.17	0.10	$8.6 \cdot 10^{-2}$
5	10^{-1}	nnz	12 552	12 552	4 980	4 332
		nnz LU	14 112	14 112	4 644	4 068
		MUMPS time	$1.7 \cdot 10^{-2}$	$2.7 \cdot 10^{-2}$	$9.5 \cdot 10^{-3}$	$9.9 \cdot 10^{-3}$
5	10^{-2}	nnz	672 840	672 840	432 600	414 456
		nnz LU	2 409 006	2 409 006	1 397 418	1 333 368
		MUMPS time	0.12	0.11	$8.6 \cdot 10^{-2}$	$6.8 \cdot 10^{-2}$

Table 6: Size of the global systems and `mumps` elapsed time for different interpolation degree k for the low Mach case with a fixed number of degrees of freedom per wavelength

Notice that the numerical experiments performed here lead to relatively small linear system. Indeed our goal was to validate the numerical method rather than showing the ability of `hawen` to handle large numerical simulations.

In TABLE 7, we review the size of the global system for various interpolation degrees k with a fixed error threshold for p_h. As we did before, we use the relative L^2-error defined by

$$\mathcal{E}_p := \frac{\|p_h - \pi_W p\|_{\mathcal{T}_h}}{\|\pi_W p\|_{\mathcal{T}_h}}.$$

We can clearly see that, when the desired error is smaller than 10^{-2}, increasing the interpolation degree and thus the order of the method leads to solving a smaller linear system to obtain the same error level. This less visible for an error threshold of 10^{-2}, as this accuracy can be obtained with relatively large elements even with a low interpolation degree. For this error level, we can also see that there is no real difference in the size of the linear systems of HDG and HDG+ methods. This is expected as there is no real difference between the different interpolation degrees for this error threshold. However for smaller error thresholds, we can see that the HDG-$\boldsymbol{q}_h$, HDG-$\boldsymbol{\sigma}_h$ and HDG+ with $(k, k-1)$ lead to linear systems with similar sizes, which is expected as they share the same order. The HDG+ method with $(k+1, k)$ also has a higher convergence rate and therefore lead to significantly smaller linear system for the same error level.

k	$\mathcal{E}_p$		HDG-$\boldsymbol{q}_h$	HDG-$\boldsymbol{\sigma}_h$	HDG+ $(k, k-1)$	HDG+ $(k+1, k)$
3	10^{-2}	nnz	49 784	39 216	37 200	26 920
		nnz LU	116 208	85 416	97 371	53 088
		MUMPS time	$3.1 \cdot 10^{-2}$	$1.8 \cdot 10^{-2}$	$2.5 \cdot 10^{-2}$	$1.3 \cdot 10^{-2}$
3	10^{-3}	nnz	99 920	102 800	92 040	49 784
		nnz LU	293 328	293 328	283 155	116 272
		MUMPS time	$3.0 \cdot 10^{-2}$	$2.8 \cdot 10^{-2}$	$3.2 \cdot 10^{-2}$	$1.9 \cdot 10^{-2}$
3	10^{-4}	nnz	210 360	267 784	255 552	140 640
		nnz LU	689 568	926 352	944 826	440 584
		MUMPS time	$4.9 \cdot 10^{-2}$	$6.2 \cdot 10^{-2}$	$8.1 \cdot 10^{-2}$	$3.63 \cdot 10^{-2}$
4	10^{-2}	nnz	28 240	28 240	26 920	28 240
		nnz LU	51 050	51 150	53 088	51 050
		MUMPS time	$2.3 \cdot 10^{-2}$	$1.4 \cdot 10^{-2}$	$1.3 \cdot 10^{-2}$	$1.4 \cdot 10^{-2}$
4	10^{-3}	nnz	58 560	58 560	49 784	42 050
		nnz LU	126 350	126 450	116 272	84 075
		MUMPS time	$1.9 \cdot 10^{-2}$	$2.0 \cdot 10^{-2}$	$1.9 \cdot 10^{-2}$	$2.2 \cdot 10^{-2}$
4	10^{-4}	nnz	106 880	151 600	140 640	77 770
		nnz LU	284 050	433 450	440 584	182 875
		MUMPS time	$2.5 \cdot 10^{-2}$	$3.2 \cdot 10^{-2}$	$3.63 \cdot 10^{-2}$	$2.1 \cdot 10^{-2}$
4	10^{-5}	nnz	255 580	255 580	372 640	156 100
		nnz LU	787 925	787 925	1 383 296	446 855
		MUMPS time	$3.9 \cdot 10^{-2}$	$4.1 \cdot 10^{-2}$	0.10	$3.4 \cdot 10^{-2}$
5	10^{-2}	nnz	24 660	24 660	28 240	24 660
		nnz LU	34 740	34 740	51 050	34 740
		MUMPS time	$1.1 \cdot 10^{-2}$	$1.4 \cdot 10^{-2}$	$1.4 \cdot 10^{-2}$	$1.9 \cdot 10^{-2}$
5	10^{-3}	nnz	40 656	40 656	42 050	40 656
		nnz LU	73 440	73 296	84 075	73 512
		MUMPS time	$1.6 \cdot 10^{-2}$	$1.4 \cdot 10^{-2}$	$2.2 \cdot 10^{-2}$	$1.3 \cdot 10^{-2}$
5	10^{-4}	nnz	84 312	88 200	77 770	60 540
		nnz LU	181 944	192 996	182 875	120 636
		MUMPS time	$2.6 \cdot 10^{-2}$	$1.9 \cdot 10^{-2}$	$2.1 \cdot 10^{-2}$	$1.5 \cdot 10^{-2}$
5	10^{-5}	nnz	148 704	153 888	156 100	111 972
		nnz LU	392 580	409 752	446 855	263 340
		MUMPS time	$3.1 \cdot 10^{-2}$	$2.9 \cdot 10^{-2}$	$3.4 \cdot 10^{-2}$	$2.5 \cdot 10^{-2}$

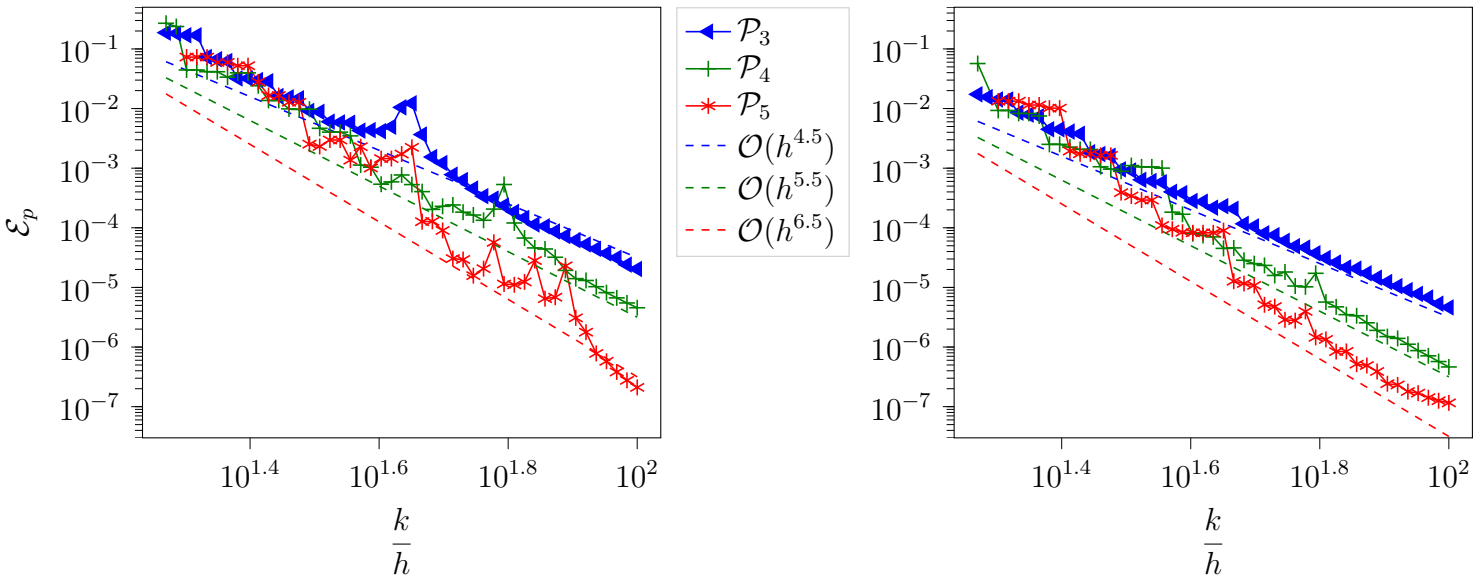
Table 7: Size of the global systems and `mumps` elapsed time for different interpolation degree k for the low Mach case with a fixed error threshold $\mathcal{E}_p$

6.1.3 Large Mach

Finally we also considered a flow with a large March number. In this case, we used the following parameters

$$n = 3, \quad \text{and} \quad M = 0.8.$$

As the simulations of acoustic wave propagation in flows with large Mach numbers is known to be more challenging, we expect to see worse performances than in the previous subsection.



(a) HDG with diffusive flux $\boldsymbol{q}$

(b) HDG with total flux $\boldsymbol{\sigma}$

Figure 13: Large Mach convergence history for the volumetric unknown p_h for both of the HDG methods

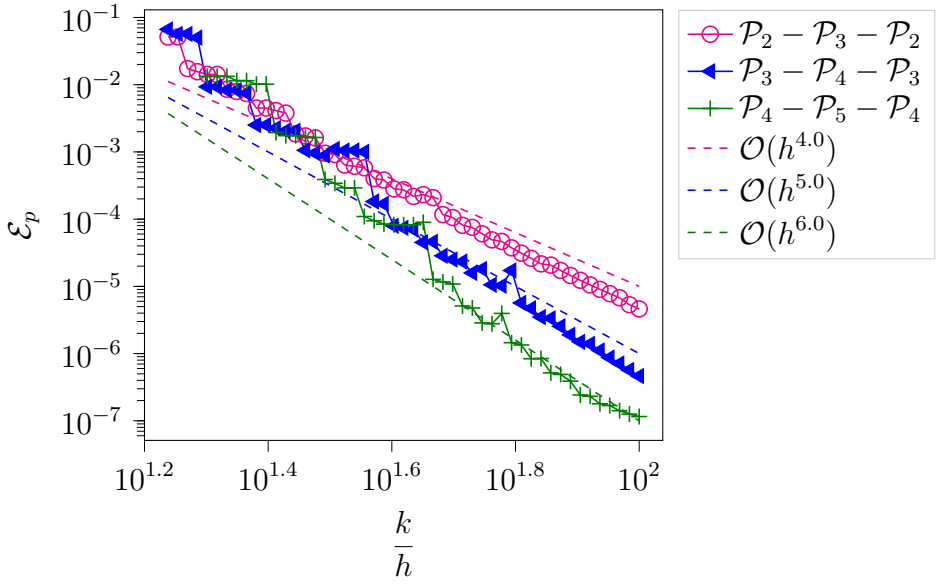
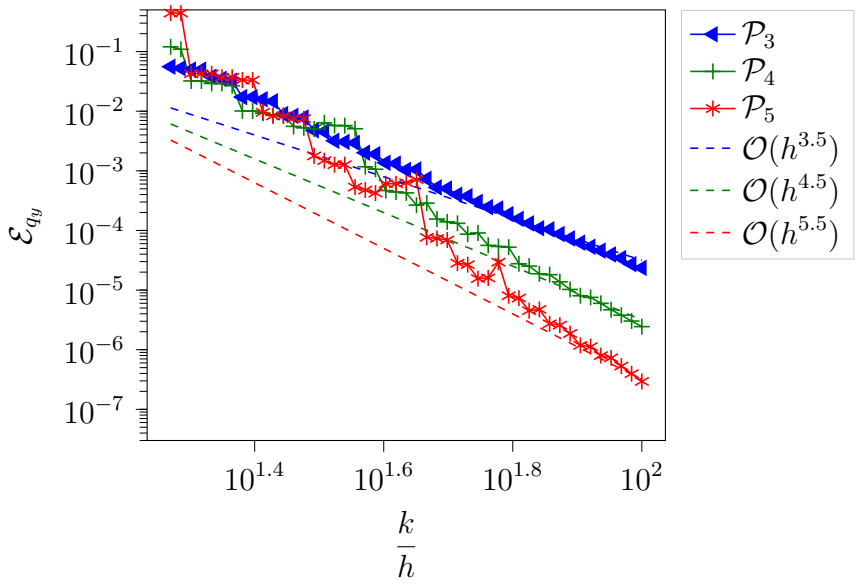
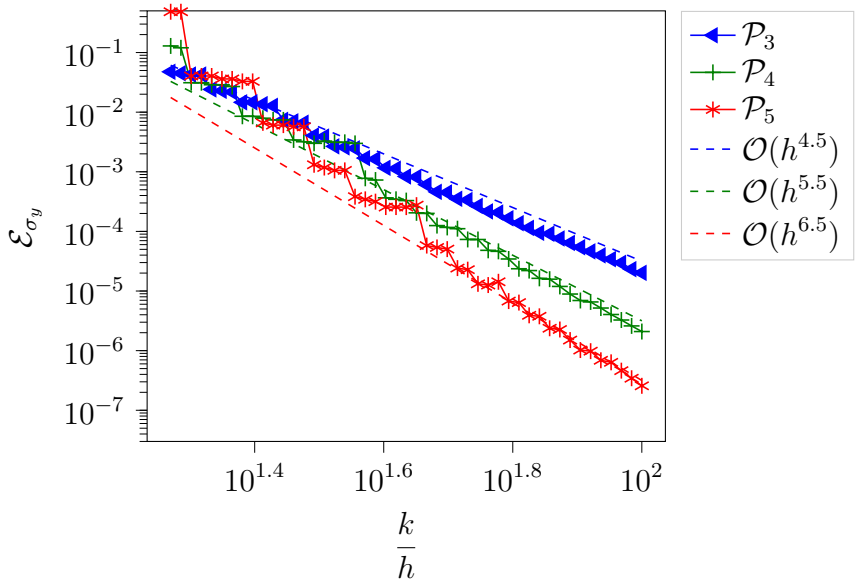


Figure 14: Large Mach convergence history for the volumetric unknown p_h for the HDG+ method

The convergence history for the volumetric unknown is depicted in FIGURE 13 for both of the HDG methods. We can see that the *total flux* formulation still achieves the convergence order of $k + 3/2$ whereas the behaviour of the *diffusive flux* formulation seems less robust. The same convergence history for HDG+ method is displayed on FIGURE 14 and we can see that it still achieves the optimal convergence rate of $k + 2$.



(a) HDG with diffusive flux $\boldsymbol{q}_h$



(b) HDG with total flux $\boldsymbol{\sigma}_h$

Figure 15: Large Mach convergence history for the second component of the volumetric flux unknown

The convergence history for the volumetric flux unknown $\boldsymbol{q}_h$ or $\boldsymbol{\sigma}_h$ is depicted in FIGURE 15 for the HDG methods and in FIGURE 16 for the HDG+ method. As in the low-Mach case, the HDG methods have a convergence rate of $k + 3/2$ and the HDG+ method has a convergence rate of $k + 1$. Notice that the HDG-$\boldsymbol{\sigma}_h$ method seems to be the most robust method for the approximation of the flux unknown for high Mach numbers.

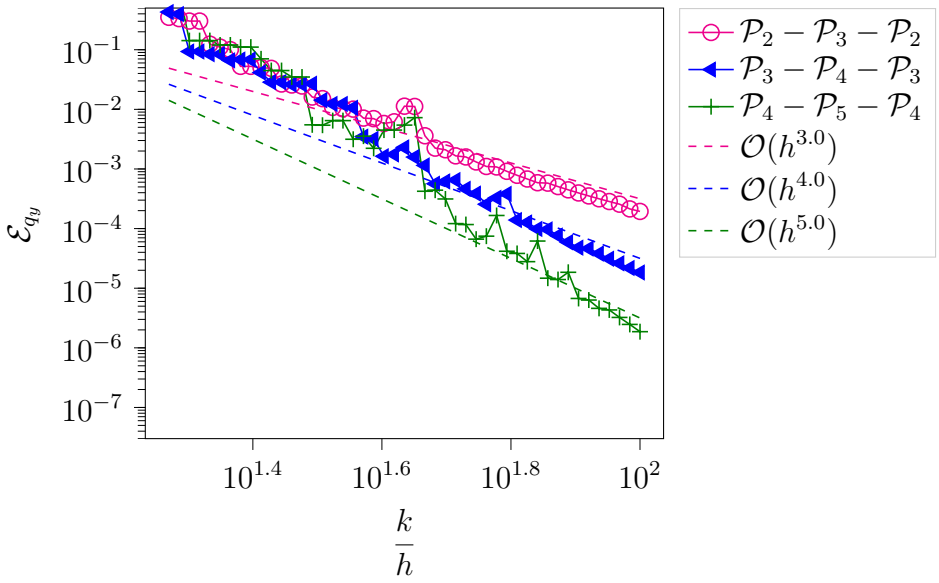


Figure 16: Large Mach convergence history for the second component of the volumetric unknown $\boldsymbol{q}_h$ for the HDG+ method

6.2 A posteriori error estimate

In this section, we will show that it is possible to compute a simple *a posteriori* error indicator. For more complete approach to *a posteriori* error analysis for HDG methods, we refer to [CZ12, CZ13]. We introduce the *relative jump error*

$$\mathcal{E}_{\text{jump}} = \frac{\sqrt{\sum_{K,\ell} \|\widehat{p_h} - p_h\|^2_{\partial K^\ell}}}{\sqrt{\sum_{K,\ell} \|\widehat{p_h}\|^2_{\partial K^\ell}}},$$

which measures the jump between p_h and $\widehat{p_h}$. For the HDG+ method, p_h should be replaced with $P_M p_h$.

Jump error and residuals for the HDG methods: To understand the importance of this quantity, we introduce the residuals

$$\delta_h := -\omega^2 \rho_0 p_h + \operatorname{div}(\boldsymbol{\sigma}_h) - s, \quad \text{and} \quad \boldsymbol{\Delta}_h := \boldsymbol{W_0}\boldsymbol{\sigma}_h + \nabla p_h + 2i\omega p_h \boldsymbol{W_0}\boldsymbol{b_0}.$$

We have chosen to work with the HDG-$\boldsymbol{\sigma}_h$ formulation as we recommend it over the HDG-$\boldsymbol{q}_h$ one, but the adaptation to the HDG-$\boldsymbol{q}_h$ formulation is immediate.
Reverting the integrations by parts in (12a)–(12b), we have

$$\begin{aligned}(\boldsymbol{\Delta}_h, \boldsymbol{r}_h)_K &= \langle p_h - \widehat{p_h}, \boldsymbol{r}_h \cdot \boldsymbol{n}\rangle_{\partial K}, \\ (\delta_h, w_h)_K &= -i\omega \langle \tau(p_h - \widehat{p_h}), w_h\rangle_{\partial K},\end{aligned}$$

as δ_h and $\boldsymbol{\Delta}_h$ are usually not polynomial quantities this means that

$$\begin{aligned}(\boldsymbol{P_V}\boldsymbol{\Delta}_h, \boldsymbol{r}_h)_K &= \langle p_h - \widehat{p_h}, \boldsymbol{r}_h \cdot \boldsymbol{n}\rangle_{\partial K}, \\ (P_W\delta_h, w_h)_K &= -i\omega \langle \tau(p_h - \widehat{p_h}), w_h\rangle_{\partial K},\end{aligned}$$

where $\boldsymbol{P_V}$ and P_W are the L^2-orthogonal projections onto $\boldsymbol{V}_h(K)$ and $W_h(K)$ respectively. Notice that if δ_h and $\boldsymbol{\Delta}_h$ are actually polynomials things are even simpler as $P_W\delta_h = \delta_h$ and $\boldsymbol{P_V}\boldsymbol{\Delta}_h = \boldsymbol{\Delta}_h$. Taking $\boldsymbol{r}_h = \boldsymbol{P_V}\boldsymbol{\Delta}_h$ and $w_h = P_W\delta_h$, and using the following inverse inequality

$$\|w_h\|_{\partial K} \leqslant C h_K^{-\frac{1}{2}} \|w_h\|_K, \quad \forall w_h \in W_h,$$

leads to

$$\|\boldsymbol{P_V}\boldsymbol{\Delta}_h\|_K \leqslant C_{\boldsymbol{\Delta}} h_K^{-\frac{1}{2}} \|p_h - \widehat{p_h}\|_{\partial K}\,,$$
$$\|P_W\delta_h\|_K \leqslant C_\delta h_K^{-\frac{1}{2}} \|p_h - \widehat{p_h}\|_{\partial K}\,.$$

As

$$\|\delta_h\|_K \leqslant \|P_W\delta_h\|_K + \|(\mathrm{Id} - P_W)\delta_h\|_K\,, \tag{97}$$

we can see that the size of the residuals and hence the quality of the approximation depends only on the size of the jump $p_h - \widehat{p_h}$ (first term of the rhs) and the approximation properties of $W_h(K)$ (second term of the rhs).
When $p \in H^s(K)$ with $s \in [\![0, k+1]\!]$, we have

$$\|p - P_W p\|_K \leqslant C h_K^s \|p\|_{s,K}\,,$$

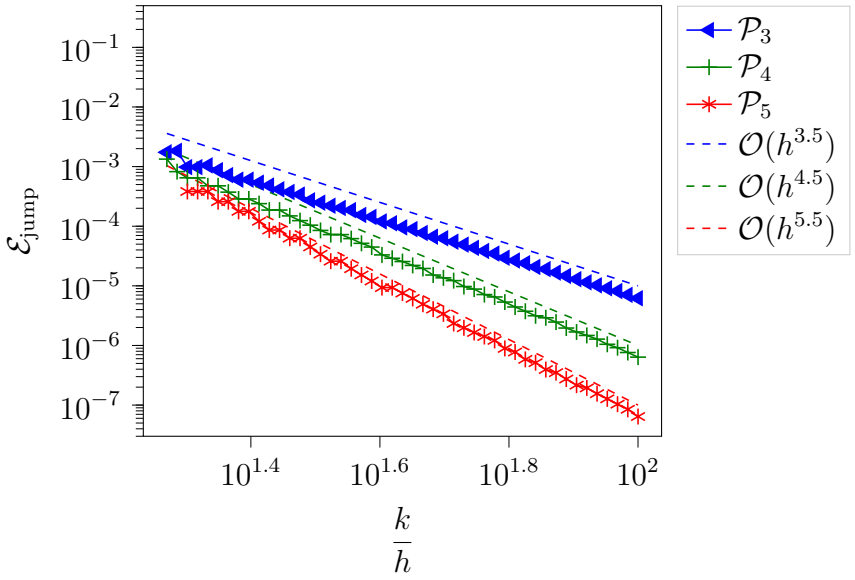
so the penalization term $\tau(p_h - \widehat{p_h})$ will ensure the stability of the method. Indeed (97) will therefore lead to

$$\|\delta_h\|_K \leqslant C_\delta h_K^{-\frac{1}{2}} \|p_h - \widehat{p_h}\|_{\partial K} + C h_K^s \|\delta_h\|_{s,K}\,,$$

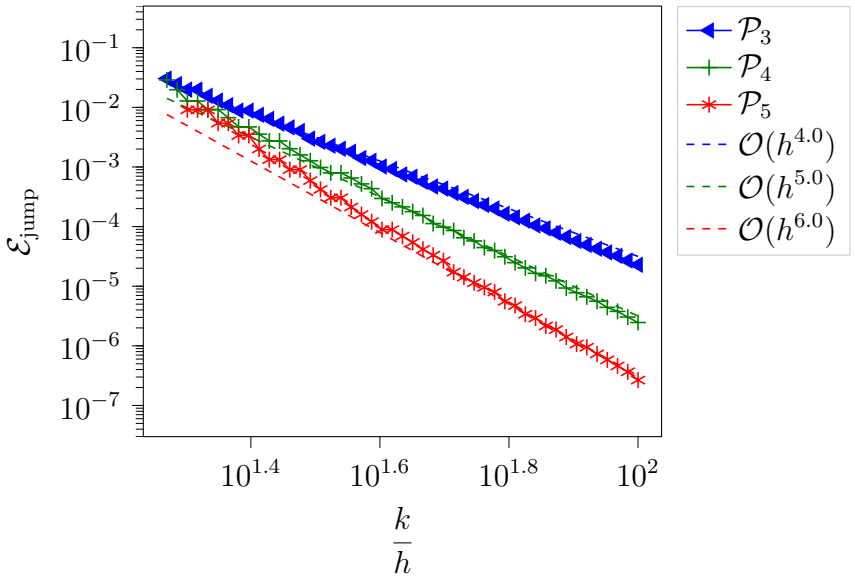
and assuming that h_K is small enough, we have

$$\|\delta_h\|_K \leqslant C h_K^{-\frac{1}{2}} \|p_h - \widehat{p_h}\|_{\partial K}\,,$$

as the last term of the right-hand side can be absorbed by the left-hand side. A similar result can be obtained for $\boldsymbol{\Delta}_h$.
To illustrate this, we have depicted $\mathcal{E}_{\text{jump}}$ for the HDG methods for a low-Mach flow in Figure 17 and for a large-Mach flow in Figure 18.

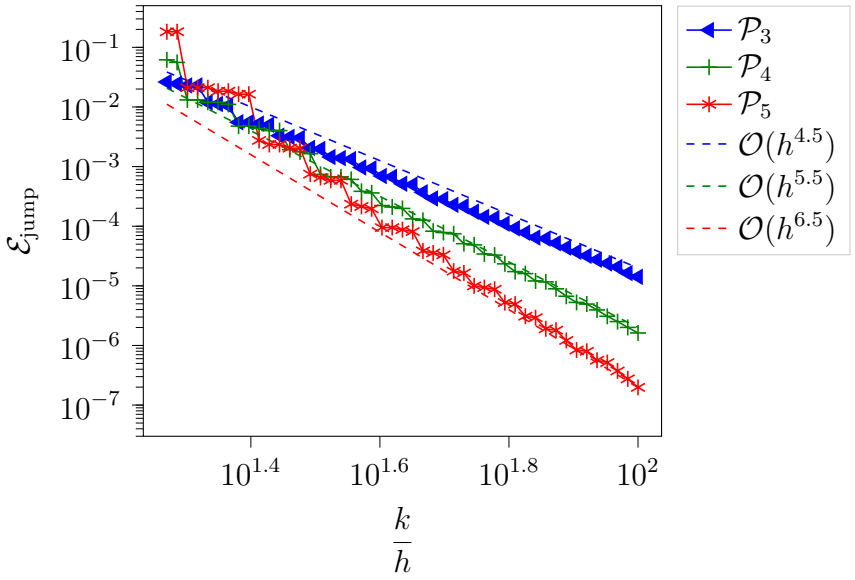


(a) HDG with diffusive flux $\boldsymbol{q}_h$

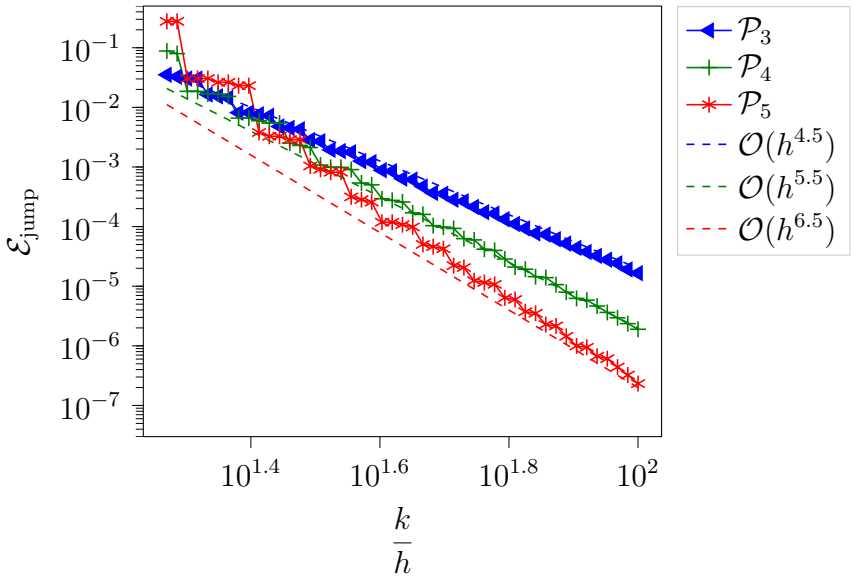


(b) HDG with total flux $\boldsymbol{\sigma}_h$

Figure 17: Low Mach convergence history for the jump error



(a) HDG with diffusive flux $\boldsymbol{q}_h$



(b) HDG with total flux $\boldsymbol{\sigma}_h$

Figure 18: Large Mach convergence history for the jump error

Adaptation to the HDG+ method: For the HDG+ method, nothing changes for the vectorial residual $\boldsymbol{\Delta}_h$ but things are a little different for the scalar residual δ_h. Indeed this time we have

$$\begin{aligned} \|P_W \delta_h\|_K &\lesssim h_K^{-\frac{1}{2}} \left(\| |\tau|(P_M p_h - \widehat{p_h}) \|_{\partial K} + \| |\tau_{\text{upw}}|(p_h - \widehat{p_h} \|_{\partial K} \right), \\ &\lesssim h_K^{-\frac{1}{2}} \| |\tau|(P_M p_h - \widehat{p_h}) \|_{\partial K}, \\ (\text{as } \tau = \mathcal{O}({h_K}^{-1})) \quad &\lesssim h_K^{-\frac{3}{2}} \| P_M p_h - \widehat{p_h} \|_{\partial K}. \end{aligned}$$

To illustrate this, we have depicted $\mathcal{E}_{\text{jump}}$ for the HDG+ method for a low-Mach flow in FIGURE 19 and for a large-Mach flow in FIGURE 20. Once again, we can see that these quantities are clearly decreasing.

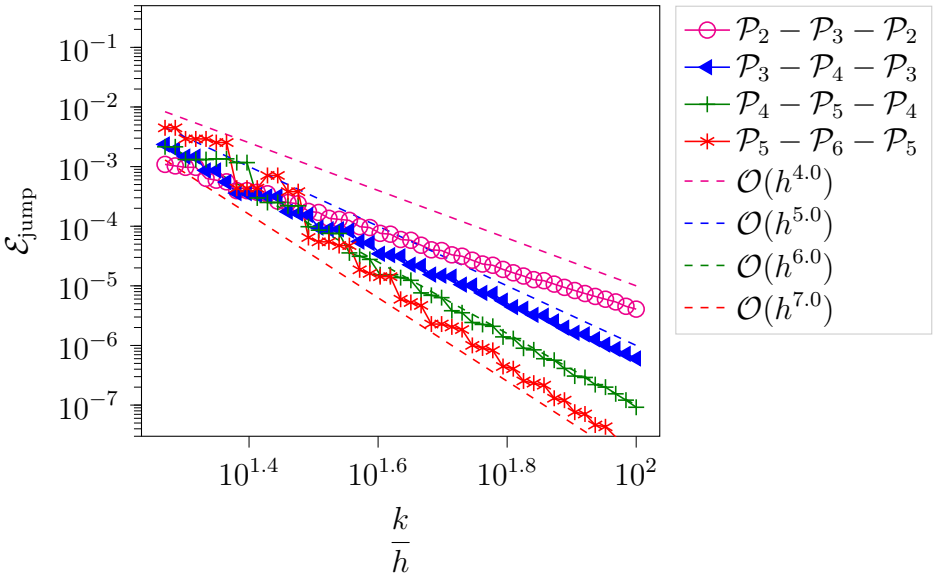


Figure 19: Low Mach convergence history for the jump error for the HDG+ method

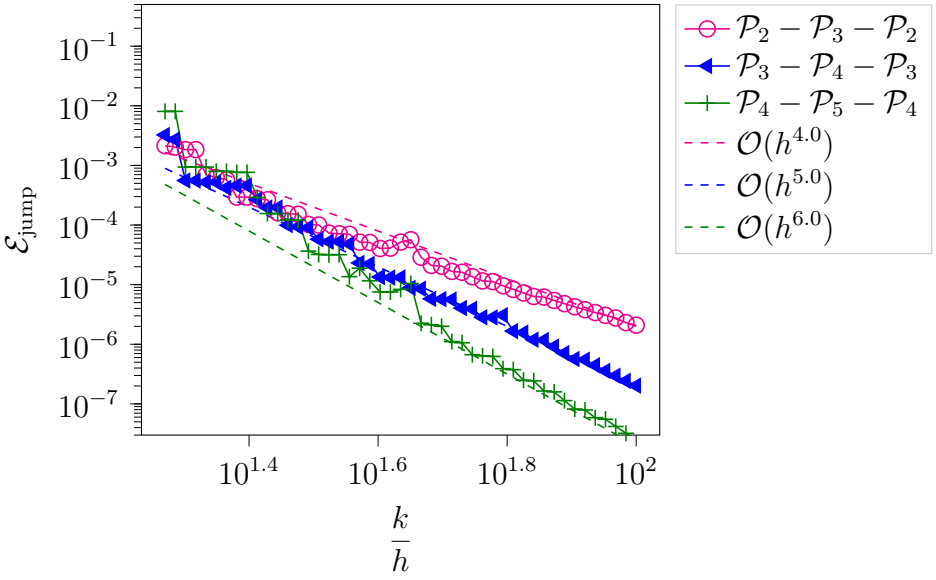


Figure 20: Large Mach convergence history for the jump error for the HDG+ method

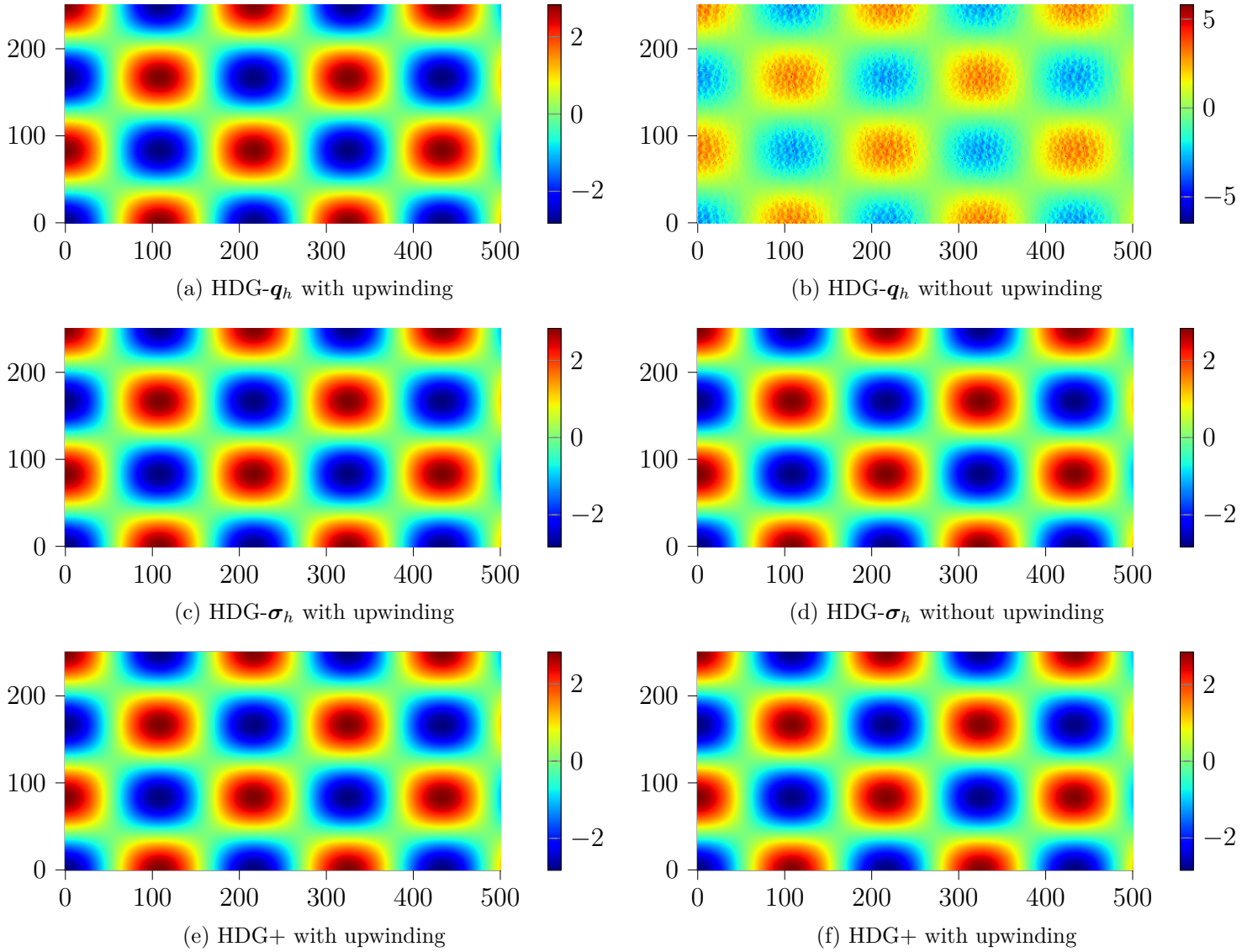
6.3 Is the upwinding mechanism necessary ?

In this section we investigate the numerical impact of upwinding mechanisms. We have performed the same numerical simulations with the three methods with and without those mechanisms. For the HDG-$\boldsymbol{q}_h$ and the HDG+ methods, deactivating the upwinding mechanism corresponds to taking $\tau_{\text{upw}} \equiv 0$. For the HDG-$\boldsymbol{\sigma}_h$ it corresponds to taking $\tau = 1$ instead of the value given by the Riemann solver.
As it can be seen on FIGURE 21 the HDG-$\boldsymbol{q}_h$ without upwinding leads to poor numerical results, whereas the two other methods seem to perform well.

Method	With upwinding	Without upwinding
HDG-$\boldsymbol{q}_h$	$4.3 \cdot 10^{-7}$	0.53
HDG-$\boldsymbol{\sigma}_h$	$4.9 \cdot 10^{-7}$	$5.8 \cdot 10^{-7}$
HDG+	$6.3 \cdot 10^{-8}$	$5.77 \cdot 10^{-7}$

Table 8: Jump error $\mathcal{E}_{\text{jump}}$ for the different method

We have also computed the jump error $\mathcal{E}_{\text{jump}}$ as quality indicator of the numerical solution. The values are given in TABLE 8. Even if the error is higher without upwinding for the HDG+ and HDG-$\boldsymbol{\sigma}_h$, it seems to remain at a reasonable level. For the HDG+ method, we can conclude that the τ_{upw} penalization seems optional. This can be understood as we need to choose $\tau = \mathcal{O}({h_K}^{-1})$ which is large, so the method seems less sensitive to changes in the penalization. On the other hand, this τ_{upw} penalization is mandatory to make the HDG-$\boldsymbol{q}_h$ formulation work. Finally, for the HDG-$\boldsymbol{\sigma}_h$ we still recommend to use the upwind penalization parameter τ as it leads to a method with no arbitrary choice to make.



(a) HDG-$\boldsymbol{q}_h$ with upwinding

(b) HDG-$\boldsymbol{q}_h$ without upwinding

(c) HDG-$\boldsymbol{\sigma}_h$ with upwinding

(d) HDG-$\boldsymbol{\sigma}_h$ without upwinding

(e) HDG+ with upwinding

(f) HDG+ with upwinding

Figure 21: p_h for the different HDG methods with and without upwind

6.4 Point-sources in a uniform flow

For many practical applications it is necessary to consider point-sources. In this section, we show that our method can be used for such computations.

Settings: We consider a uniform flow in an infinite plane. We use a Dirac point source $s = \delta_{(0,0)}$ located at the origin. We write $\boldsymbol{b_0}$ as

$$\boldsymbol{b_0} = M\left(\cos\alpha \boldsymbol{e}_x + \sin\alpha \boldsymbol{e}_y\right),$$

where M is the Mach number, α the angle between $\boldsymbol{b_0}$ and the horizontal axis, and we normalize $\rho_0, c_0 \equiv 1$.
The physical domain $\mathcal{O}_{\text{phys}}$ is surrounded with PMLs, see FIGURE 22 for a sketch of the geometric settings.

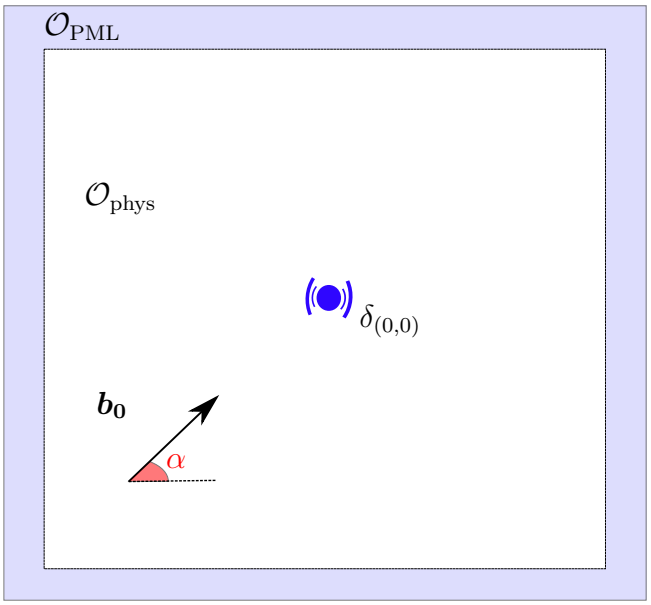
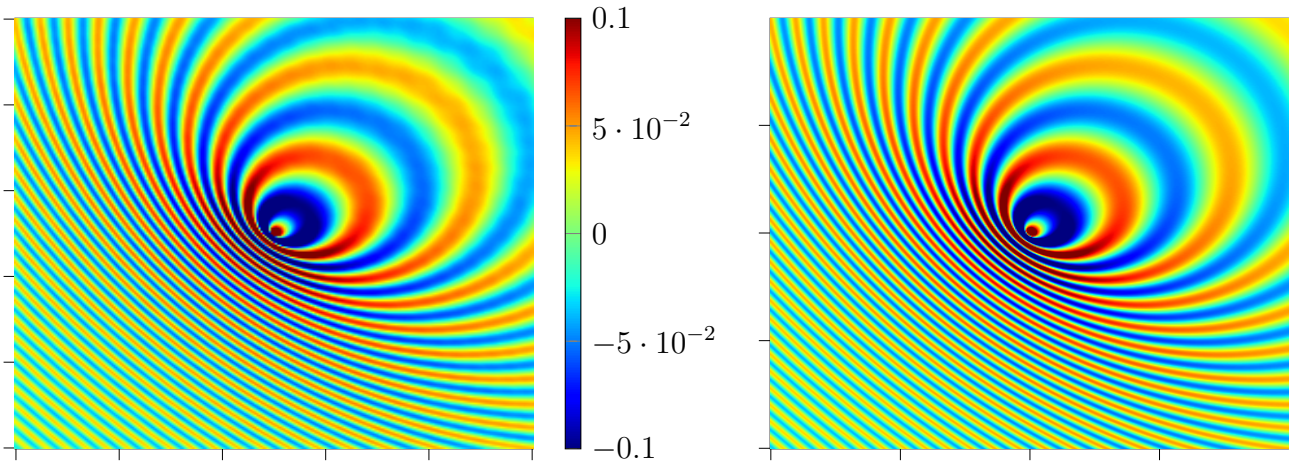


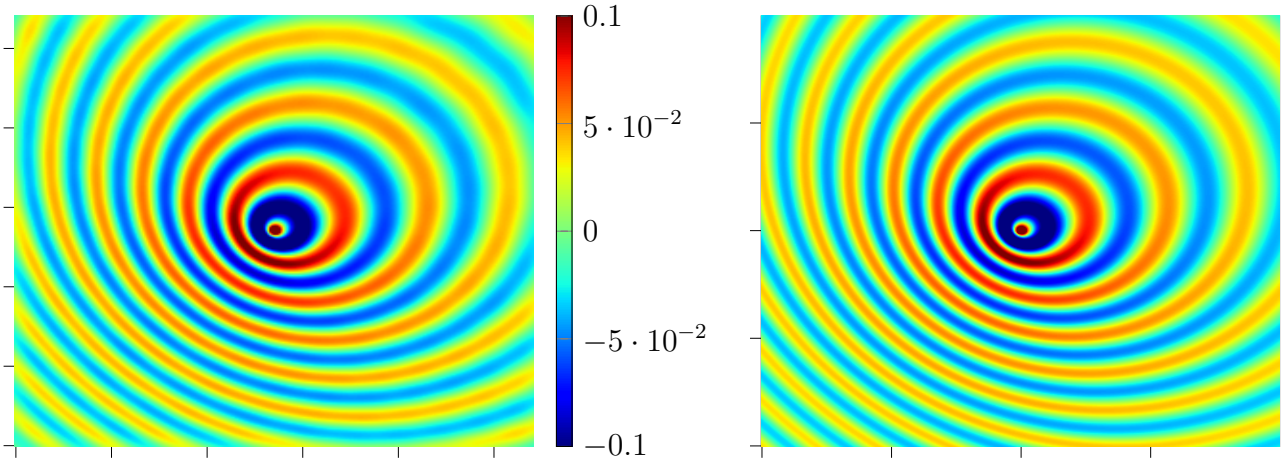
Figure 22: Geometric settings

We have depicted the solution for a large Mach number ($M = 0.8$) in FIGURE 23 and for a low Mach number ($M = 0.4$) in FIGURE 24. The presence of convection leads to a clearly visible Doppler effect : indeed because of the convection, the apparent frequency changes in the domain. The apparent frequency is higher in the bottom-left part of the domain than in the top-right part. The artifacts in the top-right part of the domain are due to the PMLs. Notice that the artifacts are more visible for the large Mach flow.



(a) Numerical result (b) Analytic solution

Figure 23: p_h computed with the HDG-$\boldsymbol{\sigma}_h$ method for $\omega = 6\pi$, $M = 0.8$ and $\alpha = \pi/4$. PMLs are not displayed. The colorbar is the same for the two pictures.



(a) Numerical result (b) Analytic solution

Figure 24: p_h computed with the HDG-$\boldsymbol{\sigma}_h$ method for $\omega = 6\pi$, $M = 0.4$ and $\alpha = \pi/4$. PMLs are not displayed. The colorbar is the same for the two pictures.

Analytic solution: In this context it is possible to write an analytic solution for (1). Following [HPN19], we first need to introduce the so-called *Prandlt-Glauert-Lorentz transformation* in the frequency domain. This transformation maps $\boldsymbol{x} = (x, y)$ to $\widetilde{\boldsymbol{x}} = (\widetilde{x}, \widetilde{y})$ and ω to $\widetilde{\omega}$, where

$$\begin{cases} \widetilde{x} = \left(1 + M^2 \dfrac{\cos^2\alpha}{\beta(1+\beta)}\right) x + M^2 \dfrac{\cos\alpha \sin\alpha}{\beta(1+\beta)} y \\ \widetilde{y} = M^2 \dfrac{\cos\alpha \sin\alpha}{\beta(1+\beta)} x + \left(1 + M^2 \dfrac{\sin^2\alpha}{\beta(1+\beta)}\right) y \\ \widetilde{\omega} = \dfrac{\omega}{\beta} \end{cases},$$

with $\beta := \sqrt{1-M^2}$. As a shorthand, we write this transformation

$$\widetilde{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x}, \quad \text{with } \boldsymbol{A} := \begin{bmatrix} \left(1 + M^2 \dfrac{\cos^2\alpha}{\beta(1+\beta)}\right) & M^2 \dfrac{\cos\alpha\sin\alpha}{\beta(1+\beta)} \\ M^2 \dfrac{\cos\alpha\sin\alpha}{\beta(1+\beta)} & \left(1 + M^2 \dfrac{\sin^2\alpha}{\beta(1+\beta)}\right) \end{bmatrix}.$$

This Prandtl-Glauert-Lorentz transformation is closely related to the Lorentz transform arising in special relativity. It is well-known that, when the background coefficients are uniform, these transformation maps the convected Helmholtz equation to a standard Helmholtz equation, see *eg.* [MBAG20, HPN19]. This can be understood as the Lorentz transformation was introduced as the transformation between two inertial frames that preserves wave equations, for a deeper insight of the connection between flow acoustics and Lorentzian geometry we refer to [Vis98]. The analytic solution is given by

$$p_{\text{exact}}(\boldsymbol{x},\omega) = -\frac{i}{4\beta} H_0^{(1)}\left(\frac{\omega}{\beta}|\boldsymbol{A}\boldsymbol{x}|\right) \exp\left[\frac{i\omega}{\beta}\boldsymbol{A}\boldsymbol{x}\cdot\boldsymbol{b_0}\right]$$

where $H_0^{(1)}$ is the Hankel function of the first kind of order 0 and

$$\widetilde{r} := \sqrt{\widetilde{x}^2 + \widetilde{y}^2}.$$

Even if there is an analytic solution in this case, we were not able to obtain meaningful convergence plots due to the bad quality of the PMLs.

Computational cost: In Table 9 we have written down the sizes of the linear systems to solve for the different HDG methods using an interpolation of order 5 for the trace variable. To give a reference, we also added the size of the system obtained when solving the convected Helmholtz equation (1) with a continuous finite element method (CG) with same interpolation degree using the `montjoie` solver[9]. We have also added the size of the system obtained when solving the standard Helmholtz equation with a Local Discontinuous Galerkin method (LDG), which is a first-order DG method, using the `montjoie` solver (as LDG methods are not implemented for the convected Helmholtz equation in `montjoie`).

Method	HDG-$\boldsymbol{q}_h$	HDG-$\boldsymbol{\sigma}_h$	HDG+	CG	LDG
k	5	5	4	5	5
nnz	12 237 750	12 237 750	6 243 750	1 353 750	46 862 808
nnz LU	67 515 161	67 515 161	34 244 845	39 008 186	259 272 979
$\mathcal{E}_{\text{jump}}$	$1.4 \cdot 10^{-3}$	$1.0 \cdot 10^{-3}$	$8.9 \cdot 10^{-4}$		

Table 9: Size of the linear system to solve for the Dirac in a uniform flow

We can see that using HDG instead of LDG leads to significantly smaller linear systems. For the convected Helmholtz equation, it is possible to use the CG method which less expensive than the HDG ones, see [CD16]. However, we would like to point out that the CG method is known to give bad numerical results for more realistic aeroacoustic models such as Galbrun's equation, see [CD18].

[9] `montjoie` is a versatile and well-tested high-order finite element solver. For more informations about the numerical method used to solve the convected Helmholtz equation, see `http://montjoie.gforge.inria.fr/helmholtz.php`.

Local refinement: To handle the point-sources, a mesh with a local refinement around the source should be used, otherwise artifacts could be present in the numerical solution. With the HDG method, those artifacts seem really limited when no local refinement is used, see FIGURE 25.

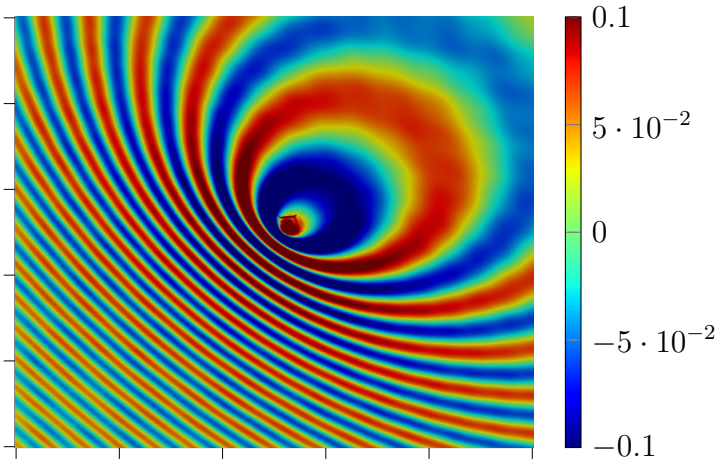


Figure 25: Point-source without local refinement

Illustrative example: To show that our method can handle more complex simulations, we consider the same test-case as before but with two point-sources located near the origin. On FIGURE 26, we can still see the changes in the apparent frequency due to the Doppler effect, but also interference patterns due to the interactions between the two sources.

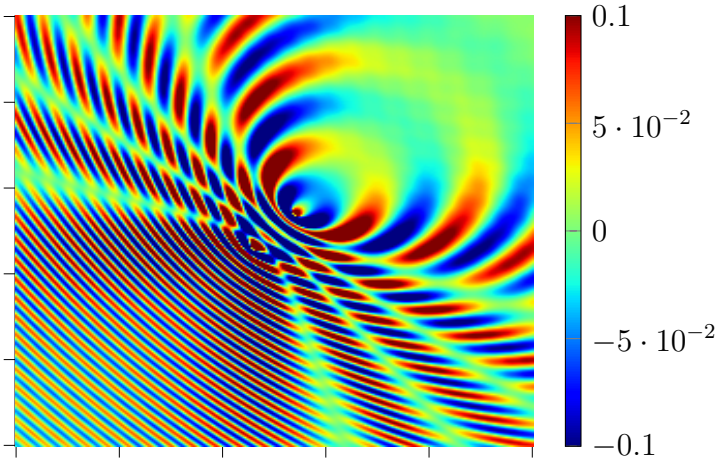


Figure 26: Interference between two point-sources

6.5 Gaussian jet

In this section we focus on a space-varying flow. We will use a *gaussian jet*, which is a common test-case in the literature, see *eg.* [MMMP17, Sec. 3]. To work with parameters and unknowns without dimension, we choose

$$\rho_0 = c_0 \equiv 1,$$

and we consider the following gaussian jet flow

$$\boldsymbol{v_0} = M_0(y)\boldsymbol{e}_x, \quad \text{where} \quad M_0(y) := M_\infty + \mu \exp\left(-\frac{y^2}{R^2}\right),$$

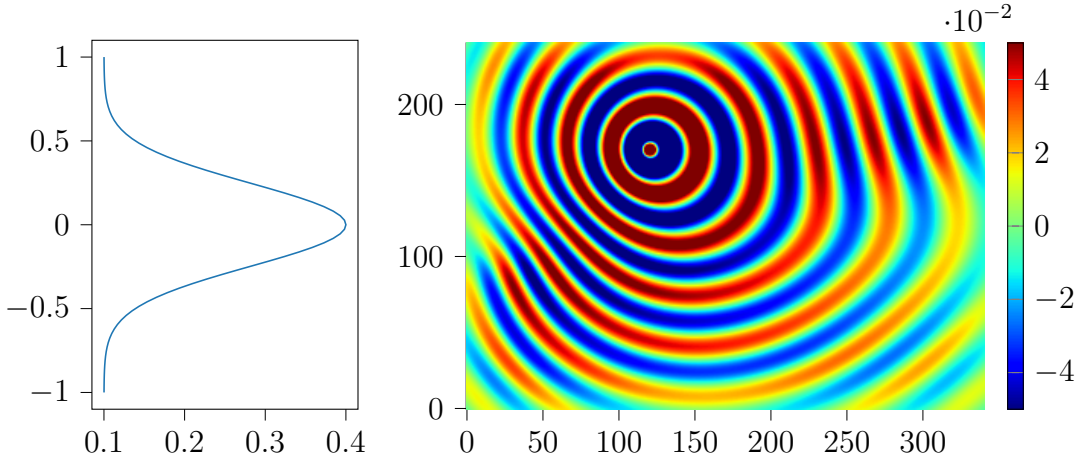
which is a gaussian perturbation of the uniform flow $\boldsymbol{v_0} = M_\infty \boldsymbol{e}_x$.
For this simulation, we work with the following values for the parameters

$$\mathcal{O}_{\text{phys}} = (0,3) \times (-1,1) \quad ; \quad \omega = 6\pi \quad ; \quad s = \delta_{(1,0.5)} \quad ;$$

and

$$M_\infty = 0.1 \quad ; \quad \mu = 0.3 \quad ; \quad R = 0.35.$$

The physical domain $\mathcal{O}_{\text{phys}}$ is surrounded by PMLs.
Notice that these kind of jet-flows are not potential flows and should therefore not be used with the convected Helmholtz equation. However, as discussed in [MMMP17], with this choice of parameters the vorticity of the flow stays small and can be neglected. The convected Helmholtz equation is therefore a good approximation of the more realistic models.



(a) $M_0(y)$ (b) Numerical solution

Figure 27: $\mathfrak{Re} p_h$ for the gaussian jet, obtained with the HDG+ method

On FIGURE 27, we can see that there is a phase-shift inside the jet, as it is expected : this refraction-like effect can be seen in the center-left and center-right parts of the domain at the limit between the jet and uniform flows. There is also a small Doppler effect as the apparent frequency changes in the domain : it is different in the the bottom-left part of the domain and in the center-right one. However this effect is less obvious than in the previous example as the Mach number is significantly lower.

Conclusion

In this paper, we have introduced three HDG methods to solve the convected Helmholtz equation. Two of them are standard HDG methods that use the same polynomial degree for the approximation of all the unknowns. The third one is less standard and uses a higher polynomial degree for the scalar unknown and a reduced stabilization process.

For all of those methods, detailed theoretical results on convergence and well-posedness are provided. It is important to note that we could not obtain the super-convergence property for the two standard HDG methods because of convection. We also provided numerical experiments that are consistent with the absence of super-convergence. The HDG+ method achieves optimal convergence. Due to the reduced stabilization this leads to a "super-convergence like behaviour" as we obtain a convergence rate of $k+2$ for the cost of a HDG method of degree k without post-processing.

During the numerical experiments, it occurred to us that the HDG+ and HDG-$\boldsymbol{\sigma}_h$ methods seemed more robust than the HDG-$\boldsymbol{q}_h$ method. In particular, they are less sensitive to the choice of penalization parameter.

Future work: Now that the use of HDG method has been validated for the convected Helmholtz equation, we will generalize this work to more realistic aeroacoustic models such as Galbrun's equation or Goldstein's equation.

We also noticed that the use of PMLs may lead to bad numerical results, and we will therefore address the construction of absorbing boundary conditions for the convected Helmholtz equation in a future paper.

Open-source implementation: An open-source implementation of those three methods in the `hawen` solver will be released soon.

Acknowledgement: The authors would like to thank Florian Faucher for his help with the numerical implementation.

Experiments presented in this paper were carried out using the PlaFRIM experimental testbed, supported by Inria, CNRS (LABRI and IMB), Université de Bordeaux, Bordeaux INP and Conseil Régional d'Aquitaine (see https://www.plafrim.fr).

Nathan Rouxelin acknowledges financial support from e2s-UPPA (see https://e2s-uppa.eu).

A Intermediate results for the error analysis of the HDG method with diffusive flux

In this appendix we state the intermediate results allowing to adapt the analysis conducted in Subsection 4.3 to prove the convergence of the HDG method stated in Theorem 6. The proofs are omitted but are very similar to the ones given in Subsection 4.3.
The main differences to keep in mind are : $s \in [1, k+1]$ instead of $s \in [1, k+2]$ and $\tau = \mathcal{O}(1)$ instead of $\tau = \mathcal{O}(h^{-1})$.

Gradient estimate: this corresponds to Lemma 4.3

Lemma A.1:
The following estimate holds

$$\|\nabla \varepsilon_h^p\|_{\mathcal{T}_h} \lesssim \|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W}_{\mathbf{0}},\mathcal{T}_h} + \|\boldsymbol{\delta}_h^{\boldsymbol{q}}\|_{\mathcal{T}_h} + h^{-\frac{1}{2}} \|\varepsilon_h^p - \widehat{\varepsilon}_h^p\|_{\partial\mathcal{T}_h}$$

Energy-like equality: this corresponds to Lemma 4.5.

Lemma A.2:
The following energy-like equality holds

$$\begin{aligned}
&\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|^2_{\boldsymbol{W_0},\mathcal{T}_h} - \omega^2 \|\varepsilon_h^p\|^2_{\rho_0,K} - 2i\omega \left(\|\varepsilon_h^p - \widehat{\varepsilon}_h^p\|^2_{\partial\mathcal{T}_h} + \left\| \left(\frac{1}{2}|\boldsymbol{b_0}\cdot\boldsymbol{n}|\right)^{\frac{1}{2}} (\varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\|^2_{\partial\mathcal{T}_h} \right) \\
&= -\omega^2 (\rho_0 \delta_h^p, \varepsilon_h^p)_{\mathcal{T}_h} + 2i\omega (\delta_h^p \boldsymbol{b_0}, \nabla\varepsilon_h^p)_{\mathcal{T}_h} + (\boldsymbol{W_0}\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}, \boldsymbol{\delta}_h^{\boldsymbol{q}})_{\mathcal{T}_h} + 2\omega\mathfrak{Im}\,(\varepsilon_h^p \boldsymbol{b_0}, \nabla\varepsilon_h^p)_{\mathcal{T}_h} \\
&+ \left\langle \boldsymbol{\delta}_h^{\boldsymbol{q}}\cdot\boldsymbol{n} + 2i\omega(\boldsymbol{b_0}\cdot\boldsymbol{n})\widehat{\delta}_h^p - 2i\omega\tau(\delta_h^p - \widehat{\delta}_h^p) + 2i\omega\tau_{\mathrm{upw}}(\delta_h^p - \widehat{\delta}_h^p), \varepsilon_h^p - \widehat{\varepsilon}_h^p \right\rangle_{\partial\mathcal{T}_h}
\end{aligned}$$

Furthermore if $p \in H^s(\mathcal{O})$ and $\boldsymbol{q} \in \boldsymbol{H^t}(\mathcal{O})$ where $s,t \in [1, k+1]$ then the following estimate holds

$$\begin{aligned}
&\left| \|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|^2_{\boldsymbol{W_0},\mathcal{T}_h} - 2i\omega \left(\|\varepsilon_h^p - \widehat{\varepsilon}_h^p\|^2_{\partial\mathcal{T}_h} + \left\| \left(\frac{1}{2}|\boldsymbol{b_0}\cdot\boldsymbol{n}|\right)^{\frac{1}{2}} (\varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\|^2_{\partial\mathcal{T}_h} \right) \right| \\
\lesssim &\omega^2 \|\varepsilon_h^p\|^2_{\mathcal{T}_h} + \omega \|\varepsilon_h^p\|_{\mathcal{T}_h} \left(\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} + h^t \|\boldsymbol{q}\|_{t,\mathcal{O}} + h^{-\frac{1}{2}} \|\varepsilon_h^p - \widehat{\varepsilon}_h^p\|_{\partial\mathcal{T}_h} + \omega h^s \|p\|_{s,\mathcal{O}} \right) \\
&+ \|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} \left(h^t \|\boldsymbol{q}\|_{t,\mathcal{O}} + \omega h^s \|p\|_{s,\mathcal{O}} \right) + h^{2t} \|\boldsymbol{q}\|^2_{t,\mathcal{O}} \\
&+ h^{-\frac{1}{2}} \|\varepsilon_h^p - \widehat{\varepsilon}_h^p\|_{\partial\mathcal{T}_h} \left(\omega h^s \|p\|_{s,\mathcal{O}} + h^t \|\boldsymbol{q}\|_{t,\mathcal{O}} \right)
\end{aligned}$$

Dual identity: this corresponds to Lemma 4.6.

Lemma A.3:
The following dual identity holds

$$\begin{aligned}
\|\varepsilon_h^p\|^2_{\mathcal{T}_h} = &- (\boldsymbol{W_0}\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}, \boldsymbol{\pi_V}\boldsymbol{\xi} - \boldsymbol{\xi})_{\mathcal{T}_h} + \omega^2 (\rho_0 \varepsilon_h^p, \pi_W\theta - \theta)_{\mathcal{T}_h} + 2i\omega (\nabla\varepsilon_h^p, (\pi_W\theta - \theta)\boldsymbol{b_0})_{\mathcal{T}_h} \\
&- 2i\omega \langle (\boldsymbol{b_0}\cdot\boldsymbol{n})\varepsilon_h^p, \pi_W\theta - \theta \rangle_{\partial\mathcal{T}_h} \\
&+ (\boldsymbol{W_0}\boldsymbol{\delta}_h^{\boldsymbol{q}}, \boldsymbol{\pi_V}\boldsymbol{\xi})_{\mathcal{T}_h} - \omega^2 (\rho_0 \delta_h^p, \pi_W\theta)_{\mathcal{T}_h} + 2i\omega (\delta_h^p \boldsymbol{b_0}, \nabla(\pi_W\theta))_{\mathcal{T}_h} \\
&+ 2i\omega \langle (\boldsymbol{b_0}\cdot\boldsymbol{n})\widehat{\varepsilon}_h^p - \tau (\varepsilon_h^p - \widehat{\varepsilon}_h^p) + \tau_{\mathrm{upw}} (\varepsilon_h^p - \widehat{\varepsilon}_h^p), \pi_W\theta - P_M\theta \rangle_{\partial\mathcal{T}_h} \\
&- 2i\omega \left\langle (\boldsymbol{b_0}\cdot\boldsymbol{n})\widehat{\delta}_h^p - \tau(\delta_h^p - \widehat{\delta}_h^p) + \tau_{\mathrm{upw}} \left(\delta_h^p - \widehat{\delta}_h^p\right), \pi_W\theta - P_M\theta \right\rangle_{\partial\mathcal{T}_h} + \langle \boldsymbol{\delta}_h^{\boldsymbol{q}}\cdot\boldsymbol{n}, \pi_W\theta - P_M\theta \rangle_{\partial\mathcal{T}_h}
\end{aligned}$$

Dual estimate: this corresponds to Lemma 4.7.

Lemma A.4:
Assuming that the regularity assumption (72) holds and that $\omega^2 h^2 \|\rho_0\|_\infty C_{\mathrm{reg}} C$ (where C is the constant of Theorem 4) is small enough, if $p \in H^s(\mathcal{O})$ and $\boldsymbol{q} \in \boldsymbol{H^t}(\mathcal{O})$ where $s,t \in [1, k+1]$ then

$$\|\varepsilon_h^p\|_{\mathcal{T}_h} \lesssim h^{t+1}(1+\omega) \|\boldsymbol{q}\|_{t,\mathcal{O}} + h^s(1+\omega+\omega^2) \|p\|_{s,\mathcal{O}} + h(1+\omega) \|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h}$$

where

$$h := \max_{K\in\mathcal{T}_h} h_K$$

References

[ALA13] Natalia C. B. Arruda, Abimael F. D. Loula, and Regina C. Almeida. Locally discontinuous but globally continuous Galerkin methods for elliptic problems. *Com-*

puter Methods in Applied Mechanics and Engineering, 255:104–120, March 2013. Cited on page 9.

[BCDL15] Marie Bonnasse-Gahot, Henri Calandra, Julien Diaz, and Stéphane Lanteri. Hybridizable Discontinuous Galerkin method for the simulation of the propagation of the elastic wave equations in the frequency domain. Research Report RR-8990, INRIA Bordeaux ; INRIA Sophia Antipolis - Méditerranée, June 2015. Cited on pages 4 and 49.

[BDL03] Eliane Bécache, Anne-Sophie Bonnet-Ben Dhia, and Guillaume Legendre. Perfectly matched layers for the convected Helmholtz equation. Technical report, 2003. Cited on page 62.

[BDMP21] Hélène Barucq, Julien Diaz, Rose-Cloé Meyer, and Ha Pham. Implementation of hybridizable discontinuous Galerkin method for time-harmonic anisotropic poroelasticity in two dimensions. *International Journal for Numerical Methods in Engineering*, 122(12):3015–3043, 2021. Cited on page 4.

[CC12] Yanlai Chen and Bernardo Cockburn. Analysis of variable-degree HDG methods for convection–diffusion equations. Part I: General nonconforming meshes. *IMA Journal of Numerical Analysis*, 32(4):1267–1293, October 2012. Cited on page 4.

[CC14] Yanlai Chen and Bernardo Cockburn. Analysis of variable-degree HDG methods for convection-diffusion equations. Part II: Semimatching nonconforming meshes. *Mathematics of Computation*, 83(285):87–111, January 2014. Cited on page 4.

[CD16] Juliette Chabassier and Marc Duruflé. High Order Finite Element Method for solving Convected Helmholtz equation in radial and axisymmetric domains. Application to Helioseismology. Research Report RR-8893, Inria Bordeaux Sud-Ouest, March 2016. Cited on page 80.

[CD18] Juliette Chabassier and Marc Duruflé. Solving time-harmonic Galbrun's equation with an arbitrary flow. Application to Helioseismology. *Rapport de recherche Inria*, 2018. Cited on page 80.

[CDG+09] Bernardo Cockburn, Bo Dong, Johnny Guzmán, Marco Restelli, and Riccardo Sacco. A Hybridizable Discontinuous Galerkin Method for Steady-State Convection-Diffusion-Reaction Problems. *SIAM Journal on Scientific Computing*, 31(5):3827–3846, January 2009. Cited on page 4.

[CDPE16] Bernardo Cockburn, Daniele A. Di Pietro, and Alexandre Ern. Bridging the hybrid high-order and hybridizable discontinuous Galerkin methods. *ESAIM: Mathematical Modelling and Numerical Analysis*, 50(3):635–650, May 2016. Cited on page 4.

[CGL09] Bernardo Cockburn, Jayadeep Gopalakrishnan, and Raytcho Lazarov. Unified Hybridization of Discontinuous Galerkin, Mixed, and Continuous Galerkin Methods for Second Order Elliptic Problems. *SIAM J. Numer. Anal*, 47:1319–1365, August 2009. Cited on page 4.

[CGS10] Bernardo Cockburn, Jayadeep Gopalakrishnan, and Francisco-Javier Sayas. A projection-based error analysis of HDG methods. *Mathematics of Computation*, 79(271):1351–1367, March 2010. Cited on pages 4, 13, 26, and 48.

[CLOS20] Liliana Camargo, Bibiana López-Rodríguez, Mauricio Osorio, and Manuel Solano. An HDG method for Maxwell's equations in heterogeneous media. *Computer Methods in Applied Mechanics and Engineering*, 368:113178, August 2020. Cited on page 4.

[Coc14] Bernardo Cockburn. Static Condensation, Hybridization, and the Devising of the HDG Methods. *springerprofessional.de*, 2014. Cited on page 4.

[CQS18] Huangxin Chen, Weifeng Qiu, and Ke Shi. A priori and computable a posteriori error estimates for an HDG method for the coercive Maxwell equations. *Computer Methods in Applied Mechanics and Engineering*, 333:287–310, May 2018. Cited on page 4.

[CQSS17] Huangxin Chen, Weifeng Qiu, Ke Shi, and Manuel Solano. A Superconvergent HDG Method for the Maxwell Equations. *Journal of Scientific Computing*, 70(3):1010–1029, March 2017. Cited on page 4.

[CS13] Bernardo Cockburn and Ke Shi. Superconvergent HDG methods for linear elasticity with weakly symmetric stresses | IMA Journal of Numerical Analysis | Oxford Academic. *IMA Journal of Numerical Analysis*, 2013. Cited on page 4.

[CZ12] Bernardo Cockburn and Wujun Zhang. A Posteriori Error Estimates for HDG Methods. *Journal of Scientific Computing*, 51(3):582–607, June 2012. Cited on page 72.

[CZ13] Bernardo Cockburn and Wujun Zhang. A Posteriori Error Analysis for Hybridizable Discontinuous Galerkin Methods for Second Order Elliptic Problems. *SIAM Journal on Numerical Analysis*, 51(1):676–693, January 2013. Cited on page 72.

[DS19] Shukai Du and Francisco-Javier Sayas. *An Invitation to the Theory of the Hybridizable Discontinuous Galerkin Method: Projections, Estimates, Tools*. SpringerBriefs in Mathematics. Springer International Publishing, Cham, 2019. Cited on pages 4, 21, 24, 25, 26, 34, and 41.

[Dub91] Moshe Dubiner. Spectral methods on triangles and other domains. *Journal of Scientific Computing*, 6(4):345–390, December 1991. Cited on page 56.

[EG04] Alexandre Ern and Jean-Luc Guermond. *Theory and Practice of Finite Elements*. Applied Mathematical Sciences. Springer-Verlag, New York, 2004. Cited on pages 8, 12, 23, 33, 35, and 41.

[Fau21] Florian Faucher. 'hawen': Time-harmonic wave modeling and inversion using hybridizable discontinuous Galerkin discretization. *Journal of Open Source Software*, 6(57):2699, January 2021. Cited on page 4.

[FCS15] G. Fu, B. Cockburn, and H. Stolarski. Analysis of an HDG method for linear elasticity. *International Journal for Numerical Methods in Engineering*, 102(3-4):551–575, 2015. Cited on page 4.

[FLd14] Cristiane O. Faria, Abimael F. D. Loula, and Antônio J. B. dos Santos. Primal stabilized hybrid and DG finite element methods for the linear elasticity problem. *Computers & Mathematics with Applications*, 68(4):486–507, August 2014. Cited on page 9.

[FS20] Florian Faucher and Otmar Scherzer. Adjoint-state method for Hybridizable Discontinuous Galerkin discretization, application to the inverse acoustic wave problem. *Computer Methods in Applied Mechanics and Engineering*, 372:113406, December 2020. Cited on pages 4 and 49.

[GM11] Roland Griesmaier and Peter Monk. Error Analysis for a Hybridizable Discontinuous Galerkin Method for the Helmholtz Equation. *Journal of Scientific Computing*, 49(3):291–310, December 2011. Cited on page 4.

[GSV18] Jay Gopalakrishnan, Manuel Solano, and Felipe Vargas. Dispersion Analysis of HDG Methods. *Journal of Scientific Computing*, 77(3):1703–1735, December 2018. Cited on page 4.

[HPN19] Fang Q. Hu, Michelle E. Pizzo, and Douglas M. Nark. On the use of a Prandtl-Glauert-Lorentz transformation for acoustic scattering by rigid bodies with a uniform flow. *Journal of Sound and Vibration*, 443:198–211, March 2019. Cited on pages 79 and 80.

[HPS17] Allan Hungria, Daniele Prada, and Francisco-Javier Sayas. HDG methods for elastodynamics. *Computers & Mathematics with Applications*, 74(11):2671–2690, December 2017. Cited on page 4.

[Hun19] Allan Hungria. *Using HDG+ to Compute Solutions of the 3D Linear Elastic and Poroelastic Wave Equations*. PhD thesis, University of Delaware, 2019. Cited on pages 4, 34, and 41.

[HW08] Jan S. Hesthaven and Tim Warburton. *Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications*. Texts in Applied Mathematics. Springer-Verlag, New York, 2008. Cited on page 16.

[Kir04] Robert C. Kirby. Algorithm 839: FIAT, a new paradigm for computing finite element basis functions. *ACM Transactions on Mathematical Software (TOMS)*, 30(4):502–516, December 2004. Cited on page 56.

[KSC12] Robert M. Kirby, Spencer J. Sherwin, and Bernardo Cockburn. To CG or to HDG: A Comparative Study. *Journal of Scientific Computing*, 51(1):183–212, April 2012. Cited on page 4.

[Leh10] Christoph Lehrenfeld. *Hybrid Discontinuous Galerkin Methods for Solving Incompressible Flow Problems*. PhD thesis, RWTH Aachen, 2010. Cited on pages 4 and 28.

[MBAG20] Philippe Marchner, Hadrien Beriot, Xavier Antoine, and Christophe Geuzaine. Stable Perfectly Matched Layers with Lorentz transformation for the convected Helmholtz equation. Technical report, 2020. Cited on page 80.

[MMMP17] Jean-François Mercier, Colin Mietka, Florence Millot, and Vincent Pagneux. Acoustic propagation in a vortical homentropic flow. page 19, 2017. Cited on pages 81 and 82.

[NPRC15] N. C. Nguyen, J. Peraire, F. Reitich, and B. Cockburn. A phase-based hybridizable discontinuous Galerkin method for the numerical solution of the Helmholtz equation. *Journal of Computational Physics*, 290:318–335, June 2015. Cited on page 4.

[Oik14] Issei Oikawa. A hybridized discontinuous Galerkin method with reduced stabilization. *arXiv:1405.2491 [math]*, November 2014. Cited on pages 4 and 56.

[Oik16] Issei Oikawa. Analysis of a Reduced-Order HDG Method for the Stokes Equations. *Journal of Scientific Computing*, 67(2):475–492, May 2016. Cited on page 4.

[Oik18] Issei Oikawa. An HDG Method with Orthogonal Projections in Facet Integrals. *Journal of Scientific Computing*, 76(2):1044–1054, August 2018. Cited on page 4.

[PE12] Daniele Antonio Di Pietro and Alexandre Ern. *Mathematical Aspects of Discontinuous Galerkin Methods*. Mathématiques et Applications. Springer-Verlag, Berlin Heidelberg, 2012. Cited on page 13.

[Pie90] Allan D. Pierce. Wave equation for sound in fluids with unsteady inhomogeneous flow. *The Journal of the Acoustical Society of America*, 87(6):2292–2299, June 1990. Cited on page 5.

[QS16a] Weifeng Qiu and Ke Shi. An HDG Method for Convection Diffusion Equation. *Journal of Scientific Computing*, 66(1):346–357, January 2016. Cited on pages 4, 28, and 41.

[QS16b] Weifeng Qiu and Ke Shi. A superconvergent HDG method for the incompressible Navier–Stokes equations on general polyhedral meshes. *IMA Journal of Numerical Analysis*, 36(4):1943–1967, October 2016. Cited on page 4.

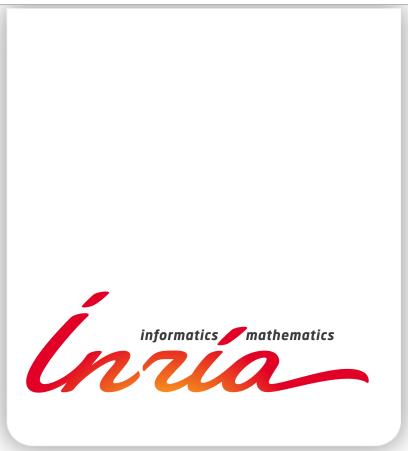
[QSS16] Weifeng Qiu, Jiguang Shen, and Ke Shi. An HDG method for linear elasticity with strong symmetric stresses. *arXiv:1312.1407 [math]*, February 2016. Cited on pages 4 and 41.

[SM50] Jack Sherman and Winifred J. Morrison. Adjustment of an Inverse Matrix Corresponding to a Change in One Element of a Given Matrix. *Annals of Mathematical Statistics*, 21(1):124–127, March 1950. Cited on page 6.

[Ste91] Rolf Stenberg. Postprocessing schemes for some mixed finite elements. *ESAIM: Mathematical Modelling and Numerical Analysis*, 25(1):151–167, 1991. Cited on pages 26 and 48.

[Vis98] Matt Visser. Acoustic black holes: Horizons, ergospheres, and Hawking radiation. *Classical and Quantum Gravity*, 15(6):1767–1791, June 1998. Cited on page 80.

[YMKS16] Sergey Yakovlev, David Moxey, Robert M. Kirby, and Spencer J. Sherwin. To CG or to HDG: A Comparative Study in 3D. *Journal of Scientific Computing*, 67(1):192–220, April 2016. Cited on page 4.



RESEARCH CENTRE
BORDEAUX – SUD-OUEST

200 avenue de la Vieille Tour

33405 Talence Cedex

Publisher
Inria
Domaine de Voluceau - Rocquencourt
BP 105 - 78153 Le Chesnay Cedex
inria.fr

ISSN 0249-6399