HAL open science

A hybridizable discontinuous Galerkin method with characteristic variables for Helmholtz problems

Axel Modave, Théophile Chaumont-Frelet

▶ **To cite this version:**

Axel Modave, Théophile Chaumont-Frelet. A hybridizable discontinuous Galerkin method with characteristic variables for Helmholtz problems. Journal of Computational Physics, 2023, 493, pp.112459. 10.1016/j.jcp.2023.112459 . hal-03909368v2

HAL Id: hal-03909368

https://hal.science/hal-03909368v2

Submitted on 12 Sep 2023

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

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

Distributed under a Creative Commons Attribution 4.0 International License

A hybridizable discontinuous Galerkin method with characteristic variables for Helmholtz problems*

Axel Modave[1] and Théophile Chaumont-Frelet[2]

[1]POEMS, CNRS, Inria, ENSTA Paris, Institut Polytechnique de Paris, 91120 Palaiseau, France, axel.modave@ensta-paris.fr
[2]Université Côte d'Azur, Inria, CNRS, LJAD, 06902 Sophia Antipolis Cedex, France, theophile.chaumont@inria.fr

Abstract

A new hybridizable discontinuous Galerkin method, named the CHDG method, is proposed for solving time-harmonic scalar wave propagation problems. This method relies on a standard discontinuous Galerkin scheme with upwind numerical fluxes and high-order polynomial bases. Auxiliary unknowns corresponding to characteristic variables are defined at the interface between the elements, and the physical fields are eliminated to obtain a reduced system. The reduced system can be written as a fixed-point problem that can be solved with stationary iterative schemes. Numerical results with 2D benchmarks are presented to study the performance of the approach. Compared to the standard HDG approach, the properties of the reduced system are improved with CHDG, which is more suited for iterative solution procedures. The condition number of the reduced system is smaller with CHDG than with the standard HDG method. Iterative solution procedures with CGNR or GMRES required smaller numbers of iterations with CHDG.

1 Introduction

Discontinuous Galerkin (DG) finite element methods have proven their strength to address realistic time-harmonic wave propagation problems, see e.g. [4, 5, 25, 45]. Due to their ability to handle unstructured and possibly non-conforming meshes, they are very versatile and can provide high-fidelity solutions to problems with complicated physical and geometrical configurations. The DG framework also allows for high-order polynomial basis functions, which limits dispersion errors occurring when considering high-frequency problems [1, 2, 46]. Besides, since the degrees of freedom (DOFs) of DG methods are only attached to cells, they can be linearly indexed in memory, which enables efficient implementation on vectorized computer architectures, including GPUs, see e.g. [41, 43, 47].

Despite their manifest advantages, the main bottleneck of DG methods (and more generally, of finite element and finite difference methods) is the numerical solution of the resulting linear system. Indeed, although the matrix is sparse, it is typically large, ill-conditioned, and indefinite, see e.g. [23]. Standard algebraic solvers perform poorly for these systems: direct solvers are prohibitively costly in large 3D applications; iterative solvers require less memory storage and allow direct parallel implementations, but the convergence of the iterative processes can be slow because of intrinsic properties of the time-harmonic wave propagation problems. Although preconditioning strategies have been proposed to speed up the convergence of iterative procedures and to reduce the computational cost, see e.g. [7, 20, 22, 24, 30, 31, 37, 55, 58], the development of fast iterative

*Published in *Journal of Computational Physics* (doi: 10.1016/j.jcp.2023.112459). Distributed under Creative Commons CC-BY 4.0 license.

finite element solvers for high-frequency wave propagation problems remains an active research area.

In this work, we focus on a DG scheme for the Helmholtz equation in first-order form with upwind fluxes, see e.g. [36, 42]. Although this approach is very popular in the time-domain, its direct use for time-harmonic problems is limited, since it involves many coupled DOFs. In order to reduce the computational cost, hybridization strategies have been introduced in the seminal work [16], and largely studied over the past decade, see e.g. [11, 33–35, 38, 44, 51]. In the resulting hybridizable discontinuous Galerkin (HDG) methods, an additional "hybrid variable" corresponding to the Dirichlet trace of the solution is introduced. This additional variable acts as a Lagrange multiplier that decouples the physical unknowns. After inverting element-wise local matrices, a reduced system involving only the Lagrange multiplier is formulated over the skeleton of the mesh. When using a direct linear solver, the advantage of this approach is straightforward, as the reduced HDG system features far less DOFs than the original DG system while preserving its sparsity pattern. On the other hand, the situation is not as clear when considering iterative solvers, since the size and filling of the matrix are no longer the main performance criteria.

Here, we propose a novel hybridization strategy in order to accelerate the solution of the large-scale linear system arising from the upwind DG discretization of time-harmonic problems with iterative procedures. This strategy, which we call the CHDG method, uses the characteristic variables defined at the interface between the elements as the hybrid variables, as opposed to the Dirichlet traces in the standard HDG method. This alternative choice of hybrid variable leads to favorable properties for the resulting reduced system and to more efficient iterative solution procedures in comparison with the standard hybridization strategy. Specifically, the reduced system can be written in the form

$$(\mathrm{I} - \Pi\mathrm{S})g = b, \tag{1}$$

where g corresponds to the characteristic variables, Π is an exchange operator swapping the variables at the interfaces, and S is a scattering operator related to the solution of local element-wise problems. The iteration operator $\Pi\mathrm{S}$ is a strict contraction, so that the system is well-posed and can be solved with a simple fixed-point iteration.

Interestingly, the form of the reduced system (1) closely resembles the ultra-weak variational formulation (UWVF) employed in Trefftz discretizations of time-harmonic problem [9]. In fact, our reduced system inherits many of the favorable properties of UWVF matrices. The advantage of our approach though, is that it simply relies on polynomial basis functions instead of local solutions. As a result, volume right-hand sides and heterogeneous media can be readily considered [36]. Besides, the mesh can be refined, and the discretization order increased without the conditioning issues typically appearing for plane wave basis functions, see e.g. [3, 28, 39, 52, 54]. To avoid these issues, quasi-Trefftz methods with polynomial basis functions are currently investigated, see e.g. [40]. The UWVF has been tested with polynomial basis functions in [27, 49].

The fixed-point system (1) also naturally appears in non-overlapping substructuring domain decomposition (DD) methods. The iteration operator $\Pi\mathrm{S}$ was already used in the seminal work of Després [21]. This formalism and the analogy with a fixed-point system have been widely used, e.g. in [8, 17, 18, 29, 48, 50, 55]. Our CHDG method can in fact be seen as an element-wise DD method. The key novelty of our approach, however, is that our discrete transmission conditions are built from the numerical fluxes naturally arising in the DG setting. In particular, cross-points where several mesh faces meet are naturally handled without any specific treatment. In contrast, standard DD algorithms based on conforming finite elements require specific (and sometimes non-local) swap operators to properly account for such cross-points [12, 13, 53].

In this work, the CHDG method with auxiliary characteristic variables is introduced and studied for the numerical solution of Helmholtz problems. We rigorously show that the resulting reduced system set on the skeleton of the mesh is well-posed and algebraically equivalent to the original upwind DG method. Moreover, we prove that this reduced system corresponds to a fixed-point problem with a strict contraction, which can therefore always be solved with the Richardson iteration. Then, the performance of CHDG is compared to the original DG scheme and its standard HDG reformulation with a sequence of numerical benchmarks. These examples show that the

standard Richardson iteration always converges without relaxation (although sometimes slowly) for the CHDG approach, whereas this approach fails to converge for DG and HDG. Finally, the convergence of standard Krylov methods is compared for the three approaches. We find that CHDG always requires fewer iterations than DG and HDG to reach a given accuracy with the GMRES and CGNR iterations.

The remainder of this work is structured as follows. In Section 2, we introduce the notations, and describe the upwind DG, the standard HDG, and the CHDG methods as well as their basic properties. In Section 3, the reduced system obtained with CHDG is analyzed in detail. We describe our numerical benchmarks in Section 4, where we also comment on the required memory space and conditioning properties of the different approaches. We study the convergence of standard iterative schemes in Section 5 and present our concluding remarks in Section 6.

2 Hybridizable discontinuous Galerkin methods

Let $\Omega \subset \mathbb{R}^d$, with $d = 2$ or 3, be a Lipschitz polytopal domain. The boundary $\partial\Omega$ of the domain is partitioned into three non-overlapping polytopal Lipschitz subsets Γ_{D}, Γ_{N} and Γ_{R}. We consider the following time-harmonic scalar wave propagation problem:

$$\begin{cases} -\imath\kappa u + \boldsymbol{\nabla}\cdot\mathbf{q} = 0, & \text{in } \Omega, \\ -\imath\kappa\mathbf{q} + \boldsymbol{\nabla} u = \mathbf{0}, & \text{in } \Omega, \\ u = s_{\mathrm{D}}, & \text{on } \Gamma_{\mathrm{D}}, \\ \mathbf{n}\cdot\mathbf{q} = s_{\mathrm{N}}, & \text{on } \Gamma_{\mathrm{N}}, \\ u - \mathbf{n}\cdot\mathbf{q} = s_{\mathrm{R}}, & \text{on } \Gamma_{\mathrm{R}}, \end{cases} \tag{2}$$

where the unknowns $u : \Omega \to \mathbb{C}$ and $\mathbf{q} : \Omega \to \mathbb{C}^d$ represent a time-harmonic wave, $\kappa > 0$ is a given real constant called the wavenumber, and $\mathbf{n}$ stands for the unit outward normal to Ω. The functions $s_{\mathrm{D}} : \Gamma_{\mathrm{D}} \to \mathbb{C}$, $s_{\mathrm{N}} : \Gamma_{\mathrm{N}} \to \mathbb{C}$ and $s_{\mathrm{R}} : \Gamma_{\mathrm{R}} \to \mathbb{C}$ are boundary data representing an incident field. Specifically, (2) is a particular case of the acoustic wave equation, where we have assumed a time dependence $e^{-\imath\omega t}$ for the data and the solution and $\kappa := \omega/c$, where ω is the angular frequency, t is the time and c is the (uniform) wave speed. For the sake of brevity, we do not consider volume right-hand sides in the two first equations of (2), but these could be included without difficulty.

2.1 Mesh, approximation spaces and inner products

We consider a conforming mesh $\mathcal{T}_h$ of the domain Ω consisting of simplicial elements K. The collection of element boundaries is denoted by $\partial\mathcal{T}_h := \{\partial K \mid K \in \mathcal{T}_h\}$, and the collection of faces is denoted by $\mathcal{F}_h$. The collection of faces of an element K is denoted by $\mathcal{F}_K$.

The approximate fields produced by DG schemes are piecewise polynomials. Here, for the sake of simplicity, we fix a polynomial degree $p \geq 0$ and introduce

$$V_h := \prod_{K\in\mathcal{T}_h} \mathcal{P}_p(K) \qquad \text{and} \qquad \mathbf{V}_h := \prod_{K\in\mathcal{T}_h} \boldsymbol{\mathcal{P}}_p(K),$$

where $\mathcal{P}_p(\cdot)$ and $\boldsymbol{\mathcal{P}}_p(\cdot)$ denote spaces of scalar and vector complex-valued polynomials of degree smaller or equal to p. By convention, the restrictions of $u_h \in V_h$ and $\mathbf{u}_h \in \mathbf{V}_h$ on K are denoted u_K and $\mathbf{u}_K$, respectively.

We introduce the sesquilinear forms

$$(u,v)_K := \int_K u\overline{v}\,d\mathbf{x}, \qquad (\mathbf{u},\mathbf{v})_K := \int_K \mathbf{u}\cdot\overline{\mathbf{v}}\,d\mathbf{x}, \qquad \langle u,v\rangle_{\partial K} := \sum_{F\in\mathcal{F}_K}\int_F u\overline{v}\,d\sigma(\mathbf{x}),$$

$$(u,v)_{\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h}(u,v)_K, \qquad (\mathbf{u},\mathbf{v})_{\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h}(\mathbf{u},\mathbf{v})_K, \qquad \langle u,v\rangle_{\partial\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h}\langle u,v\rangle_{\partial K}.$$

By convention, the quantities used in the surface integral $\langle \cdot, \cdot \rangle_{\partial K}$ correspond to the restriction of fields defined on K (e.g. v_K and $\mathbf{v}_K$) or quantities associated to the faces of K (e.g. $\mathbf{n}_{K,F}$ with $F \in \mathcal{F}_K$).

2.2 Standard DG formulation and numerical fluxes

The general DG formulation of system (2) reads:

Problem 2.1. *Find* $(u_h, \mathbf{q}_h) \in V_h \times \mathbf{V}_h$ *such that, for all* $(v_h, \mathbf{p}_h) \in V_h \times \mathbf{V}_h$,

$$\begin{cases} -\imath\kappa(u_h, v_h)_{\mathcal{T}_h} - (\mathbf{q}_h, \boldsymbol{\nabla} v_h)_{\mathcal{T}_h} + \langle \mathbf{n} \cdot \widehat{\mathbf{q}}(u_h, \mathbf{q}_h), v_h \rangle_{\partial \mathcal{T}_h} = 0, \\ -\imath\kappa(\mathbf{q}_h, \mathbf{p}_h)_{\mathcal{T}_h} - (u_h, \boldsymbol{\nabla} \cdot \mathbf{p}_h)_{\mathcal{T}_h} + \langle \widehat{u}(u_h, \mathbf{q}_h), \mathbf{n} \cdot \mathbf{p}_h \rangle_{\partial \mathcal{T}_h} = 0, \end{cases}$$

where the numerical fluxes $\widehat{u}(u_h, \mathbf{q}_h)$ *and* $\mathbf{n} \cdot \widehat{\mathbf{q}}(u_h, \mathbf{q}_h)$ *are defined face by face below.*

The properties of DG formulations intrinsically depend on the choice of the numerical fluxes. In this work, we consider *upwind fluxes*. For an interior face $F \not\subset \partial\Omega$ of an element K, these fluxes can be written as

$$\begin{cases} \widehat{u}_F := \dfrac{u_K + u_{K'}}{2} + \mathbf{n}_{K,F} \cdot \left(\dfrac{\mathbf{q}_K - \mathbf{q}_{K'}}{2} \right), \\ \mathbf{n}_{K,F} \cdot \widehat{\mathbf{q}}_F := \mathbf{n}_{K,F} \cdot \left(\dfrac{\mathbf{q}_K + \mathbf{q}_{K'}}{2} \right) + \dfrac{u_K - u_{K'}}{2}, \end{cases} \tag{3a}$$

where K' is the neighboring element and $\mathbf{n}_{K,F}$ is the unit outward normal to K on F. For a boundary face $F \subset \partial\Omega$ of an element K, the fluxes are defined as

$$\begin{cases} \widehat{u}_F := s_{\mathrm{D}}, \\ \mathbf{n}_{K,F} \cdot \widehat{\mathbf{q}}_F := \mathbf{n}_{K,F} \cdot \mathbf{q}_K + (u_K - s_{\mathrm{D}}), \end{cases} \quad \text{if } F \subset \Gamma_{\mathrm{D}}, \tag{3b}$$

$$\begin{cases} \widehat{u}_F := u_K + (\mathbf{n}_{K,F} \cdot \mathbf{q}_K - s_{\mathrm{N}}), \\ \mathbf{n}_{K,F} \cdot \widehat{\mathbf{q}}_F := s_{\mathrm{N}}, \end{cases} \quad \text{if } F \subset \Gamma_{\mathrm{N}}, \tag{3c}$$

$$\begin{cases} \widehat{u}_F := (u_K + \mathbf{n}_{K,F} \cdot \mathbf{q}_K + s_{\mathrm{R}})/2, \\ \mathbf{n}_{K,F} \cdot \widehat{\mathbf{q}}_F := (u_K + \mathbf{n}_{K,F} \cdot \mathbf{q}_K - s_{\mathrm{R}})/2, \end{cases} \quad \text{if } F \subset \Gamma_{\mathrm{R}}. \tag{3d}$$

The upwind fluxes are consistent, which means that $\widehat{u}(u, \mathbf{q}) = u$ and $\mathbf{n} \cdot \widehat{\mathbf{q}}(u, \mathbf{q}) = \mathbf{n} \cdot \mathbf{q}$ on both interior and boundary faces when u and $\mathbf{q}$ are the solution of Problem (2). Under standard assumptions, the method achieves the optimal convergence rate for the numerical fields u_h and $\mathbf{q}_h$ in L^2-norm, i.e. $p+1$ where p is the polynomial degree of the basis functions. Error estimates have been derived for HDG formulations, equivalent to the DG formulation above, for the Helmholtz problem with a Dirichlet boundary condition in [35] and a Robin boundary condition in [19, 26]. By using a post-processing, the convergence rate for u_h can be increased by one, see e.g. [15].

2.3 Hybridization with numerical trace — Standard HDG method

In standard HDG formulations, an additional variable $\widehat{u}_h$ corresponding to the numerical flux $\widehat{u}$ is introduced at the interface between the elements and on the boundary faces. The discrete unknowns associated to the fields u_h and $\mathbf{q}_h$ are eliminated in the solution procedure, leading to a reduced system with discrete unknowns associated to $\widehat{u}_h$ on the skeleton, see e.g. [14, 16].

The additional variable, which is called the *numerical trace* in the HDG literature, belongs to the space $\widehat{V}_h$ defined as

$$\widehat{V}_h := \prod_{F \in \mathcal{F}_h} \mathcal{P}_p(F).$$

For any field $\widehat{u}_h \in \widehat{V}_h$, there is one set of scalar unknowns associated to each face of the mesh. After observing that

$$\mathbf{n} \cdot \widehat{\mathbf{q}}(u_h, \mathbf{q}_h) = u_h + \mathbf{n} \cdot \mathbf{q}_h - \widehat{u}_h,$$

we obtain the following HDG formulation, where the numerical trace appears as a hybrid variable:

Problem 2.2. *Find* $(u_h, \mathbf{q}_h, \widehat{u}_h) \in V_h \times \mathbf{V}_h \times \widehat{V}_h$ *such that, for all* $(v_h, \mathbf{p}_h, \widehat{v}_h) \in V_h \times \mathbf{V}_h \times \widehat{V}_h$,

$$\begin{cases} -\imath\kappa(u_h, v_h)_{\mathcal{T}_h} - (\mathbf{q}_h, \boldsymbol{\nabla} v_h)_{\mathcal{T}_h} + \langle u_h + \mathbf{n} \cdot \mathbf{q}_h - \widehat{u}_h, v_h \rangle_{\partial\mathcal{T}_h} = 0, \\ -\imath\kappa(\mathbf{q}_h, \mathbf{p}_h)_{\mathcal{T}_h} - (u_h, \boldsymbol{\nabla} \cdot \mathbf{p}_h)_{\mathcal{T}_h} + \langle \widehat{u}_h, \mathbf{n} \cdot \mathbf{p}_h \rangle_{\partial\mathcal{T}_h} = 0 \end{cases}$$

and

$$\begin{aligned} \langle \widehat{u}_h, \widehat{v}_h \rangle_{\mathcal{F}_h} - \langle \tfrac{1}{2}(u_h + \mathbf{n} \cdot \mathbf{q}_h), \widehat{v}_h \rangle_{\partial\mathcal{T}_h \setminus \partial\Omega} - \langle u_h + \mathbf{n} \cdot \mathbf{q}_h, \widehat{v}_h \rangle_{\Gamma_\mathrm{N}} - \langle \tfrac{1}{2}(u_h + \mathbf{n} \cdot \mathbf{q}_h), \widehat{v}_h \rangle_{\Gamma_\mathrm{R}} \\ = \langle s_\mathrm{D}, \widehat{v}_h \rangle_{\Gamma_\mathrm{D}} - \langle s_\mathrm{N}, \widehat{v}_h \rangle_{\Gamma_\mathrm{N}} + \langle \tfrac{1}{2} s_\mathrm{R}, \widehat{v}_h \rangle_{\Gamma_\mathrm{R}}. \end{aligned}$$

This formulation is equivalent to the standard DG formulation (Problem 2.1) in the sense that the discrete solutions u_h and $\mathbf{q}_h$ are identical, see e.g. [44].

In the HDG literature [16, 35, 44], a generalization of the above formulation is often considered with

$$\mathbf{n} \cdot \widehat{\mathbf{q}}(u_h, \mathbf{q}_h) = \mathbf{n} \cdot \mathbf{q}_h + \tau(u_h - \widehat{u}_h),$$

where τ is the so-called *stabilization function*. In this work, we focus on the case where $\tau = 1$, which corresponds to the standard upwind fluxes and is widely used in practice.

Remark 2.3 (Source projection). The numerical trace $\widehat{u}_h$ is a polynomial function on every face, whereas the numerical flux $\widehat{u}$ introduced in the previous section may be a more general function at any boundary face where the boundary data does not belong to $\mathcal{P}_p(F)$. Nevertheless, in practice, equations (3b)-(3d) are still valid for $\widehat{u}_h$ if the boundary data are projected into the polynomial spaces.

Local element-wise discrete problems

In the solution procedure, the fields u_h and $\mathbf{q}_h$ are eliminated by solving local element-wise problems, where the numerical trace $\widehat{u}_h$ is considered as a given data.

For each element K, the local problem reads:

Problem 2.4. *Find* $(u_K, \mathbf{q}_K) \in \mathcal{P}_p(K) \times \boldsymbol{\mathcal{P}}_p(K)$ *such that, for all* $(v_K, \mathbf{p}_K) \in \mathcal{P}_p(K) \times \boldsymbol{\mathcal{P}}_p(K)$,

$$\begin{cases} -\imath\kappa(u_K, v_K)_K - (\mathbf{q}_K, \boldsymbol{\nabla} v_K)_K + \displaystyle\sum_{F \in \mathcal{F}_K} \langle u_K + \mathbf{n}_{K,F} \cdot \mathbf{q}_K, v_K \rangle_F = \sum_{F \in \mathcal{F}_K} \langle \widehat{u}_F, v_K \rangle_F, \\ -\imath\kappa(\mathbf{q}_K, \mathbf{p}_K)_K - (u_K, \boldsymbol{\nabla} \cdot \mathbf{p}_K)_K = -\displaystyle\sum_{F \in \mathcal{F}_K} \langle \widehat{u}_F, \mathbf{n}_{K,F} \cdot \mathbf{p}_K \rangle_F, \end{cases}$$

for given surface data $\widehat{u}_F \in \mathcal{P}_p(F)$ *for all* $F \in \mathcal{F}_K$.

This local discrete problem is similar to a Helmholtz problem defined on K with a non-homogeneous Dirichlet boundary condition on ∂K. The discrete problem is well-posed without any condition, as shown e.g. in [34]. We include the proof here for the sake of completeness.

Theorem 2.5 (Well-posedness of the local discrete problem)**.** *Problem 2.4 is well-posed.*

Proof. We simply have to prove that, if $\widehat{u}_F = 0$ for all $F \in \mathcal{F}_K$, the unique solution of Problem 2.4 is $u_K = 0$ and $\mathbf{q}_K = \mathbf{0}$. For the sake of brevity, the subscripts K and F are omitted for the local

fields, the test functions, the unit outgoing normal and the surface data. Taking both equations of Problem 2.4 with $v = u$ and $\mathbf{p} = \mathbf{q}$ gives

$$\begin{aligned} -\imath\kappa(u,u)_K - (\mathbf{q}, \boldsymbol{\nabla} u)_K + \langle u + \mathbf{n}\cdot\mathbf{q}, u\rangle_{\partial K} &= 0, \\ -\imath\kappa(\mathbf{q},\mathbf{q})_K - (u, \boldsymbol{\nabla}\cdot\mathbf{q})_K &= 0. \end{aligned}$$

Integrating by parts in both equations and taking the complex conjugate lead to

$$\begin{aligned} \imath\kappa(u,u)_K + (u, \boldsymbol{\nabla}\cdot\mathbf{q})_K + \langle u, u\rangle_{\partial K} &= 0, \\ \imath\kappa(\mathbf{q},\mathbf{q})_K + (\mathbf{q}, \boldsymbol{\nabla} u)_K - \langle \mathbf{n}\cdot\mathbf{q}, u\rangle_{\partial K} &= 0. \end{aligned}$$

Adding the four previous equations yields $\langle u, u\rangle_{\partial K} = 0$, and then $u = 0$ on ∂K. By using this result in Problem 2.4, one has

$$\begin{cases} -\imath\kappa(u,v)_K + (\boldsymbol{\nabla}\cdot\mathbf{q}, v)_K = 0, \\ -\imath\kappa(\mathbf{q},\mathbf{p})_K + (\boldsymbol{\nabla} u, \mathbf{p})_K = 0, \end{cases}$$

for all $[v, \mathbf{p}] \in \mathcal{P}_p(K) \times \boldsymbol{\mathcal{P}}_p(K)$. We conclude that

$$\begin{aligned} -i\kappa u + \boldsymbol{\nabla}\cdot\mathbf{q} &= 0, \\ -i\kappa\mathbf{q} + \boldsymbol{\nabla} u &= \mathbf{0}, \end{aligned}$$

in a strong sense. Because there is no non-trivial polynomial solution to the previous equations, this yields the result. □

Remark 2.6 (Conditioning). At the continuous level, Helmholtz problems with Dirichlet boundary conditions are ill-posed if the frequency corresponds to an eigenvalue of the Laplace operator. Here, the Dirichlet conditions are weakly imposed through penalization, so that the discrete problem are always well-posed. Nevertheless, we shall see in Section 4.4 that the matrices of the local systems becomes ill-conditioned as kh goes to zero.

2.4 Hybridization with characteristic variables — CHDG method

We propose a new hybridization procedure where the additional variable is associated to incoming and outgoing fluxes at every face of the mesh. More precisely, the additional variable corresponds to the incoming characteristic variable relative to each element. Similarly to the standard HDG method, the discrete unknowns associated to the fields u_h and $\mathbf{q}_h$ are eliminated in the solution procedure, leading to a reduced system with discrete unknowns associated to the incoming characteristic variable on the skeleton.

Characteristic variables

At each interior face $F \not\subset \partial\Omega$ of an element K, the *outgoing characteristic variable* $g^{\oplus}_{K,F}$ and the *incoming characteristic variable* $g^{\ominus}_{K,F}$ are defined as

$$\begin{aligned} g^{\oplus}_{K,F} &:= u_K + \mathbf{n}_{K,F}\cdot\mathbf{q}_K, \\ g^{\ominus}_{K,F} &:= u_{K'} - \mathbf{n}_{K,F}\cdot\mathbf{q}_{K'}, \end{aligned} \tag{4}$$

respectively, where K' is the neighboring element. Let us highlight that the outgoing characteristic variable depends only on values corresponding to element K, whereas the incoming one depends only on values corresponding to the neighboring element K'. The outgoing characteristic variable of one side corresponds to the incoming one of the other side, i.e. $g^{\oplus}_{K,F} = g^{\ominus}_{K',F}$ and $g^{\ominus}_{K,F} = g^{\oplus}_{K',F}$. The notations are illustrated on Figure 1.

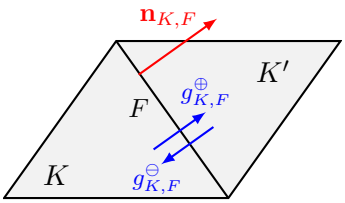

Figure 1: Notations for the outgoing and incoming characteristic variables (resp. $g^{\oplus}_{K,F}$ and $g^{\ominus}_{K,F}$) at the face F shared by an element K and a neighboring element K'. Let us note that $g^{\oplus}_{K,F} = g^{\ominus}_{K',F}$ and $g^{\ominus}_{K,F} = g^{\oplus}_{K',F}$.

The characteristic variables can be interpreted as information transported towards the exterior and the interior of K, respectively. Indeed, let us consider the time-domain version of the governing equations. Assuming there is no source and the fields are varying only in direction $\mathbf{n}$, we get

$$\begin{cases} \partial_t u + c\,\partial_n(\mathbf{n}\cdot\mathbf{q}) = 0, \\ \partial_t(\mathbf{n}\cdot\mathbf{q}) + c\,\partial_n u = 0. \end{cases}$$

A simple linear combination gives the transport equations

$$\begin{cases} \partial_t(u + \mathbf{n}\cdot\mathbf{q}) + c\,\partial_n(u + \mathbf{n}\cdot\mathbf{q}) = 0, \\ \partial_t(u - \mathbf{n}\cdot\mathbf{q}) - c\,\partial_n(u - \mathbf{n}\cdot\mathbf{q}) = 0. \end{cases}$$

Therefore, $g^{\oplus} = u + \mathbf{n}\cdot\mathbf{q}$ and $g^{\ominus} = u - \mathbf{n}\cdot\mathbf{q}$ correspond to quantities transported in the domain in directions $+\mathbf{n}$ (downstream) and $-\mathbf{n}$ (upstream), respectively, at velocity c. In the CFD community, the variables $g^{\oplus}$ and $g^{\ominus}$ are generally called *characteristic variables* (see e.g. [59]), and they are used to define *upwind fluxes* for solving time-dependent problems. For more general problems, characteristic variables and upwind fluxes are obtained by solving local Riemann problems along the normal direction, see e.g. [36, 59].

The numerical fluxes (3a) can be rewritten with the characteristic variables as

$$\begin{cases} \widehat{u}_F = (g^{\oplus}_{K,F} + g^{\ominus}_{K,F})/2, \\ \mathbf{n}_{K,F}\cdot\widehat{\mathbf{q}}_F = (g^{\oplus}_{K,F} - g^{\ominus}_{K,F})/2. \end{cases}$$

If F is a boundary face, i.e. $F \subset \partial\Omega$, the numerical fluxes and the outgoing characteristic variable can be defined with (4), but the incoming characteristic variable must be defined differently because there is no neighboring element. It is defined as

$$g^{\ominus}_{K,F} := 2s_{\mathrm{D}} - g^{\oplus}_{K,F}, \qquad \text{if } F \subset \Gamma_{\mathrm{D}}, \tag{5a}$$

$$g^{\ominus}_{K,F} := g^{\oplus}_{K,F} - 2s_{\mathrm{N}}, \qquad \text{if } F \subset \Gamma_{\mathrm{N}}, \tag{5b}$$

$$g^{\ominus}_{K,F} := s_{\mathrm{R}}, \qquad \text{if } F \subset \Gamma_{\mathrm{R}}. \tag{5c}$$

By using these definitions, the numerical fluxes corresponding to the boundary conditions, i.e. equations (3b)-(3d), are recovered. Therefore, the boundary conditions are prescribed directly in the definition of the incoming characteristic variables.

CHDG formulation

In the proposed method, the additional variable, denoted $g^{\ominus}_h$, corresponds to the incoming characteristic variable at the boundary of all the elements. The variable $g^{\ominus}_h$ belongs to the space G_h defined as

$$G_h := \prod_{K\in\mathcal{T}_h} \prod_{F\in\mathcal{F}_K} \mathcal{P}_p(F).$$

For any $g_h^{\ominus} \in G_h$, there are two sets of unknowns at each interior face of the mesh, which correspond to the incoming characteristic variable associated to the neighboring elements. In the following, the method is called the CHDG method. The first letter of the name refers to the "c" in "characteristic variable".

The CHDG formulation reads:

Problem 2.7. *Find* $(u_h, \mathbf{q}_h, g_h^{\ominus}) \in V_h \times \mathbf{V}_h \times G_h$ *such that, for all* $(v_h, \mathbf{p}_h, \xi_h) \in V_h \times \mathbf{V}_h \times G_h$,

$$\begin{cases} -\imath\kappa(u_h, v_h)_{\mathcal{T}_h} - (\mathbf{q}_h, \boldsymbol{\nabla} v_h)_{\mathcal{T}_h} + \langle \frac{1}{2}(g^{\oplus}(u_h, \mathbf{q}_h) - g_h^{\ominus}), v_h \rangle_{\partial\mathcal{T}_h} = 0, \\ -\imath\kappa(\mathbf{q}_h, \mathbf{p}_h)_{\mathcal{T}_h} - (u_h, \boldsymbol{\nabla} \cdot \mathbf{p}_h)_{\mathcal{T}_h} + \langle \frac{1}{2}(g^{\oplus}(u_h, \mathbf{q}_h) + g_h^{\ominus}), \mathbf{n} \cdot \mathbf{p}_h \rangle_{\partial\mathcal{T}_h} = 0, \end{cases}$$

and

$$\langle g_h^{\ominus} - \Pi(g^{\oplus}(u_h, \mathbf{q}_h)), \xi_h \rangle_{\partial\mathcal{T}_h} = \langle b, \xi_h \rangle_{\partial\mathcal{T}_h}, \tag{6}$$

with $g^{\oplus}(u_h, \mathbf{q}_h) := u_h + \mathbf{n} \cdot \mathbf{q}_h$.

The operator $\Pi : G_h \longrightarrow G_h$ used in equation (6) is the *global exchange operator*. It is the key mechanism to enforce the weak coupling of the element-wise problems at the interior faces and to enforce the boundary conditions at the boundary faces. At interior faces, it simply swaps the outgoing characteristics of the two neighboring elements. This definition is suitably modified at boundary faces to account for boundary conditions. For each face F of each element K, Π is defined as

$$\Pi(g^{\oplus})|_{K,F} = \begin{cases} g_{K',F}^{\oplus} & \text{if } F \not\subset \partial\Omega \text{ is shared by } K \text{ and } K', \\ -g_{K,F}^{\oplus} & \text{if } F \subset \Gamma_{\mathrm{D}}, \\ g_{K,F}^{\oplus} & \text{if } F \subset \Gamma_{\mathrm{N}}, \\ 0 & \text{if } F \subset \Gamma_{\mathrm{R}}, \end{cases} \tag{7}$$

for any $g^{\oplus} \in G_h$. For each face F of each element K, the *global right-hand side* b is given by

$$b|_{K,F} = \begin{cases} 0 & \text{if } F \not\subset \partial\Omega, \\ 2s_{\mathrm{D}} & \text{if } F \subset \Gamma_{\mathrm{D}}, \\ -2s_{\mathrm{N}} & \text{if } F \subset \Gamma_{\mathrm{N}}, \\ s_{\mathrm{R}} & \text{if } F \subset \Gamma_{\mathrm{R}}. \end{cases}$$

Therefore, Equation (6) is equivalent to the following relations:

$$\begin{aligned} \langle g_{K,F}^{\ominus}, \xi_{K,F} \rangle_F - \langle u_{K'} + \mathbf{n}_{K',F} \cdot \mathbf{q}_{K'}, \xi_{K,F} \rangle_F &= 0, && \text{if } F \not\subset \partial\Omega, \\ \langle g_{K,F}^{\ominus}, \xi_{K,F} \rangle_F + \langle u_K + \mathbf{n}_{K,F} \cdot \mathbf{q}_K, \xi_{K,F} \rangle_F &= \langle 2s_{\mathrm{D}}, \xi_{K,F} \rangle_F, && \text{if } F \subset \Gamma_{\mathrm{D}}, \\ \langle g_{K,F}^{\ominus}, \xi_{K,F} \rangle_F - \langle u_K + \mathbf{n}_{K,F} \cdot \mathbf{q}_K, \xi_{K,F} \rangle_F &= -\langle 2s_{\mathrm{N}}, \xi_{K,F} \rangle_F, && \text{if } F \subset \Gamma_{\mathrm{N}}, \\ \langle g_{K,F}^{\ominus}, \xi_{K,F} \rangle_F &= \langle s_{\mathrm{R}}, \xi_{K,F} \rangle_F, && \text{if } F \subset \Gamma_{\mathrm{R}}, \end{aligned}$$

for each face F of each element K. The first relation enforces that the incoming characteristic variable of an element is the outgoing one of the neighboring element, and vice versa, for each interior face. The other relations enforce the boundary conditions.

The CHDG formulation is equivalent to the standard DG formulation (Problem 2.1), and thus to the standard HDG formulation (Problem 2.2). Similar to the standard HDG formulation, the additional variable $g_h^{\ominus}$ is a polynomial function on each face, whereas the incoming characteristic variable introduced previously could be a more general function on the boundary of the domain. Nevertheless, equations (5a)-(5c) still hold up to projecting the right-hand sides onto piecewise polynomials.

Local element-wise discrete problems

The hybridization procedure leads to a reduced system with discrete unknowns associated to the incoming characteristic variable $g_h^{\ominus}$ on the skeleton. This elimination is achieved by solving local element-wise problems, where the incoming characteristic variable is considered as a given data.

For each element K, the local problem reads:

Problem 2.8. *Find* $(u_K, \mathbf{q}_K) \in \mathcal{P}_p(K) \times \boldsymbol{\mathcal{P}}_p(K)$ *such that, for all* $(v_K, \mathbf{p}_K) \in \mathcal{P}_p(K) \times \boldsymbol{\mathcal{P}}_p(K)$,

$$\begin{cases} -\imath\kappa(u_K, v_K)_K - (\mathbf{q}_K, \boldsymbol{\nabla} v_K)_K + \sum_{F\in\mathcal{F}_K} \langle \frac{1}{2} g_{K,F}^{\oplus}, v_K\rangle_F = \sum_{F\in\mathcal{F}_K} \langle \frac{1}{2} g_{K,F}^{\ominus}, v_K\rangle_F, \\ -\imath\kappa(\mathbf{q}_K, \mathbf{p}_K)_K - (u_K, \boldsymbol{\nabla}\cdot \mathbf{p}_K)_K + \sum_{F\in\mathcal{F}_K} \langle \frac{1}{2} g_{K,F}^{\oplus}, \mathbf{n}_{K,F}\cdot \mathbf{p}_K\rangle_F = -\sum_{F\in\mathcal{F}_K} \langle \frac{1}{2} g_{K,F}^{\ominus}, \mathbf{n}_{K,F}\cdot \mathbf{p}_K\rangle_F, \end{cases}$$

with $g_{K,F}^{\oplus} = u_K + \mathbf{n}_{K,F}\cdot \mathbf{q}_K$, *for given surface data* $g_{K,F}^{\ominus} \in \mathcal{P}_p(F)$ *for all* $F \in \mathcal{F}_K$.

The local problem can be interpreted as a discretized Helmholtz problem defined on K with a non-homogeneous Robin boundary condition on ∂K. We show hereafter that this discrete problem is well-posed.

Theorem 2.9 (Well-posedness of the local discrete problem)**.** *Problem 2.8 is well-posed.*

Proof. We simply have to prove that, if $g_{K,F}^{\ominus} = 0$ for all $F \in \mathcal{F}_K$, the unique solution of Problem 2.8 is $u_K = 0$ and $\mathbf{q}_K = \mathbf{0}$. For the sake of brevity, the subscripts K and F are omitted for the local fields, the test functions, the unit outgoing normal and the surface data. Taking both equations of Problem (2.8) with $v = u$ and $\mathbf{p} = \mathbf{q}$ gives

$$\begin{aligned} -\imath\kappa(u,u)_K - (\mathbf{q}, \boldsymbol{\nabla} u)_K + \langle \tfrac{1}{2}(u + \mathbf{n}\cdot\mathbf{q}), u\rangle_{\partial K} &= 0, \\ -\imath\kappa(\mathbf{q},\mathbf{q})_K - (u, \boldsymbol{\nabla}\cdot\mathbf{q})_K + \langle \tfrac{1}{2}(u + \mathbf{n}\cdot\mathbf{q}), \mathbf{n}\cdot\mathbf{q}\rangle_{\partial K} &= 0. \end{aligned}$$

Integrating by parts in both equations and taking the complex conjugate lead to

$$\begin{aligned} \imath\kappa(u,u)_K + (u, \boldsymbol{\nabla}\cdot\mathbf{q})_K + \langle u, \tfrac{1}{2}(u - \mathbf{n}\cdot\mathbf{q})\rangle_{\partial K} &= 0, \\ \imath\kappa(\mathbf{q},\mathbf{q})_K + (\mathbf{q}, \boldsymbol{\nabla} u)_K - \langle \mathbf{n}\cdot\mathbf{q}, \tfrac{1}{2}(u - \mathbf{n}\cdot\mathbf{q})\rangle_{\partial K} &= 0. \end{aligned}$$

Adding the four previous equations yields $\langle u, u\rangle_{\partial K} + \langle \mathbf{n}\cdot\mathbf{q}, \mathbf{n}\cdot\mathbf{q}\rangle_{\partial K} = 0$, which gives $u = 0$ and $\mathbf{n}\cdot\mathbf{q} = 0$ on ∂K. By using these boundary conditions in Problem (2.8), we have that the fields should be a solution of the strong problem. Because there is no solution with both homogeneous Neumann and Dirichlet boundary conditions, this yields the result. □

Remark 2.10 (Conditioning). In contrast to Helmholtz problems with Dirichlet boundary conditions, the local problems with Robin boundary conditions are always well-posed at the continuous level. We shall see in Section 4.4 that the matrices of the local systems stays well-conditioned as kh goes to zero for low-order finite elements, and that the condition number is smaller than with HDG for high-order finite elements.

3 Analysis of the reduced system for the CHDG method

In this section, we introduce and study the reduced version of the hybridized formulation with characteristic variables (Problem 2.7). This version is obtained by solving the local element-wise problems (Problem 2.8) and then eliminating the physical variables u_h and $\mathbf{q}_h$ from the system.

3.1 Formulation of the reduced system

In order to write the problem in a reduced formulation, we introduce the *global scattering operator* $\mathrm{S} : G_h \longrightarrow G_h$ defined such that, for each face F of each element K,

$$\mathrm{S}(g_h^{\ominus})\big|_{K,F} := u_K(g_h^{\ominus}) + \mathbf{n}_{K,F} \cdot \mathbf{q}_K(g_h^{\ominus}), \tag{8}$$

where $(u_K, \mathbf{q}_K)$ is the solution of Problem 2.8 with the incoming characteristic data $(g_{K,F}^{\ominus})_{F\in\mathcal{F}_K}$ contained in $g_h^{\ominus}$ as a given surface data. This operator can be interpreted as an *"incoming characteristic variable to outgoing characteristic variable"* operator.

By using the operator S, Problem 2.7 is rewritten as:

Problem 3.1. *Find $g_h^{\ominus} \in G_h$ such that, for all $\xi_h \in G_h$,*

$$\langle g_h^{\ominus}, \xi_h\rangle_{\partial\mathcal{T}_h} - \langle \Pi(\mathrm{S}(g_h^{\ominus})), \xi_h\rangle_{\partial\mathcal{T}_h} = \langle b, \xi_h\rangle_{\partial\mathcal{T}_h}.$$

In order to write the problem in a more compact form, we introduce the *global projected right-hand side* $b_h := P_h b \in G_h$, where $P_h : L^2(\partial\mathcal{T}_h) \longrightarrow G_h$ is the projection operator defined such that $\langle P_h b, \xi_h\rangle_{\partial\mathcal{T}_h} = \langle b, \xi_h\rangle_{\partial\mathcal{T}_h}$ for all $\xi_h \in G_h$. Problem 3.1 can then be rewritten as:

Problem 3.2. *Find $g_h^{\ominus} \in G_h$ such that*

$$(\mathrm{I} - \Pi\mathrm{S})g_h^{\ominus} = b_h.$$

Problems 3.1 and 3.2 are equivalent to Problem 2.7 because the element-wise local problems (Problem 2.8) are well-posed. As discussed in the introduction, Problem 3.2 is similar to formulations obtained for DD and UWVF methods to solve Helmholtz problems.

3.2 Fixed-point problem

Problem 3.2 corresponds to a fixed-point problem. In this section, we prove that the operator $\Pi\mathrm{S}$ is a strict contraction. As a consequence, the fixed-point problem is always well-posed, and it can (at least in principle) be solved with stationary iterative procedures. The algebraic version of this system is discussed in Sections 4.1 and 5.1.

The properties of S and Π are established by using a norm on G_h defined as

$$\|g_h^{\ominus}\| := \sqrt{\sum_{K\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_K} \|g_{K,F}^{\ominus}\|_F^2},$$

where $\|\cdot\|_F^2$ is the natural norm of $L^2(F)$. We start with a technical lemma.

Lemma 3.3. *(i) The solution of Problem 2.8 verifies*

$$\sum_{F\in\mathcal{F}_K} \|u_K + \mathbf{n}_{K,F}\cdot\mathbf{q}_K\|_F^2 + \sum_{F\in\mathcal{F}_K} \|u_K - \mathbf{n}_{K,F}\cdot\mathbf{q}_K - g_{K,F}^{\ominus}\|_F^2 = \sum_{F\in\mathcal{F}_K} \|g_{K,F}^{\ominus}\|_F^2. \tag{9}$$

(ii) The second term in the left-hand side of (9) vanishes if and only if $g_{K,F}^{\ominus} = 0$.

Proof. For the sake of brevity, the subscripts K and F are omitted for the local fields, the test functions, the unit outgoing normal and the surface data.

(i) Taking both equations of Problem 2.8 with $v = u$ and $\mathbf{p} = \mathbf{q}$ gives

$$\begin{aligned} -\imath\kappa(u,u)_K - (\mathbf{q}, \boldsymbol{\nabla} u)_K + \langle \tfrac{1}{2}(u + \mathbf{n}\cdot\mathbf{q}), u\rangle_{\partial K} &= \langle \tfrac{1}{2} g^{\ominus}, u\rangle_{\partial K} \\ -\imath\kappa(\mathbf{q},\mathbf{q})_K - (u, \boldsymbol{\nabla}\cdot\mathbf{q})_K + \langle \tfrac{1}{2}(u + \mathbf{n}\cdot\mathbf{q}), \mathbf{n}\cdot\mathbf{q}\rangle_{\partial K} &= -\langle \tfrac{1}{2} g^{\ominus}, \mathbf{n}\cdot\mathbf{q}\rangle_{\partial K}. \end{aligned}$$

Integrating by parts in both equations and taking the complex conjugate lead to

$$\begin{aligned} \imath\kappa(u,u)_K + (u, \boldsymbol{\nabla}\cdot\mathbf{q})_K + \langle u, \tfrac{1}{2}(u - \mathbf{n}\cdot\mathbf{q})\rangle_{\partial K} &= \langle u, \tfrac{1}{2} g^{\ominus}\rangle_{\partial K} \\ \imath\kappa(\mathbf{q},\mathbf{q})_K + (\mathbf{q}, \boldsymbol{\nabla} u)_K - \langle \mathbf{n}\cdot\mathbf{q}, \tfrac{1}{2}(u - \mathbf{n}\cdot\mathbf{q})\rangle_{\partial K} &= -\langle \mathbf{n}\cdot\mathbf{q}, \tfrac{1}{2} g^{\ominus}\rangle_{\partial K}. \end{aligned}$$

Adding the four previous equations yields

$$\frac{1}{2}\langle (u+\mathbf{n}\cdot\mathbf{q}),(u+\mathbf{n}\cdot\mathbf{q})\rangle_{\partial K}+\frac{1}{2}\langle (u-\mathbf{n}\cdot\mathbf{q}),(u-\mathbf{n}\cdot\mathbf{q})\rangle_{\partial K} \\ =\frac{1}{2}\langle g^{\ominus},(u-\mathbf{n}\cdot\mathbf{q})\rangle_{\partial K}+\frac{1}{2}\langle (u-\mathbf{n}\cdot\mathbf{q}),g^{\ominus}\rangle_{\partial K},$$

and then

$$\|u+\mathbf{n}\cdot\mathbf{q}\|^2_{\partial K}+\|u-\mathbf{n}\cdot\mathbf{q}\|^2_{\partial K}=\|u-\mathbf{n}\cdot\mathbf{q}\|^2_{\partial K}-\|u-\mathbf{n}\cdot\mathbf{q}-g^{\ominus}\|^2_{\partial K}+\|g^{\ominus}\|^2_{\partial K},$$

which gives the result (9).

(ii) If the second term in the left-hand side of (9) vanishes, then $g^{\ominus}=u-\mathbf{n}\cdot\mathbf{q}$ on ∂K. Using this relation in Problem 2.8, we see that u and $\mathbf{q}$ must satisfy

$$\begin{cases} -\imath\kappa(u,v)_K-(\mathbf{q},\boldsymbol{\nabla}v)_K+\langle \mathbf{n}\cdot\mathbf{q},v\rangle_{\partial K}=0,\\ -\imath\kappa(\mathbf{q},\mathbf{p})_K-(u,\boldsymbol{\nabla}\cdot\mathbf{p})_K+\langle u,\mathbf{n}\cdot\mathbf{p}\rangle_{\partial K}=0\end{cases}$$

for all $v\in\mathcal{P}_p(K)$ and $\mathbf{p}\in\boldsymbol{\mathcal{P}}_p(K)$, and integration by parts shows that u and $\mathbf{q}$ solve the Helmholtz equation in strong form. But as we have already seen in the proof of Theorem 2.5, there is no non-trivial polynomial solution, meaning that $u=0$ and $\mathbf{q}=\mathbf{0}$, and then $g^{\ominus}=0$. The converse statement is direct, because the local problem is well-posed. □

Theorem 3.4. *The scattering operator* S *is a strict contraction, i.e.*

$$\|\mathrm{S}(g_h^{\ominus})\|<\|g_h^{\ominus}\|,\quad \forall g_h^{\ominus}\in G_h\backslash\{0\}.$$

Proof. Let $g_h^{\ominus}\in G_h\backslash\{0\}$. By Lemma 3.3, one has

$$\sum_{F\in\mathcal{F}_K}\|u_K+\mathbf{n}_{K,F}\cdot\mathbf{q}_K\|^2_F<\sum_{F\in\mathcal{F}_K}\|g^{\ominus}_{K,F}\|^2_F.$$

The equality cannot happen because $g_h^{\ominus}\neq 0$. Then, by using the definition of S (i.e. equation (8)), one has

$$\sum_{F\in\mathcal{F}_K}\|\mathrm{S}(g_h^{\ominus})|_{K,F}\|^2_F<\sum_{F\in\mathcal{F}_K}\|g^{\ominus}_{K,F}\|^2_F.$$

Summing this estimate over all $K\in\mathcal{T}_h$ gives the result. □

The global scattering operator S is always strictly contracting whereas, in a continuous context, it preserves energy. The proof of Theorem 3.4 uses the fact that there are no polynomial solution to the Helmholtz equation, and therefore, the strict contraction property of S is a numerical artifact that is not physical. This is related to the fact the upwind DG scheme is a dissipative method to start with [2].

Theorem 3.5. *The exchange operator* Π *is a contraction, i.e.*

$$\|\Pi(g_h^{\ominus})\|\leq\|g_h^{\ominus}\|,\quad \forall g_h^{\ominus}\in G_h.$$

In addition, if $\Gamma_{\mathrm{R}}=\emptyset$, Π *is an involution, i.e.* $\Pi^2=\mathrm{I}$, *and an isometry, i.e.*

$$\|\Pi(g_h^{\ominus})\|=\|g_h^{\ominus}\|,\quad \forall g_h^{\ominus}\in G_h.$$

Proof. These results are straightforward consequences of the definition of Π. □

As a consequence of the two previous theorems, we have the following result.

Corollary 3.6. *The operator* $\Pi\mathrm{S}$ *is a strict contraction, i.e.*

$$\|\Pi\mathrm{S}(g_h^{\ominus})\| < \|g_h^{\ominus}\|, \quad \forall g_h^{\ominus} \in G_h \backslash \{0\}.$$

The strict contraction property of Corollary 3.6 is due to the fact that Π and/or S dissipate energy. Actually, the global scattering operator S is always strictly contracting. As discussed above, this can be related to the fact the upwind DG scheme is a dissipative method. On the other hand, the global exchange operator Π can only dissipate energy in the presence of a Robin boundary (see the last line of (7)), otherwise it is an involution. Therefore, we may identify two possible sources of dissipation. The first source is numerical dissipation which is always present, but may become small as the mesh is refined, leading to possibly slow convergence of fixed point iterations in energy-preserving problem. The other source of dissipation comes from physical absorption and should lead to faster convergence rates on fine meshes. The numerical examples we present in Section 5.1 clearly depict how the presence or absence of physical dissipation impact the convergence rates of fixed point iterations.

Let us note that, for conservative methods (including standard conforming finite elements), where S does not dissipate, $\Pi\mathrm{S}$ should preserve energy if there is no physical dissipation. In fact, the convergence of standard DD algorithms is proven only for energy-preserving problems with relaxation, e.g. [13]. It has been proven recently in [54] that the iteration matrix of a Trefftz DG method is also a strict contraction for a configuration with a Robin boundary condition. To the best of our knowledge, this is the only other example of finite element method that can be written with a strictly contracting iterative matrix for Helmholtz problems.

4 Linear algebraic systems

In this section, the algebraic systems resulting from the DG discretization and its two possible hybridizations are studied for two-dimensional problems. After a description of the polynomial basis and reference benchmarks in Sections 4.1 and 4.2, respectively, the required memory storage is discussed in Section 4.3. The condition numbers of the local element-wise matrices and the global reduced matrices are discussed in Sections 4.4 and 4.5, respectively.

4.1 Polynomial basis functions

The physical fields u_h and $\mathbf{q}_h$ are represented with standard hierarchical shape functions. These functions are built with tensor products of Lobatto shape functions (see e.g. [57, section 2.2.3] and [6]). For triangular elements, they are classified into vertex, edge, and bubble functions. Since the bubble functions vanish on the edges of the triangle, only the degrees of freedom associated to vertex and edge functions are involved in the boundary and interface integrals of the variational formulations. In remainder of this work, the edges of the triangular elements are called "faces" in order to follow the general terminology.

The fields defined on the skeleton, i.e. $\widehat{u}_h$ for HDG and $g_h^{\ominus}$ for CHDG, are univariate polynomials. A possible choice for the shape functions would be the Lobatto shape functions, which correspond to the restriction of the shape functions used for the physical fields. Instead, we consider scaled Legendre shape functions, which are orthogonal in $L^2(F)$ for each face F. For each element, they are scaled in such a way that the local mass matrices are the identity matrix, i.e.

$$(\phi_i^F, \phi_j^F)_F = \delta_{ij}, \quad \text{for } i, j = 1, \ldots, N_{\text{dof·per·fce}},$$

where ϕ_i^F and ϕ_j^F are the shape functions associated with the face F, and $N_{\text{dof·per·fce}}$ is the number of degrees of freedom per face.

The Lobatto functions and the scaled Legendre functions give rigorously the same numerical solution (up to floating point errors), as they are two equivalent sets of basis functions, but they lead to different algebraic systems. Let us consider the algebraic system resulting from the finite element discretization of Problem 3.1. With the Lobatto functions, the first term of this problem

corresponds to a mass matrix in the algebraic system. By contrast, with the scaled Legendre functions, it corresponds to an identity matrix as the shape functions are orthonormal. In fact, the system corresponding to the scaled Legendre functions, denoted $\mathsf{A}\mathbf{g} = \mathbf{b}$, can be obtained from the system corresponding to the Lobatto functions, denoted $\mathsf{A}_{\mathrm{Lob}}\mathbf{g}_{\mathrm{Lob}} = \mathbf{b}_{\mathrm{Lob}}$, by using a symmetric preconditioning:

$$\underbrace{\left(\mathsf{M}_{\mathrm{Lob}}^{-1/2}\mathsf{A}_{\mathrm{Lob}}\mathsf{M}_{\mathrm{Lob}}^{-1/2}\right)}_{\mathsf{A}}\underbrace{\left(\mathsf{M}_{\mathrm{Lob}}^{1/2}\mathbf{g}_{\mathrm{Lob}}\right)}_{\mathbf{g}} = \underbrace{\left(\mathsf{M}_{\mathrm{Lob}}^{-1/2}\mathbf{b}_{\mathrm{Lob}}\right)}_{\mathbf{b}}, \tag{10}$$

where $\mathsf{M}_{\mathrm{Lob}}$ is the mass matrix associated to the faces.

In preliminary comparison studies (not shown), we have observed that, for both HDG and CHDG methods, the convergence of the iterative solution procedures (without preconditioning strategy) is faster with the scaled Legendre functions than with the Lobatto functions. Here is a partial explanation. With the scaled Legendre functions, the scalar product $(\cdot,\cdot)_F$ of two fields is equal to the algebraic inner product on the corresponding components. Similarly, the L^2-norm of a field is equal to the 2-norm of its components. Therefore, the inner product and the norm used in the standard iterative solution procedures are in some sense *"natural"* for the considered problems. Note that this approach is rigorously equivalent to using the Lobbato functions with a symmetric preconditioning with the mass matrix $\mathsf{M}_{\mathrm{Lob}}$, see equation (10). In fact, that preconditioning approach is equivalent to using $\mathsf{M}_{\mathrm{Lob}}$ as a left preconditioner and using the scalar $(\cdot,\cdot)_F$ as inner product in weighted Krylov methods.

For the sake of brevity, only results with the scaled Legendre functions are presented in the remainder of this article.

4.2 Reference benchmarks

To study the properties of the algebraic systems and the convergence of iterative solution procedures, we consider three benchmarks corresponding to different physical configurations, already used in [10]. Snapshots of the real part of the solutions are shown in Figure 2. The numerical simulations have been performed with a dedicated MATLAB code. The mesh generation and the visualization have been done with gmsh [32] (version 4.11.1). In all the cases, third-degree polynomial bases, i.e. $p = 3$, have been used. The parameter h is the element size provided in gmsh.

Benchmark 1 (Plane wave). The first benchmark is a simple plane wave propagating in the unit square domain $\Omega =]0,1[\times]0,1[$. The reference solution reads

$$u_{\mathrm{ref}}(\mathbf{x}) = e^{\imath\kappa\mathbf{d}\cdot\mathbf{x}},$$

with the propagation direction $\mathbf{d} = (\cos\theta, \sin\theta)$ and a given angle θ. A non-homogeneous Robin condition is prescribed on the boundary of the domain (i.e. $\Gamma_{\mathrm{R}} := \partial\Omega$) with the appropriate right-hand side term. By default, the parameters are $\kappa = 15\pi$ and $h = 1/16$. We have also considered a wavenumber twice larger, $\kappa = 30\pi$, with a spatial step $h = 1/34$ corresponding to a relative error close to the one with the default parameters.

Benchmark 2 (Cavity). The second benchmark is a cavity problem. The computational domain is again the unit square domain $\Omega =]0,1[\times]0,1[$. A homogeneous Dirichlet condition is prescribed on the boundary of the domain (i.e. $\Gamma_{\mathrm{D}} := \partial\Omega$), and a unit source term is used in the Helmholtz equation:

$$\begin{cases} -\Delta u - \kappa^2 u = 1, & \text{in } \Omega, \\ u = 0, & \text{on } \Gamma_{\mathrm{D}}. \end{cases}$$

The reference solution is real. The eigenvalues and eigenmodes of this problem are $\kappa_{n,m}^2 := (n^2+m^2)\pi^2$ and $u_{n,m} := \sin(n\pi x_1)\sin(m\pi x_2)$, respectively, for all $m, n > 0$. The reference solution

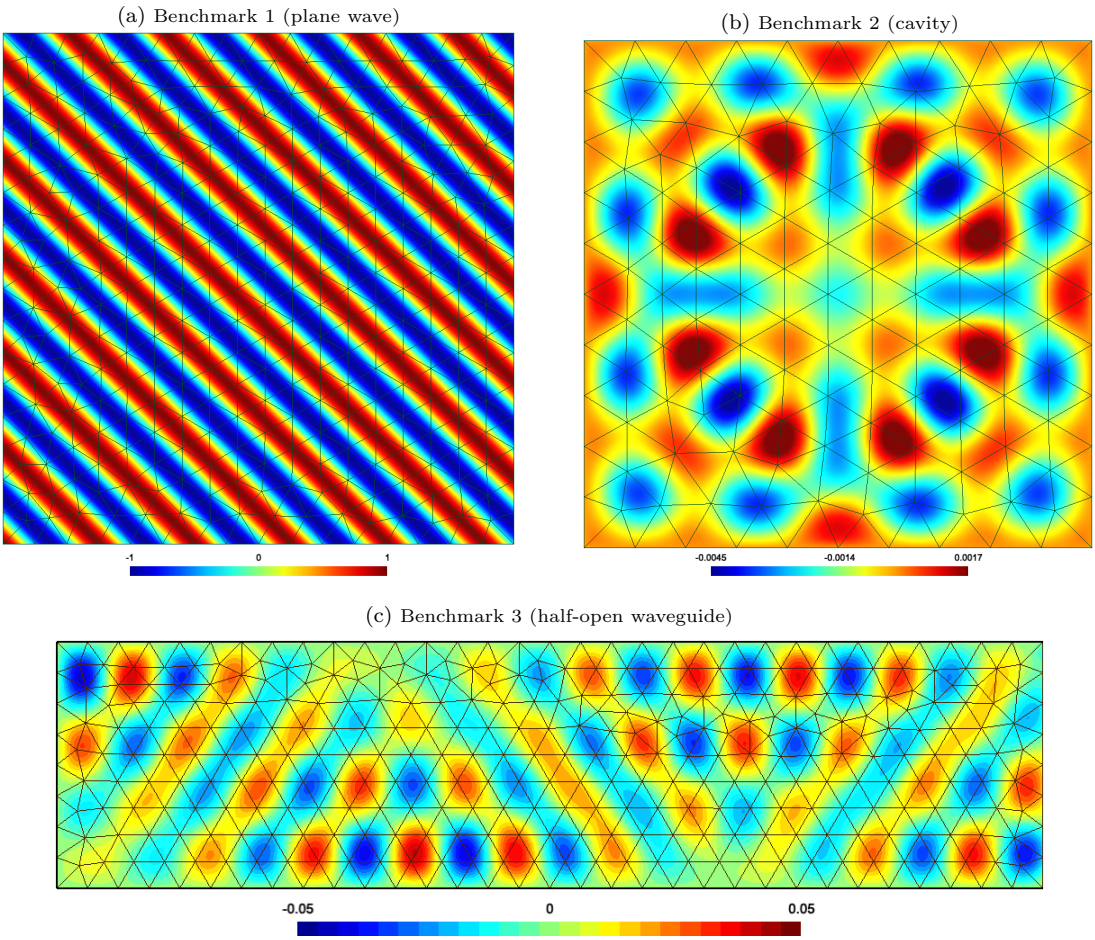

Figure 2: Snapshots of the real part of the solution for the three benchmarks with the default parameters.

is obtained semi-analytically by truncating the Fourier expansion (see e.g. [10]). By default, the parameters are $\kappa = (7 + 1/10)\sqrt{2}\pi$ and $h = 1/10$. We have also considered a wavenumber closer to an eigenvalue, $\kappa = (7 + 1/100)\sqrt{2}\pi$, with a spatial step $h = 1/15$ corresponding to a relative error close to the one with the default parameters.

Benchmark 3 (Waveguide). The last benchmark is a half open waveguide problem. The domain is $\Omega =]0, 4[\times]0, 1[$, with a given length L. The open side of the waveguide corresponds to the right side of Ω. An incident plane wave is prescribed at the open side by using a non-homogeneous Robin condition:

$$\partial_n u - \imath\kappa u = e^{\imath\kappa \mathbf{d}\cdot\mathbf{x}}, \quad \text{on } \Gamma_{\mathrm{R}} := \{4\} \times]0, 1[,$$

with the propagation direction $\mathbf{d} = (\cos\theta, \sin\theta)$ and a given angle θ. A homogeneous Dirichlet condition is prescribed on the other sides of Ω. The reference solution is computed by using a semi-analytical approach described in [10]. By default, the parameters are $\kappa = 6\pi$ and $h = 1/8$. We have also considered a wavenumber twice larger, $\kappa = 12\pi$, with a spatial step $h = 1/17$ corresponding to a relative error close to the one with the default parameters.

Table 1: Number of degrees of freedom (#dof) and number of non-zero entries (#nnz) in $\mathbf{A}$ for the different DG methods (i.e. standard DG without hybridization, HDG and CHDG).

		#dof	#nnz
Benchmark 1	DG	18420	734626
	HDG	3812	74192
	CHDG	7368	109082
Benchmark 2	DG	7260	286715
	HDG	1532	27970
	CHDG	2904	43830
Benchmark 3	DG	19260	765006
	HDG	4012	75149
	CHDG	7704	116202

4.3 Memory storage

The total numbers of degrees of freedom (DOFs) with the DG, HDG and CHDG methods are given respectively by

$$\begin{aligned}
\#(\mathrm{dof}_{\mathrm{DG}}) &= 3N_{\mathrm{tri}}N_{\mathrm{dof\cdot per\cdot tri}},\\
\#(\mathrm{dof}_{\mathrm{HDG}}) &= N_{\mathrm{fce}}N_{\mathrm{dof\cdot per\cdot fce}},\\
\#(\mathrm{dof}_{\mathrm{CHDG}}) &= 3N_{\mathrm{tri}}N_{\mathrm{dof\cdot per\cdot fce}} = \left(N_{\mathrm{fce\cdot bnd}} + 2N_{\mathrm{fce\cdot int}}\right)N_{\mathrm{dof\cdot per\cdot fce}},
\end{aligned}$$

with the number of faces N_{fce}, the number of boundary faces $N_{\mathrm{fce\cdot bnd}}$, the number of interior faces $N_{\mathrm{fce\cdot int}}$ and the number of triangles N_{tri}. Let us note that $N_{\mathrm{fce}} = N_{\mathrm{fce\cdot bnd}} + N_{\mathrm{fce\cdot int}}$ and $3N_{\mathrm{tri}} = N_{\mathrm{fce\cdot bnd}} + 2N_{\mathrm{fce\cdot int}}$. For a scalar field, the numbers of DOFs per triangle and per face are given respectively by $N_{\mathrm{dof\cdot per\cdot tri}} = (p+1)(p+2)/2$ and $N_{\mathrm{dof\cdot per\cdot fce}} = p+1$, where p is the polynomial degree.

The number of DOFs is obviously far smaller with the hybridizable methods. It is nearly twice larger with CHDG than with HDG because there are two characteristic variables per interior face and only one numerical trace. The results would be similar in three dimensions.

Upper bounds for the numbers of non-zero elements in the global sparse matrix $\mathbf{A}$ of the DG, HDG and CHDG systems are given respectively by

$$\begin{aligned}
\#(\mathrm{nnz}_{\mathrm{DG}}) &\lesssim N_{\mathrm{tri}}\left(7N^2_{\mathrm{dof\cdot per\cdot tri}} + 54N^2_{\mathrm{dof\cdot per\cdot fce}}\right),\\
\#(\mathrm{nnz}_{\mathrm{HDG}}) &\lesssim N_{\mathrm{fce}}\left(5N^2_{\mathrm{dof\cdot per\cdot fce}}\right),\\
\#(\mathrm{nnz}_{\mathrm{CHDG}}) &\lesssim N_{\mathrm{fce}}\left(8N^2_{\mathrm{dof\cdot per\cdot fce}}\right).
\end{aligned}$$

For the hybridizable methods, the matrix $\mathbf{A}$ is obtained after the elimination of the physical unknowns. These bounds have been computed by using the rough approximation $N_{\mathrm{fce\cdot bnd}} \ll N_{\mathrm{fce\cdot int}}$, which is valid only for large benchmarks. Under this approximation, we have

$$\frac{\#(\mathrm{nnz}_{\mathrm{CHDG}})}{\#(\mathrm{nnz}_{\mathrm{HDG}})} \approx 1.6.$$

For the matrices of the reference benchmarks with the default parameters, this ratio varies between 1.54 and 1.66 (see Table 1). For three-dimensional problems with tetrahedral elements, a similar reasoning leads to a ratio equal to 1.43. Therefore, although there are nearly twice as many DOFs with CHDG than with HDG, the number of non-zero elements is not increased as much.

4.4 Conditioning of the local matrices

With the hybridizable approaches, the construction of the matrix $\mathbf{A}$, and the application of $\mathbf{A}$ in matrix-free iterative procedures, requires the solution of local element-wise algebraic systems. For

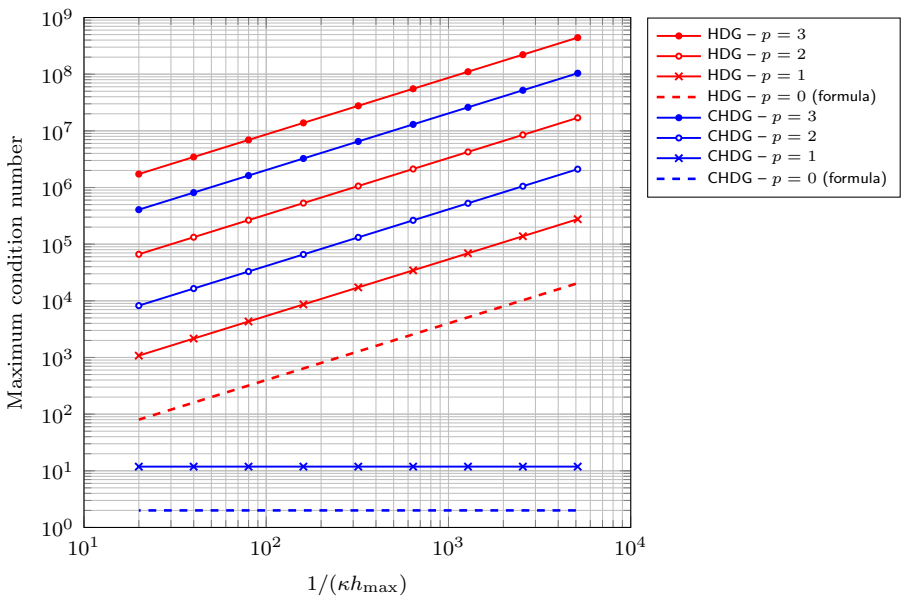

Figure 3: Maximum condition number of the local matrices with HDG (red curves) and CHDG (blue curves) as a function of $1/(\kappa h_{\max})$ for basis functions with polynomial degrees $p = 1, 2$ and 3, where $h_{\max}$ is the length of the longest edge. The condition numbers corresponding to formulas (11) and (12) are plotted with dashed lines.

the HDG and CHDG methods, these systems correspond to Problems 2.4 and 2.8, respectively. A bad conditioning of these systems could impact the quality of the numerical solution, regardless of the solution procedure.

As a preliminary study of the conditioning of the local systems, we first consider an elementary configuration used in [34]. The local systems are defined on a square element K of side length h with the lowest polynomial degree, i.e. $p = 0$. With the HDG method, the local matrix corresponding to Problems 2.4 with the shape functions $\phi_1 = 1$, $\boldsymbol{\phi}_1 = [1, 0]^\top$ and $\boldsymbol{\phi}_2 = [0, 1]^\top$ reads

$$\mathbf{A}_{\text{loc}} = \text{diag}(4h - \imath\kappa h^2, -\imath\kappa h^2, -\imath\kappa h^2)$$

and the condition number of this matrix is

$$\text{cond}(\mathbf{A}_{\text{loc}}) = \sqrt{1 + 16/(\kappa h)^2}. \tag{11}$$

With the CHDG method, the local matrix corresponding to Problem 2.8 reads

$$\mathbf{A}_{\text{loc}} = \text{diag}(2h - \imath\kappa h^2, h - \imath\kappa h^2, h - \imath\kappa h^2)$$

and the condition number of this matrix is

$$\text{cond}(\mathbf{A}_{\text{loc}}) = \sqrt{((\kappa h)^2 + 4)/((\kappa h)^2 + 1)}. \tag{12}$$

The condition number of the HDG local matrix is always the largest. In addition, this matrix becomes ill-conditioned as κh goes to zero, whereas the CHDG local matrix stays well-conditioned with $\text{cond}(\mathbf{A}_{\text{loc}}) \approx 2$ for small values of κh. Although this simple setting is not representative of practical situations, it already highlights the influence of the variables used in the hybridization on the conditioning of the local matrices.

To continue the study, we consider a non-structured mesh for the unit square $\Omega =]0, 1[^2$. This mesh is made of 1478 triangles and the length of the longest edge is close to $h_{\max} = 0.05$. The condition number of the corresponding local element-wise systems is computed for both HDG and CHDG, with different polynomial degrees $p = 1, 2, 3$ and different wavenumbers κ.

The maximum condition number is plotted as a function of $1/(\kappa h_{\max})$ on Figure 3 for the different configurations. The value $1/(\kappa h_{\max})$ is a measure of the mesh density in the denser region of the mesh. We observe that the condition number increases linearly with $1/(\kappa h_{\max})$ in all the cases, except for CHDG with $p = 1$. Therefore, refining the mesh for a given wavenumber, or using a smaller wavenumber with a given mesh, increases the condition number of the local matrices. Comparing the results with the HDG and CHDG methods for a given polynomial degree p, we observe that the condition number is always higher with HDG than with CHDG. Increasing p increases the condition number in all the cases.

4.5 Conditioning of the global matrices

The condition number of the global matrix $\mathbf{A}$ is plotted a function of $1/(\kappa h_{\max})$ for the DG, HDG and CHDG methods on Figure 4. For each benchmark, two wavenumbers have been considered: the default wavenumber of the benchmark (denoted κ_1), and a second wavenumber corresponding to a more challenging case (denoted κ_2). The second wavenumber is twice larger for benchmarks 1 and 3, and closer to a resonance mode for benchmark 2. The condition number has been computed with the function `condest` in MATLAB. For all the results, the relative error on the numerical solution is smaller than 10^{-1}. The black squares correspond to configurations with a relative error close to 10^{-2}.

We observe on Figure 4 that the condition number is always smaller with CHDG than with HDG and DG by one or two orders of magnitude in nearly all the cases. Moreover, the condition number increases nearly linearly with $1/(\kappa h_{\max})$ for DG and CHDG, while the increase is nearly quadratic for HDG.

The influence of κ on the condition number is similar for HDG and CHDG. Indeed, for each benchmark, the condition number is larger with the larger wavenumber. By contrast, the condition number for the DG method without hybridization does not vary much with κ.

5 Iterative solution procedures

In this section, we study the efficiency of iterative procedures for solving the linear systems resulting from the DG discretization and the two hybridization strategies. With the CHDG approach, the fixed-point iterative procedure can be considered thanks to the specific structure of the global matrix, which we analyzed in Section 3. The convergence of the fixed-point iterative scheme with CHDG is discussed in Section 5.1. The performance of DG, HDG and CHDG with standard iterative schemes is discussed in Section 5.2.

5.1 Convergence of the fixed-point iterative scheme for CHDG

We consider the algebraic system obtained by using the CHDG approach (Problem 3.2) with the discretization described in Section 4.1. This CHDG system can be written as

$$(\mathbf{I} - \mathbf{\Pi S})\mathbf{g} = \mathbf{b},$$

where $\mathbf{I}$, $\mathbf{\Pi}$ and $\mathbf{S}$ are the identity, exchange and scattering matrices, respectively. As the operator ΠS is a strict contraction (Corollary 3.6), the spectral radius of $\mathbf{\Pi S}$ is strictly lower than 1, i.e. $\rho(\mathbf{\Pi S}) < 1$. Therefore, the Richardson iterative scheme applied to this system shall converge without relaxation (see e.g. [56]). For a given initial guess $\mathbf{g}^{(0)}$, the procedure reads

$$\mathbf{g}^{(\ell+1)} = \mathbf{\Pi S}\mathbf{g}^{(\ell)} + \mathbf{b}, \qquad \text{for } \ell = 0, 1, \ldots$$

If the eigenvalues of the iteration operator are far from the unit disk, this procedure will converge rapidly. As discussed in Section 3.2, this will depend on both the dissipative properties of the upwind DG scheme and on the physical dissipation in the problem under consideration.

As a preliminary verification, we discuss the eigenvalues of the iteration matrix $\mathbf{\Pi S}$ and the spectral radius $\rho(\mathbf{\Pi S})$ by using the numerical benchmarks. The eigenvalues of the iteration matrix

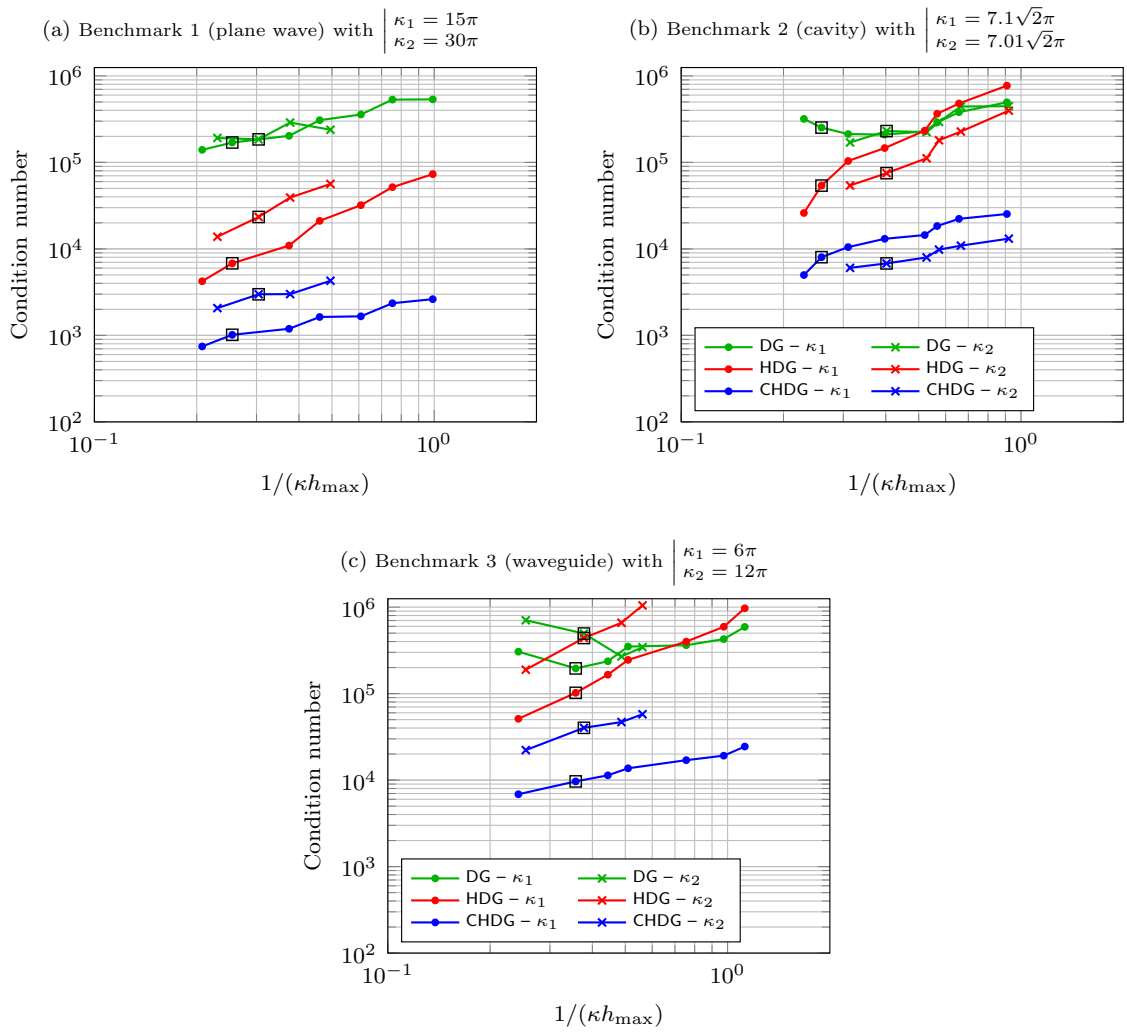

Figure 4: Condition number of $\mathbf{A}$, the matrix of the physical system (for DG) or the matrix of the reduced system (for HDG and CHDG), as a function of $1/(\kappa h_{\max})$ for the three benchmarks, where $h_{\max}$ is the length of the longest edge. For each benchmark, two wavenumbers are considered, κ_1 and κ_2. The black squares correspond to configurations with a relative error close to 10^{-2}.

Table 2: Spectral radius ρ of the iteration matrix $\mathbf{\Pi S}$ of the fixed-point iterative scheme for the three benchmarks with different parameters and the CHDG method.

	Benchmark 1 (plane wave)			Benchmark 2 (cavity)			Benchmark 3 (waveguide)		
κ	15π	15π	30π	$7.1\sqrt{2}\pi$	$7.1\sqrt{2}\pi$	$7.01\sqrt{2}\pi$	6π	6π	12π
h	1/16	1/34	1/34	1/10	1/15	1/15	1/8	1/17	1/17
κh	2.95	1.39	2.77	3.15	2.10	2.08	2.36	1.11	2.22
$1-\rho(\mathbf{\Pi S})$	$2.9\ 10^{-3}$	$7.8\ 10^{-5}$	$5.5\ 10^{-4}$	$2.8\ 10^{-4}$	$1.5\ 10^{-5}$	$1.4\ 10^{-5}$	$5.5\ 10^{-5}$	$2.5\ 10^{-6}$	$2.9\ 10^{-5}$

are represented on Figure 5 for the three benchmarks with the default parameters. The values of $1-\rho(\mathbf{\Pi S})$ are given in Table 2 for different sets of parameters. The eigenvalues and the spectral radius are obtained by using the function `eigs` in MATLAB.

In all the cases, the eigenvalues are strictly inside the unit circle, which is in agreement with the theoretical result. We shall also observe in the next section that the iterative process effectively converges. Nevertheless, some eigenvalues are close to the unit circle, so that the spectral radius is

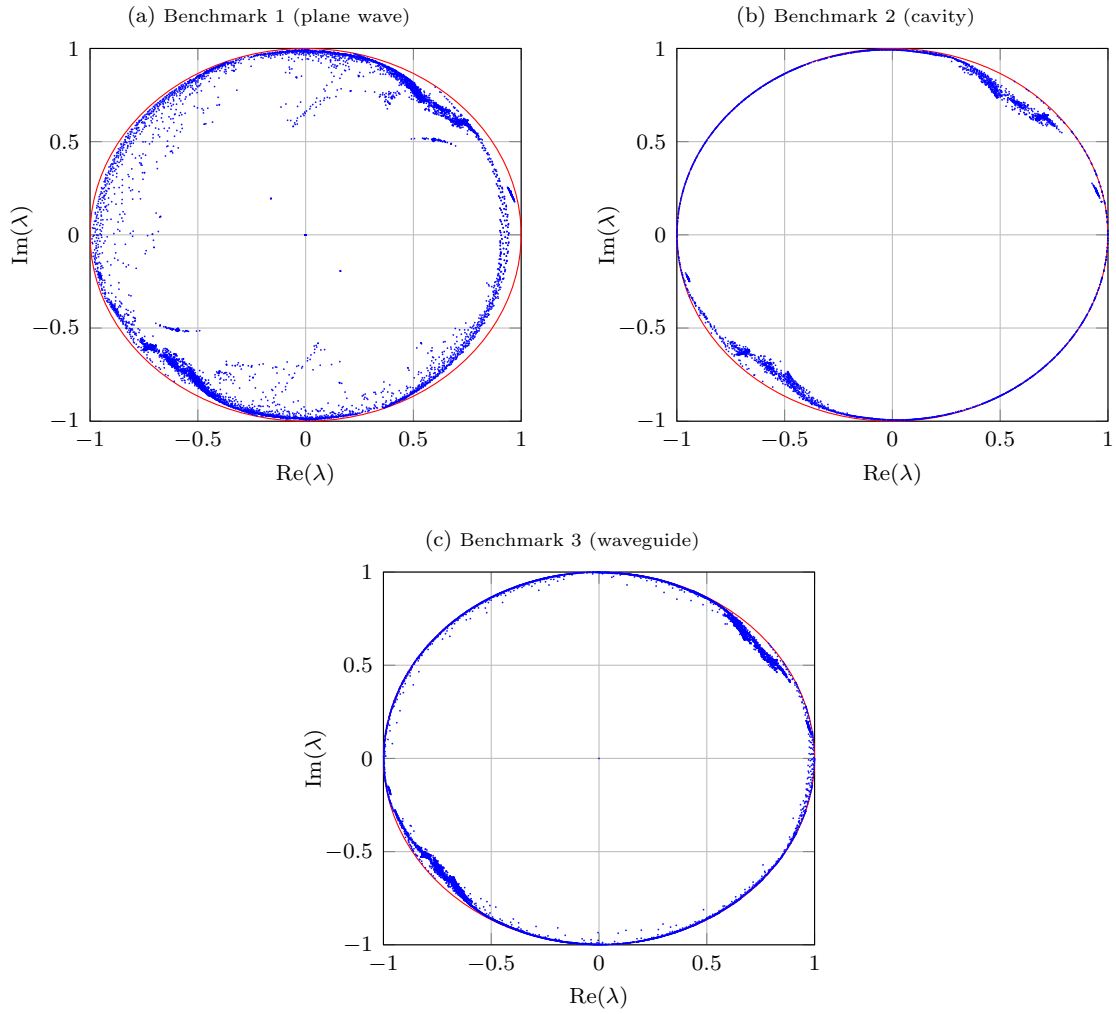

Figure 5: Spectrum of the iteration matrix $\mathbf{\Pi S}$ of the fixed-point iterative scheme for the three benchmarks with the default parameters and the CHDG method. The unit circle is plotted in red.

close to one. For every benchmark, we observe that the spectral radius is closer to one when using a finer mesh (second column of each benchmark in Table 2) or when using the second wavenumber with the fine mesh (third column).

5.2 Comparison of DG, HDG and CHDG with standard iterative schemes

In practice, the iterative procedures to solve large-scale time-harmonic problems can be rather sophisticated, because the corresponding algebraic linear systems are generally non-Hermitian and ill-conditioned. The GMRES *(generalized minimal residual)* method with restart and preconditioning strategies is one of the most widely used approach. For the standard version without restart, the convergence is guaranteed, but the computational cost increases with the number iterations, both in terms of memory storage and floating-point operations. Alternative Krylov methods are frequently considered, with smaller computational cost per iteration and smaller memory footprint, but at the price of a larger number of iterations and/or a convergence that is not always guaranteed.

For the sake of brevity, we only consider three standard iterative schemes to compare the DG methods: the fixed-point iterative scheme (for CHDG only), the GMRES iteration without

restart and the CGNR *(conjugate gradient normal)* method. The CGNR iteration corresponds to the conjugate gradient method applied to the normal equation $\mathbf{A}^*\mathbf{A}\mathbf{g} = \mathbf{A}^*\mathbf{b}$. For a given initial solution $\mathbf{g}^{(0)}$, both GMRES and CGNR produce an approximate solution $\mathbf{g}^{(\ell)}$ at step ℓ that belongs to a certain Krylov subspace and that minimizes the 2-norm of the residual, i.e. $\mathbf{g}^{(\ell)}$ minimizes $f(\mathbf{g}) = \|\mathbf{b} - \mathbf{A}\mathbf{g}\|_2$. The approximate solution belongs to $\mathbf{g}^{(0)} + \mathcal{K}_\ell(\mathbf{A}, \mathbf{r}^{(0)})$ with GMRES and to $\mathbf{g}^{(0)} + \mathcal{K}_\ell(\mathbf{A}^*\mathbf{A}, \mathbf{A}^*\mathbf{r}^{(0)})$ with CGNR, where $\mathcal{K}_\ell$ is the Krylov subspace of order ℓ (see e.g. [56]). The convergence rate of the CGNR iterative process depends on the condition number of $\mathbf{A}$. The convergence can be slow if the condition number is large. Nevertheless, we have observed that the condition number is nearly always smaller with CHDG than with the other approaches (see Section 4.5).

To study the efficiency of the iterative schemes with the different methods, we consider the relative error of the physical fields defined as

$$\sqrt{\frac{\|u_h - u_{\text{ref}}\|_\Omega^2 + \|\mathbf{q}_h - \mathbf{q}_{\text{ref}}\|_\Omega^2}{\|u_{\text{ref}}\|_\Omega^2 + \|\mathbf{q}_{\text{ref}}\|_\Omega^2}},$$

where u_{ref} and $\mathbf{q}_{\text{ref}}$ correspond to the reference analytical or semi-analytical solution. The history of relative error is plotted in Figure 6 for CGNR (lines with marker $\circ$), GMRES (lines with marker $\bullet$) and the fixed-point iteration in the CHDG case (lines with marker $\times$). The results have been obtained for DG without hybridization (green lines), HDG (red lines) and CHDG (blue lines). The relative error obtained with a direct solver is indicated by the horizontal dashed line.

First, let us analyze the results obtained with CHDG and fixed-point iterations (blue lines with marker $\times$). The following observations can be made:

- For benchmark 1 (plane wave), the convergence of the iterative process is very fast. The decay of error is slightly slower with the higher wavenumber. Compared to the other approaches, CHDG with fixed-point iterations provides nearly the fastest convergence.

- By contrast, for benchmark 2 (cavity), the convergence of the fixed-point iterations is very slow. This can be explained by the fact that this benchmark does not feature any physical absorption. Therefore, as discussed in Section 3.2, the only source of dissipation comes from the DG scheme. The decay of error is much slower for the wavenumber closer to the resonance. Compared to the other methods, this approach provides the slowest convergence.

- For benchmark 3 (half-open waveguide) with the first set of parameters (Figure 6e), the relative error decays relatively rapidly during the 500 first iterations, then the decay slows down dramatically, and the relative error is only about 10^{-1} at iteration 4,000. With the higher wavenumber (Figure 6f), the relative error decays more rapidly until approximately 10^{-2} at iteration 4,000.

- The asymptotic regime of convergence has been reached in three cases, and the slopes of error decay are coherent to the spectral radii obtained in Table 2: $\rho = 1 - 2.8\,10^{-4}$ for Figure 6c, $\rho = 1 - 1.5\,10^{-5}$ for Figure 6d, and $\rho = 1 - 5.5\,10^{-5}$ for Figure 6e. The asymptotic regime starts at the beginning of the iterations in the cavity case.

To summarize, the fixed-point iterative process effectively converges for CHDG, but the performance strongly depends on the physical setting. The convergence can be very fast for purely propagating cases, and very slow for cavity or waveguide cases. In the latter cases, the asymptotic regime, which can start relatively quickly, is rather slow.

We then discuss the convergence of the CGNR and GMRES schemes with the different approaches, i.e. DG without hybridization, HDG and CHDG. We can make the following comments:

- When using CGNR (lines with marker $\circ$ on Figure 6), the convergence is much faster with CHDG than with HDG and DG without hybridization in all the cases. Comparing the last two approaches, the convergence is faster with HDG than with DG without hybridization on Figures 6a, 6b and 6c, and the converse is true on Figures 6d, 6e and 6f.

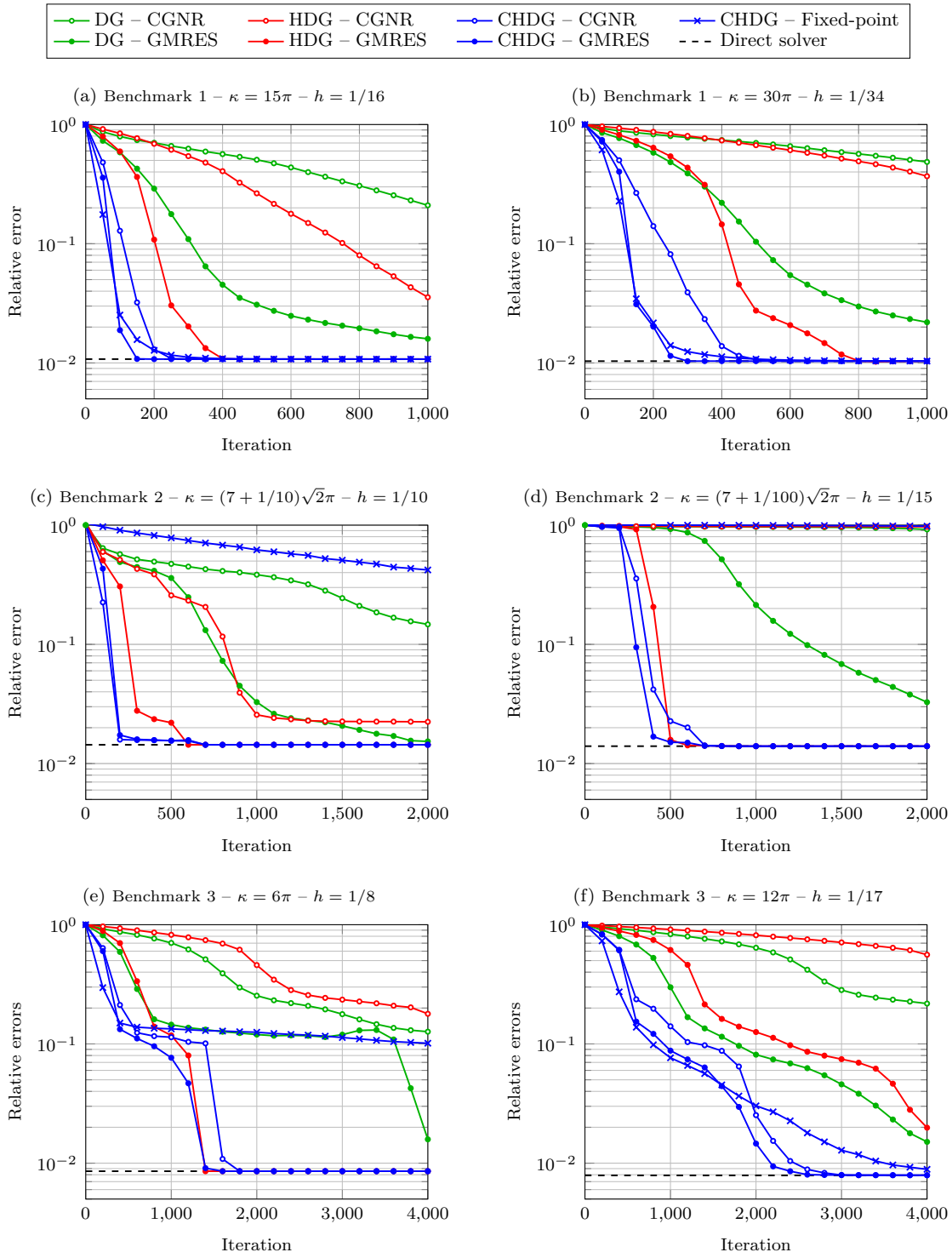

Figure 6: Error history for the three benchmarks with different iterative schemes and different DG schemes. The dashed horizontal lines correspond to the relative errors obtained with a direct solver.

- When using GMRES (lines with marker • on Figure 6), the fastest convergence is still obtained with CHDG in all the cases, but the convergence is rather close with HDG for the cavity benchmark (Figures 6c-6d) and the first waveguide benchmark (Figures 6e). The convergence is generally faster with HDG than with DG without hybridization, but the converse is true for the second waveguide benchmark (Figures 6f).

To summarize, if the problem is solved with either CGNR or GMRES, the convergence of the iterative process is always faster with the CHDG method. Using the standard HDG method generally speeds up the convergence in comparison with the DG method without hybridization, but the converse is true for several cases.

Finally, let us compare the performance of CGNR and GMRES when the CHDG method is used (blue lines with markers ∘ and • on Figure 6). The convergence is always slightly faster with GMRES than with CGNR, but the difference is not very large. In the worst case (Figure 6b), the number of iterations to achieve the reference relative error (obtained with the direct solver) is twice larger with CGNR than with GMRES. Considering the computational cost of GMRES, which increases at each iteration, the CGNR is a potential good candidate for realistic cases. The complete analysis of the runtimes and computational costs, which depend on the implementation, will be performed in future works

6 Conclusion

In this work, we propose a new hybridization technique, which we call the CHDG method, for solving time-harmonic problems with upwind DG discretizations. The auxiliary unknowns used in the CHDG method correspond to characteristic variables, whereas the auxiliary unknowns used in the standard approach correspond to a Dirichlet trace. At the price of increasing the required memory storage for the reduced linear system, this choice largely improves its properties and makes it more suitable for iterative solution procedures.

We study the properties of the local element-wise problems and the global linear systems for the standard HDG method and the CHDG method. In order to investigate how the original DG scheme and its hybridized versions interplay with usual iterative solvers, we provide a set of 2D numerical results where the auxiliary unknowns are discretized with scaled Legendre basis functions. The key properties of the CHDG may be summarized as follows.

With CHDG, the reduced system can be written in the form $(\mathrm{I} - \Pi\mathrm{S})g = b$, where the operator $\Pi\mathrm{S}$ is a strict contraction. It can be solved with a fixed-point iteration without relaxation. This fixed-point iteration converges quickly in open domains, but unfortunately, the convergence becomes slow when waves are trapped, like in waveguides or cavities.

The memory storage required to store an unknown vector of the reduced system is twice larger with CHDG than with the standard HDG method. Similarly, the number of non-zero entries in the CHDG matrix is multiplied by about 1.6 in 2D and 1.4 in 3D as compared to the HDG matrix, with a similar filling pattern. In return, the condition number of the matrices of the local element-wise systems is always smaller with CHDG than with the standard HDG method. Similarly, the condition number of the global reduced matrix is also always smaller with CHDG than with HDG. It is also smaller than the condition number of the global matrix of the DG system without hybridization.

For the iterative solution procedure, we have employed the usual GMRES iteration (without restart) and the CGNR iteration. In both cases, the convergence of the iterative process is always faster with CHDG than with HDG or DG without hybridization. Focusing on the CHDG system, the number of CGNR iterations is always larger than the number of GMRES iterations, but the difference is rather limited for the benchmarks considered in this article. Since restart must be employed for GMRES in practice, and since each GMRES iteration is typically more costly than the corresponding CGNR iteration, we believe that CGNR may be a competitive approach to solve the CHDG system.

Although we focus on 2D benchmarks here, the definition of the method is valid for 3D cases. Besides, the method is in principle not restricted to scalar problems, and electromagnetic or elastic

waves should be accessible as well because similar DG schemes with upwind fluxes are already available. In future works, we will investigate in more depth the computational aspects for solving iteratively 3D cases, high-order transmission conditions, and combinations with preconditioning techniques and domain decomposition methods to accelerate further the convergence of the iterative solution procedures.

Acknowledgments. This work was supported in part by the ANR JCJC project *WavesDG* (research grant ANR-21-CE46-0010). The authors thank X. Antoine, H. Bériot, X. Claeys, G. Gabard, C. Geuzaine and S. Pescuma for helpful discussions and remarks.

References

[1] M. Ainsworth. Discrete dispersion relation for *hp*-version finite element approximation at high wave number. *SIAM Journal on Numerical Analysis*, 42(2):553–575, 2004.

[2] M. Ainsworth, P. Monk, and W. Muniz. Dispersive and dissipative properties of discontinuous Galerkin finite element methods for the second-order wave equation. *Journal of Scientific Computing*, 27(1–3), 2006.

[3] H. Barucq, A. Bendali, J. Diaz, and S. Tordeux. Local strategies for improving the conditioning of the plane-wave Ultra-Weak Variational Formulation. *Journal of Computational Physics*, 441:110449, 2021.

[4] H. Barucq, J. Diaz, R.-C. Meyer, and H. 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.

[5] H. Barucq, N. Rouxelin, and S. Tordeux. Construction and analysis of a HDG solution for the total-flux formulation of the convected Helmholtz equation. *Mathematics of Computation*, 92(343): 2097–2131, 2023.

[6] H. Bériot, A. Prinn, and G. Gabard. Efficient implementation of high-order finite elements for Helmholtz problems. *International Journal for Numerical Methods in Engineering*, 106(3):213–240, 2016.

[7] N. Bootland, V. Dolean, P. Jolivet, and P.-H. Tournier. A comparison of coarse spaces for Helmholtz problems in the high frequency regime. *Computers & Mathematics with Applications*, 98:239–253, 2021.

[8] Y. Boubendir, X. Antoine, and C. Geuzaine. A quasi-optimal non-overlapping domain decomposition algorithm for the Helmholtz equation. *Journal of Computational Physics*, 231(2):262–280, 2012.

[9] O. Cessenat and B. Despres. Application of an ultra weak variational formulation of elliptic PDEs to the two-dimensional Helmholtz problem. *SIAM journal on numerical analysis*, 35(1):255–299, 1998.

[10] T. Chaumont-Frelet, M. J. Grote, S. Lanteri, and J. H. Tang. A controllability method for maxwell's equations. *SIAM Journal on Scientific Computing*, 44(6):A3700–A3727, 2022.

[11] H. Chen, P. Lu, and X. Xu. A hybridizable discontinuous Galerkin method for the Helmholtz equation with high wave number. *SIAM Journal on Numerical Analysis*, 51(4):2166–2188, 2013.

[12] X. Claeys. Non-local variant of the optimised Schwarz method for arbitrary non-overlapping subdomain partitions. *ESAIM: Mathematical Modelling and Numerical Analysis*, 55(2):429–448, 2021.

[13] X. Claeys and E. Parolin. Robust treatment of cross-points in optimized Schwarz methods. *Numerische Mathematik*, pages 1–38, 2022.

[14] B. Cockburn. Static condensation, hybridization, and the devising of the hdg methods. *Building bridges: connections and challenges in modern approaches to numerical partial differential equations*, pages 129–177, 2016.

[15] B. Cockburn, B. Dong, and J. Guzmán. A superconvergent LDG-hybridizable Galerkin method for second-order elliptic problems. *Mathematics of Computation*, 77(264):1887–1916, 2008.

[16] B. Cockburn, J. Gopalakrishnan, and R. Lazarov. Unified hybridization of discontinuous galerkin, mixed, and continuous Galerkin methods for second order elliptic problems. *SIAM Journal on Numerical Analysis*, 47(2):1319–1365, 2009.

[17] F. Collino, S. Ghanemi, and P. Joly. Domain decomposition method for harmonic wave propagation: a general presentation. *Computer methods in applied mechanics and engineering*, 184(2-4):171–211, 2000.

[18] F. Collino, P. Joly, and M. Lecouvez. Exponentially convergent non overlapping domain decomposi-

tion methods for the Helmholtz equation. *ESAIM: Mathematical Modelling and Numerical Analysis*, 54(3):775–810, 2020.

[19] J. Cui and W. Zhang. An analysis of HDG methods for the Helmholtz equation. *IMA Journal of Numerical Analysis*, 34(1):279–295, 2014.

[20] R. Dai, A. Modave, J.-F. Remacle, and C. Geuzaine. Multidirectional sweeping preconditioners with non-overlapping checkerboard domain decomposition for Helmholtz problems. *Journal of Computational Physics*, 453:110887, 2022.

[21] B. Després. *Une méthodes de décomposition de domaine pour les problèmes de propagation d'ondes en régime harmonique. Le théorème de Borg pour l'équation de Hill vectorielle.* PhD thesis, Université de Paris-IX, 1991.

[22] Y. A. Erlangga, C. Vuik, and C. W. Oosterlee. On a class of preconditioners for solving the Helmholtz equation. *Applied Numerical Mathematics*, 50(3-4):409–425, 2004.

[23] O. G. Ernst and M. J. Gander. Why it is difficult to solve Helmholtz problems with classical iterative methods. In *Numerical analysis of multiscale problems*, pages 325–363. Springer, 2012.

[24] C. Farhat, R. Tezaur, and J. Toivanen. A domain decomposition method for discontinuous Galerkin discretizations of Helmholtz problems with plane waves and Lagrange multipliers. *International journal for numerical methods in engineering*, 78(13):1513–1531, 2009.

[25] F. Faucher and O. 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, 2020.

[26] X. Feng and Y. Xing. Absolutely stable local discontinuous Galerkin methods for the Helmholtz equation with large wave number. *Mathematics of Computation*, 82(283):1269–1296, 2013.

[27] H. S. Fure, S. Pernet, M. Sirdey, and S. Tordeux. A discontinuous Galerkin Trefftz type method for solving the two dimensional Maxwell equations. *SN Partial Differential Equations and Applications*, 1:1–25, 2020.

[28] G. Gabard. Discontinuous Galerkin methods with plane waves for time-harmonic problems. *Journal of Computational Physics*, 225(2):1961–1984, 2007.

[29] M. Gander, F. Magoules, and F. Nataf. Optimized Schwarz methods without overlap for the Helmholtz equation. *SIAM Journal on Scientific Computing*, 24(1):38–60, 2002.

[30] M. J. Gander and H. Zhang. A class of iterative solvers for the Helmholtz equation: Factorizations, sweeping preconditioners, source transfer, single layer potentials, polarized traces, and optimized Schwarz methods. *SIAM Review*, 61(1):3–76, 2019.

[31] M. J. Gander and H. Zhang. Schwarz methods by domain truncation. *Acta Numerica*, 31:1–134, 2022.

[32] C. Geuzaine and J.-F. Remacle. Gmsh: A 3-D finite element mesh generator with built-in pre- and post-processing facilities. *International journal for numerical methods in engineering*, 79(11): 1309–1331, 2009.

[33] G. Giorgiani, S. Fernández-Méndez, and A. Huerta. Hybridizable discontinuous Galerkin p-adaptivity for wave propagation problems. *International Journal for Numerical Methods in Fluids*, 72(12):1244–1262, 2013.

[34] J. Gopalakrishnan, S. Lanteri, N. Olivares, and R. Perrussel. Stabilization in relation to wavenumber in HDG methods. *Advanced Modeling and Simulation in Engineering Sciences*, 2(1):1–24, 2015.

[35] R. Griesmaier and P. Monk. Error analysis for a hybridizable discontinuous Galerkin method for the Helmholtz equation. *Journal of Scientific Computing*, 49(3):291–310, 2011.

[36] J. S. Hesthaven and T. Warburton. *Nodal discontinuous Galerkin methods: algorithms, analysis, and applications.* Springer Science & Business Media, 2007.

[37] M. Huber and J. Schöberl. Hybrid domain decomposition solvers for the Helmholtz equation. In *Domain Decomposition Methods in Science and Engineering XXI*, pages 351–358. Springer, 2014.

[38] A. Huerta, A. Angeloski, X. Roca, and J. Peraire. Efficiency of high-order elements for continuous and discontinuous Galerkin methods. *International Journal for numerical methods in Engineering*, 96(9):529–560, 2013.

[39] T. Huttunen, P. Monk, and J. P. Kaipio. Computational aspects of the ultra-weak variational formulation. *Journal of Computational Physics*, 182(1):27–46, 2002.

[40] L.-M. Imbert-Gerard and G. Sylvand. Three types of quasi-Trefftz functions for the 3D convected Helmholtz equation: construction and approximation properties. *arXiv preprint arXiv:2201.12993*, 2023.

[41] A. Karakus, N. Chalmers, K. Świrydowicz, and T. Warburton. A GPU accelerated discontinuous Galerkin incompressible flow solver. *Journal of Computational Physics*, 390:380–404, 2019.

[42] G. E. Karniadakis, G. Karniadakis, and S. Sherwin. *Spectral/hp element methods for computational fluid dynamics*. Oxford University Press on Demand, 2005.
[43] A. Klöckner, T. Warburton, J. Bridge, and J. S. Hesthaven. Nodal discontinuous Galerkin methods on graphics processors. *Journal of Computational Physics*, 228(21):7863–7882, 2009.
[44] L. Li, S. Lanteri, and R. Perrussel. Numerical investigation of a high order hybridizable discontinuous Galerkin method for 2d time-harmonic Maxwell's equations. *COMPEL*, 32(3):1112–1138, 2013.
[45] L. Li, S. Lanteri, and R. Perrussel. A hybridizable discontinuous Galerkin method combined to a Schwarz algorithm for the solution of 3D time-harmonic Maxwell's equation. *Journal of Computational Physics*, 256:563–581, 2014.
[46] J. Melenk and S. Sauter. Wavenumber explicit convergence analysis for Galerkin discretizations of the Helmholtz equation. *SIAM Journal on Numerical Analysis*, 49(3):1210–1243, 2011.
[47] A. Modave, A. St-Cyr, and T. Warburton. GPU performance analysis of a nodal discontinuous Galerkin method for acoustic and elastic models. *Computers & Geosciences*, 91:64–76, 2016.
[48] A. Modave, A. Royer, X. Antoine, and C. Geuzaine. A non-overlapping domain decomposition method with high-order transmission conditions and cross-point treatment for Helmholtz problems. *Computer Methods in Applied Mechanics and Engineering*, 368:113162, 2020.
[49] P. Monk, J. Schöberl, and A. Sinwel. Hybridizing Raviart-Thomas elements for the Helmholtz equation. *Electromagnetics*, 30(1-2):149–176, 2010.
[50] F. Nataf, F. Rogier, and E. de Sturler. Optimal interface conditions for domain decomposition methods. Technical report, CMAP (Ecole Polytechnique), 1994.
[51] N. C. Nguyen, J. Peraire, and B. Cockburn. High-order implicit hybridizable discontinuous Galerkin methods for acoustics and elastodynamics. *Journal of Computational Physics*, 230(10):3695–3718, 2011.
[52] E. Parolin, D. Huybrechs, and A. Moiola. Stable approximation of Helmholtz solutions by evanescent plane waves. *arXiv preprint arXiv:2202.05658*, 2022.
[53] C. Pechstein. A unified theory of non-overlapping Robin–Schwarz methods: Continuous and discrete, including cross points. *Journal of Scientific Computing*, 96(2):60, 2023.
[54] S. Pernet, M. Sirdey, and S. Tordeux. Ultra-weak variational formulation for heterogeneous Maxwell problem in the context of high performance computing. *HAL preprint hal-03642116*, 2022.
[55] A. Royer, C. Geuzaine, E. Béchet, and A. Modave. A non-overlapping domain decomposition method with perfectly matched layer transmission conditions for the Helmholtz equation. *Computer Methods in Applied Mechanics and Engineering*, 395:115006, 2022.
[56] Y. Saad. *Iterative methods for sparse linear systems*. SIAM, 2003.
[57] P. Solin, K. Segeth, and I. Dolezel. *Higher-order finite element methods*. CRC Press, 2003.
[58] M. Taus, L. Zepeda-Núñez, R. J. Hewett, and L. Demanet. L-Sweeps: A scalable, parallel preconditioner for the high-frequency Helmholtz equation. *Journal of Computational Physics*, 420:109706, 2020.
[59] E. F. Toro. *Riemann solvers and numerical methods for fluid dynamics: a practical introduction*. Springer Science & Business Media, 2013.