ESAIM: M2AN 58 (2024) 1087–1106
https://doi.org/10.1051/m2an/2024019

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

DUALITY ANALYSIS OF INTERIOR PENALTY DISCONTINUOUS GALERKIN METHODS UNDER MINIMAL REGULARITY AND APPLICATION TO THE *A PRIORI* AND *A POSTERIORI* ERROR ANALYSIS OF HELMHOLTZ PROBLEMS

Théophile Chaumont-Frelet*

Abstract. We consider interior penalty discontinuous Galerkin discretizations of time-harmonic wave propagation problems modeled by the Helmholtz equation, and derive novel *a priori* and *a posteriori* estimates. Our analysis classically relies on duality arguments of Aubin–Nitsche type, and its originality is that it applies under minimal regularity assumptions. The estimates we obtain directly generalize known results for conforming discretizations, namely that the discrete solution is optimal in a suitable energy norm and that the error can be explicitly controlled by *a posteriori* estimators, provided the mesh is sufficiently fine.

Mathematics Subject Classification. 35J05, 65N12, 65N15, 65N30.

Received August 30, 2022. Accepted March 13, 2024.

1. Introduction

Discontinuous Galerkin methods are a popular approach to discretize PDE boundary value problems [19]. Similar to conforming finite element methods [15, 24], they can handle general meshes, which allows to account for complicated geometries. In addition, the use of discontinuous polynomial shape functions facilitates the presence of hanging nodes and varying polynomial degree. As a result, discontinuous Galerkin methods are especially suited for *hp*-adaptivity [17, 33]. Furthermore, since all their degrees of freedom are attached with mesh cells, discontinuous Galerkin methods allow linear memory access, which is crucial for efficient computer implementations, in particular on GPUs [8].

In the context of wave propagation problems, discontinuous Galerkin methods further display specific advantages. For time-dependent problems, the resulting mass-matrix is block-diagonal [2, 31], which enables explicit time-stepping schemes without mass-lumping [16]. For time-harmonic wave propagation too, discontinuous Galerkin formulations are attractive as they exhibit additional stability as compared to conforming alternatives [6, 26, 27]. These interesting properties have henceforth motivated a large number of works considering discontinuous Galerkin discretizations of wave propagation problems, and a non-exhaustive list includes [2, 17, 21, 26, 27, 31, 32, 40].

Keywords and phrases. *a priori* error estimates, *a posteriori* error estimates, Aubin–Nitsche trick, discontinuous Galerkin, Helmholtz problems, interior penalty, minimal regularity.

Inria Univ. Lille and Laboratoire Paul Painlevé, 59655 Villeneuve-d'Ascq, France.
*Corresponding author: theophile.chaumont@inria.fr

In this work, we consider the acoustic Helmholtz equation, which is probably the simplest model problem relevant to the difficulties of wave propagation. Specifically, given a domain $\Omega \subset \mathbb{R}^d$, $d = 2$ or 3, and $f : \Omega \to \mathbb{C}$, the unknown $u : \Omega \to \mathbb{C}$ should satisfy

$$\begin{cases} -\omega^2 \mu u - \boldsymbol{\nabla} \cdot (\boldsymbol{A} \boldsymbol{\nabla} u) = f & \text{in } \Omega, \\ u = 0 & \text{on } \Gamma_{\mathrm{D}}, \\ \boldsymbol{A} \boldsymbol{\nabla} u \cdot \boldsymbol{n} = 0 & \text{on } \Gamma_{\mathrm{N}}, \\ \boldsymbol{A} \boldsymbol{\nabla} u \cdot \boldsymbol{n} - i\omega\gamma u = 0 & \text{on } \Gamma_{\mathrm{R}}, \end{cases} \tag{1.1}$$

where $\overline{\Gamma_{\mathrm{D}}} \cup \overline{\Gamma_{\mathrm{N}}} \cup \overline{\Gamma_{\mathrm{R}}} = \partial\Omega$ is a partition of the boundary and $\mu, \boldsymbol{A}$ and γ are given coefficients. Precise assumptions are listed in Section 2.1.

Our interest lies in interior penalty discontinuous Galerkin (IPDG) discretizations of (1.1). In particular, we focus on the "minimal regularity" case, where we do not assume any specific smoothness for the solution u. To the best of our knowledge, this problem has not been considered in the literature, and available works essentially assumes that the solution belongs to $H^{3/2+\varepsilon}$, so that the traces of $\boldsymbol{\nabla} u$ are well-defined on mesh faces, see *e.g.* equation (4.5) from [32] and Lemma 2.6 from [40]. Unfortunately, such assumptions rule out important configurations of coefficients and boundary conditions, which may bring the regularity of the solution arbitrarily close to H^1 (see the appendix of [18] for instance).

When considering conforming finite element discretizations of (1.1), the so-called "Schatz argument" enables to show that the discrete solution u_h is quasi-optimal if the mesh is fine enough [41]. Assuming for simplicity that $\Gamma_{\mathrm{R}} = \varnothing$ and introducing the energy norm

$$\|v\|_{\omega,\Omega}^2 := \omega^2 \|v\|_{\mu,\Omega}^2 + \|\boldsymbol{\nabla} v\|_{\boldsymbol{A},\Omega}^2 \quad v \in H^1_{\Gamma_{\mathrm{D}}}(\Omega),$$

we have

$$\|u - u_h\|_{\omega,\Omega} \leqslant \frac{1}{1 - \gamma_{\mathrm{ba}}^2} \min_{v_h \in V_h} \|u - v_h\|_{\omega,\Omega}, \tag{1.2}$$

whenever the approximation factor

$$\gamma_{\mathrm{ba}} := 2\omega \max_{\substack{\psi \in L^2(\Omega) \\ \|\psi\|_{\mu,\Omega} = 1}} \min_{v_h \in V_h} \|u_\psi - v\|_{\omega,\Omega} \tag{1.3}$$

is strictly less than one (all these notations are detailed in Section 2 below). Above, u_ψ solves (1.1) with right-hand side $\mu\psi$ instead of f. Similarly, when considering *a posteriori* error estimation [14], we have

$$\|u - u_h\|_{\omega,\Omega} \leqslant \sqrt{1 + \gamma_{\mathrm{ba}}^2}\,\eta, \tag{1.4}$$

where η is a suitable *a posteriori* estimator.

Estimates similar to (1.2) and (1.4) are available for IPDG discretizations [37, 40], but with energy norms involving the normal trace of the gradient on faces, thus essentially requiring $H^{3/2+\varepsilon}$ regularity of u_ψ. Here, in contrast, we extend (1.2) and (1.4) to IPDG discretizations without additional regularity assumptions on the solutions u_ψ. As detailed below, our key finding is that it can be achieved by redefining the approximation factor as

$$\gamma_{\mathrm{ba}}^2 := 4\omega^2 \max_{\substack{\psi \in L^2(\Omega) \\ \|\psi\|_{\mu,\Omega} = 1}} \left(\min_{v_h \in V_h} \|u_\psi - v_h\|_{\dagger,1,\mathcal{T}_h}^2 + \min_{\boldsymbol{w}_h \in \boldsymbol{W}_h} \|\boldsymbol{A}\boldsymbol{\nabla} u_\psi - \boldsymbol{w}_h\|_{\dagger,\mathrm{div},\mathcal{T}_h}^2 \right), \tag{1.5}$$

where $\boldsymbol{W}_h$ is the BDM finite element space built using the same mesh and polynomial degree than V_h, and $\|\cdot\|_{\dagger,1,\mathcal{T}_h}$ and $\|\cdot\|_{\dagger,\mathrm{div},\mathcal{T}_h}$ are $H^1_{\Gamma_{\mathrm{D}}}(\Omega)$ and $\boldsymbol{H}_{\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$ norms appropriately scaled by the mesh size. The additional term in (1.5) as compared to (1.3) is necessary to account for the non-conformity of the scheme.

Actually, the interest of the subject exceeds time-harmonic wave propagation. In fact, the *a priori* and *a posteriori* error analysis of finite element discretizations to (1.1) rely on duality arguments of Aubin-Nitsche type [41]. Such techniques are crucial in time-harmonic wave propagation [36,40], but there also useful in other contexts to establish convergence in weak norms, see *e.g.* Section 5.1 from [4]. To the best of our knowledge, duality analysis for IPDG discretizations under minimal regularity has not been addressed in the literature, and it is our goal to do so here.

The remainder of this work is organized as follows. In Section 2, we precise the setting and introduce key notations. Section 3 presents the key argument that enables duality techniques for IPDG under minimal regularity. Sections 4 and 5 then employ the aforementioned reasoning to perform the *a priori* and *a posteriori* error analysis of IPDG discretization for Helmholtz problems. Finally, for the sake of completeness, we provide an estimate for the approximation factor $\gamma_{\rm ba}$ in Appendix A.

2. Setting and preliminary results

2.1. Setting

Throughout this work, $\Omega \subset \mathbb{R}^d$, with $d = 2$ or 3, is a Lipschitz polytopal domain. The boundary of Ω is partitioned into three open, Lipschitz and disjoint polytopal subsets $\Gamma_{\rm D}$, $\Gamma_{\rm N}$ and $\Gamma_{\rm R}$ such that $\partial\Omega = \overline{\Gamma_{\rm D}} \cup \overline{\Gamma_{\rm N}} \cup \overline{\Gamma_{\rm R}}$. We employ the notation $\boldsymbol{n}$ for the unit vector normal to $\partial\Omega$ pointing outside Ω.

We consider coefficients $\mu : \Omega \to \mathbb{R}$, $\boldsymbol{A} : \Omega \to \mathbb{R}^{d\times d}$ and $\gamma : \Gamma_{\rm R} \to \mathbb{R}$ satisfying the following properties. We assume that there exists a partition $\mathscr{P}$ of Ω into a finite number of disjoint open polytopal subdomains such that $\mu|_P = \mu_P \in \mathbb{R}$ and $\boldsymbol{A}|_P = \boldsymbol{A}_P \in \mathbb{R}^{d\times d}$ take constant values for all $P \in \mathscr{P}$. Similarly, there exists a finite partition $\mathscr{Q}$ of $\Gamma_{\rm R}$ consisting of open polytopal subsets such that $\gamma|_Q = \gamma_Q \in \mathbb{R}$ is constant for each $Q \in \mathscr{Q}$. For each $P \in \mathscr{P}$, we assume that $\boldsymbol{A}_P$ is symmetric, and let $\alpha_P := \min_{\boldsymbol{\xi}\in\mathbb{R}^d; |\boldsymbol{\xi}|=1} \boldsymbol{A}_P\boldsymbol{\xi}\cdot\boldsymbol{\xi}$. We then classically require that $\min_{P\in\mathscr{P}} \mu_P > 0$, $\min_{P\in\mathscr{P}} \alpha_P > 0$, and $\min_{Q\in\mathscr{Q}} \gamma_Q > 0$.

We also fix a real number $\omega > 0$ representing the (angular) frequency.

2.2. Functional spaces

If $D \subset \mathbb{R}^d$ is an open set, we denote by $L^2(D)$ the Lebesgue space of complex-valued square integrable functions defined over D, and we set $\boldsymbol{L}^2(D) := [L^2(D)]^d$ for vector-valued functions. The notations $(\cdot,\cdot)_D$ and $\|\cdot\|_D$ then stand for the usual inner-product and norm of $L^2(D)$ or $\boldsymbol{L}^2(D)$. In addition, if $w : D \to \mathbb{R}$ is a measurable function satisfying $0 < \operatorname{ess\,inf}_D w$ and $\operatorname{ess\,sup}_D w < +\infty$, then $\|\cdot\|^2_{w,D} := (w\cdot,\cdot)_D$ defines a norm on $L^2(D)$ equivalent to the standard one. We use the same notation in $\boldsymbol{L}^2(D)$ with matrix-valued weights. Besides, we employ similar notations for $d-1$ manifolds.

$H^1(D)$ is the Sobolev space of functions $v \in L^2(D)$ such that $\boldsymbol{\nabla} v \in \boldsymbol{L}^2(D)$, where $\boldsymbol{\nabla}$ denotes the weak gradient defined in the sense of distributions. If $\Gamma \subset \partial D$ is a relatively open subset of the boundary of D, then $H^1_\Gamma(D)$ stands for the subset of functions $v \in H^1(D)$ such that $v|_\Gamma = 0$ in the sense of traces. We refer the reader to [1] for a detailed description of the above spaces. On $H^1_{\Gamma_{\rm D}}(\Omega)$, we will often employ the following "energy" norm:

$$\|v\|^2_{\omega,\Omega} := \omega^2\|v\|^2_{\mu,\Omega} + \omega\|v\|^2_{\gamma,\Gamma_{\rm R}} + \|\boldsymbol{\nabla} v\|^2_{\boldsymbol{A},\Omega} \quad \forall v \in H^1_{\Gamma_{\rm D}}(\Omega). \tag{2.1}$$

It turns out that Sobolev spaces of vector-valued functions will be key in the duality analysis we are about to perform. Specifically, we will need the space $\boldsymbol{H}(\mathrm{div}, D)$ of functions $\boldsymbol{v} \in \boldsymbol{L}^2(D)$ such that $\boldsymbol{\nabla}\cdot\boldsymbol{v} \in L^2(D)$ where $\boldsymbol{\nabla}\cdot$ is the weak divergence operator [29]. Following, *e.g.*, [28], if $\Gamma \subset \partial D$ is a relatively open set, the normal trace $(\boldsymbol{w}\cdot\boldsymbol{n})|_\Gamma$ of $\boldsymbol{w} \in \boldsymbol{H}(\mathrm{div}, D)$ can be defined in a weak sense, and $\boldsymbol{H}_\Gamma(\mathrm{div}, D)$ will stand for the space of $\boldsymbol{w} \in \boldsymbol{H}(\mathrm{div}, D)$ such that $(\boldsymbol{w}\cdot\boldsymbol{n})|_\Gamma = 0$.

We are studying "Robin-type" boundary conditions which are not naturally handled in the $\boldsymbol{H}(\mathrm{div})$ setting. We follow the standard remedy [10,39] and introduce the space

$$\boldsymbol{X}(\mathrm{div}, \Omega) := \left\{\boldsymbol{w} \in \boldsymbol{H}(\mathrm{div}, \Omega) \mid (\boldsymbol{w}\cdot\boldsymbol{n})|_{\Gamma_{\rm R}} \in L^2(\Gamma_{\rm R})\right\},$$

where additional normal trace regularity is enforced on the Robin boundary. The notation $\boldsymbol{X}_{\Gamma_{\rm N}}(\mathrm{div}, \Omega) := \boldsymbol{X}(\mathrm{div}, \Omega) \cap \boldsymbol{H}_{\Gamma_{\rm N}}(\mathrm{div}, \Omega)$ will also be useful.

2.3. Helmholtz problem

Central to our considerations will be the sesquilinear form

$$b(\phi, v) := -\omega^2(\mu\phi, v)_\Omega - i\omega(\gamma\phi, v)_{\Gamma_{\rm R}} + (\boldsymbol{A}\boldsymbol{\nabla}\phi, \boldsymbol{\nabla} v)_\Omega \qquad \forall \phi, v \in H^1_{\Gamma_{\rm D}}(\Omega), \tag{2.2}$$

corresponding to the weak formulation of (1.1). We will assume throughout this work that the considered Helmholtz problem is well-posed, meaning that there exists $\gamma_{\rm st} > 0$ such that

$$\min_{\substack{\phi \in H^1_{\Gamma_{\rm D}}(\Omega) \\ \|\phi\|_{\omega,\Omega}=1}} \max_{\substack{v \in H^1_{\Gamma_{\rm D}}(\Omega) \\ \|v\|_{\omega,\Omega}=1}} \operatorname{Re} b(\phi, v) = \frac{1}{\gamma_{\rm st}}\cdot \tag{2.3}$$

In our setting, (2.3) always holds as soon as $\Gamma_{\rm R}$ has a positive measure due to the unique continuation principle. On the other hand, if $\Gamma_{\rm R} = \emptyset$, then (2.3) fails to hold if and only if ω^2 is an eigenvalue of the resulting self-adjoint operator. In general, it is hard to quantitatively estimate $\gamma_{\rm st}$, but a reasonable assumption is that it grows polynomially with the frequency [35]. It is also worth mentioning that the constant may be explicitly controlled in some specific configurations [5, 9, 12, 38].

In the remainder of this work, we fix a right-hand side $f \in L^2(\Omega)$, and let $u \in H^1_{\Gamma_{\rm D}}(\Omega)$ be the unique element satisfying

$$b(u, v) = (f, v)_\Omega \quad \forall v \in H^1_{\Gamma_{\rm D}}(\Omega). \tag{2.4}$$

2.4. Computational mesh

We consider a mesh $\mathcal{T}_h$ of Ω consisting of non-overlapping (closed) simplicial elements K. We classically assume that $\mathcal{T}_h$ is a matching mesh meaning that the intersection of two distinct elements either is empty, or it is a full face, edge or vertex of the two elements (see, *e.g.* [15], Sect. 2.2 or [24], Def. 6.11). The set of faces of the mesh is denoted by $\mathcal{F}_h$. We also employ the standard notations h_K and ρ_K for the diameter of $K \in \mathcal{T}_h$ and the diameter of the largest ball contained in K (see [15], Thm. 3.1.3 or [24], Def. 6.4). Then, $\kappa_K := h_K/\rho_K$ is the shape-regularity parameter of $K \in \mathcal{T}_h$, and $\kappa := \max_{K \in \mathcal{T}_h} \kappa_K$. We also introduce the global mesh size $h := \max_{K \in \mathcal{T}_h} h_K$. Similarly, h_F stands for the diameter of the face $F \in \mathcal{F}_h$.

We further assume that the mesh $\mathcal{T}_h$ conforms with the partition of the boundary and coefficients in the sense that, for each $K \in \mathcal{T}_h$, there exists a (unique) $P \in \mathscr{P}$ such that $K \subset \overline{P}$ and for each $F \in \mathcal{F}_h$ with $F \subset \partial\Omega$, we have either $F \subset \overline{\Gamma_{\rm D}}$, $F \subset \overline{\Gamma_{\rm N}}$ or $F \subset \overline{\Gamma_{\rm R}}$. If $F \subset \overline{\Gamma_{\rm R}}$, we additionally require that $F \subset \overline{Q}$ for some (unique) $Q \in \mathscr{Q}$. We respectively denote by $\mathcal{F}_h^{\rm D}$, $\mathcal{F}_h^{\rm N}$ and $\mathcal{F}_h^{\rm R}$ the set of faces $F \in \mathcal{F}_h$ such that $F \subset \overline{\Gamma_{\rm D}}$, $\overline{\Gamma_{\rm N}}$ or $\overline{\Gamma_{\rm R}}$. We also set $\mathcal{F}_h^{\rm e} := \mathcal{F}_h^{\rm D} \cup \mathcal{F}_h^{\rm N} \cup \mathcal{F}_h^{\rm R}$, and $\mathcal{F}^{\rm i} := \mathcal{F}_h \backslash \mathcal{F}_h^{\rm e}$. For $K \in \mathcal{T}_h$, we can then set $\mu_K := \mu_P$, $\alpha_K := \alpha_P$, and $\vartheta_K := \sqrt{\alpha_K/\mu_K}$, where $P \in \mathscr{P}$ contains K. Similarly, if $F \in \mathcal{F}_h^{\rm e}$, we set $\alpha_F := \alpha_K$ where $K \in \mathcal{T}_h$ is the only element having F as a face. If $F \in \mathcal{F}_h^{\rm R}$, we also introduce $\gamma_F := \gamma_Q$ and $\vartheta_F := \alpha_F/\gamma_F$ where Q is the unique subset in $\mathscr{Q}$ containing F.

Finally, we associate with each $F \in \mathcal{F}_h$ a unit normal vector $\boldsymbol{n}_F$. If $F \in \mathcal{F}_h^{\rm e}$, we require that $\boldsymbol{n}_F = \boldsymbol{n}$. Otherwise, the orientation of $\boldsymbol{n}_F$ is arbitrary, but fixed.

Remark 2.1 (Hanging nodes)**.** For the sake of simplicity, our assumptions on the mesh rule out the presence of hanging nodes. However, this is not essential for the analysis performed in the manuscript. Instead, the key assumption required is that the conforming Lagrange and Raviart-Thomas finite element spaces associated with the mesh are sufficiently rich, which is known to be the case in a variety of situations, see *e.g.* [3, 7]. Since the precise assumptions are intricate, we have chosen to restrict our attention to matching meshes to ease the presentation. The interested reader will find more details in Appendix A below. Related to this comment, notice that the present analysis would be harder to extend to polytopal meshes.

2.5. Polynomial spaces

In the remainder of this work, we fix a polynomial degree $p \geqslant 1$. For $K \in \mathcal{T}_h$, $\mathcal{P}_p(K)$ is the set of polynomial functions $K \to \mathbb{C}$ of total degree less than or equal to q. For vector-valued functions, we also set $\boldsymbol{\mathcal{P}}_p(K) := [\mathcal{P}_p(K)]^3$. If $\mathcal{T} \subset \mathcal{T}_h$ is a collection of elements covering the (open) domain Θ, then $\mathcal{P}_p(K) := \{v \in L^2(\Theta);\ v|_K \in \mathcal{P}_p(K)\ \forall K \in \mathcal{T}\}$ and $\boldsymbol{\mathcal{P}}_p(\mathcal{T}) := [\mathcal{P}_p(\mathcal{T})]^3$.

2.6. Broken Sobolev spaces

We define $H^1(\mathcal{T}_h)$ as the subset of functions $v \in L^2(\Omega)$ such that $v|_K \in H^1(K)$ for all $K \in \mathcal{T}_h$. For functions in $H^1(\mathcal{T}_h)$, we still employ the notations $\boldsymbol{\nabla}$ for the element-wise weak gradient, so that $\boldsymbol{\nabla}(H^1(\mathcal{T}_h)) \subset \boldsymbol{L}^2(\Omega)$. To avoid confusions, we will employ the alternative notations $(\cdot,\cdot)_{\mathcal{T}_h}$ and $\|\cdot\|_{\mathcal{T}_h}$ for the $L^2(\Omega)$ and $\boldsymbol{L}^2(\Omega)$ norms and inner-products when working with broken functions (we also employ the same notation for weighted norms). Similarly, we write

$$\|v\|^2_{\gamma,\mathcal{F}_h^{\mathrm{R}}} := \sum_{F\in\mathcal{F}_h^{\mathrm{R}}} \|v\|^2_{\gamma,F} \quad \forall v \in L^2(\Gamma_{\mathrm{R}}).$$

The notation

$$(\phi, v)_{\mathcal{F}_h} := \sum_{F\in\mathcal{F}_h} (\phi, v)_F \qquad \forall \phi, v \in \bigoplus_{F\in\mathcal{F}_h} L^2(F)$$

will also be useful.

2.7. Jumps, averages, lifting and discrete gradient

Considering $\phi \in H^1(\mathcal{T}_h)$, its jump through a face $F \in \mathcal{F}_h^{\mathrm{i}}$ is defined by

$$[\![\phi]\!]_F := \phi_+|_F \boldsymbol{n}_+ \cdot \boldsymbol{n}_F + \phi_-|_F \boldsymbol{n}_- \cdot \boldsymbol{n}_F$$

with $K_\pm \in \mathcal{T}_h$ the two elements such that $F = \partial K_+ \cap \partial K_-$, $\phi|_\pm := \phi|_{K_\pm}$ and $\boldsymbol{n}_\pm := \boldsymbol{n}_{K_\pm}$. For an exterior face $F \in \mathcal{F}_h^{\mathrm{e}}$, we set instead

$$[\![\phi]\!]_F := \phi|_F \text{ if } F \in \mathcal{F}_h^{\mathrm{D}} \quad \text{ and } \quad [\![\phi]\!]_F := 0 \text{ otherwise.}$$

Similarly, if $\boldsymbol{w} \in \boldsymbol{\mathcal{P}}_p(\mathcal{T}_h)$, its average on $F \in \mathcal{F}_h$ is given by

$$\{\!\{\boldsymbol{w}\}\!\}_F := \frac{1}{2}(\boldsymbol{w}_+|_F + \boldsymbol{w}_-|_F) \quad \text{ if } F \in \mathcal{F}_h^{\mathrm{i}} \quad \text{ and } \quad \{\!\{\boldsymbol{w}\}\!\}_F := \boldsymbol{w}|_F \quad \text{ if } F \in \mathcal{F}_h^{\mathrm{e}},$$

where $\boldsymbol{w}_\pm := \boldsymbol{w}|_{K_\pm}$ for the two elements $K_\pm \in \mathcal{T}_h$ such that $F = K_- \cap K_+$ in the case of an interior face.

A key part of our analysis will be to give a meaning to the terms $([\![\phi]\!], \{\!\{\boldsymbol{A}\boldsymbol{\nabla} v\}\!} \cdot \boldsymbol{n}_F)_{\mathcal{F}_h}$ appearing in the IPDG form, for functions ϕ and v only belonging to $H^1(\mathcal{T}_h)$. This is subtle, since the normal trace of $\boldsymbol{\nabla} v$ is actually not defined on faces. Following Section 4.3 from [19], the solution is to introduce a lifting operator defined as follows.

For $\phi \in H^1(\mathcal{T}_h)$ we define the lifting $\boldsymbol{\ell}_h([\![\phi]\!])$ as the unique element of $\boldsymbol{\mathcal{P}}_p(\mathcal{T}_h)$ such that

$$(\boldsymbol{\ell}_h([\![\phi]\!]), \boldsymbol{w}_h)_{\mathcal{T}_h} = ([\![\phi]\!]_F, \{\!\{\boldsymbol{w}_h\}\!\} \cdot \boldsymbol{n}_F)_{\mathcal{F}_h} \qquad \forall \boldsymbol{w}_h \in \boldsymbol{\mathcal{P}}_p(\mathcal{T}_h).$$

Notice that we can then write $(\boldsymbol{A}\boldsymbol{\ell}_h([\![\phi]\!]), \boldsymbol{\nabla} v)_{\mathcal{T}_h}$ instead of $([\![\phi]\!], \{\!\{\boldsymbol{A}\boldsymbol{\nabla} v\}\!\} \cdot \boldsymbol{n}_F)_{\mathcal{F}_h}$ for functions $\phi, v \in \mathcal{P}_p(\mathcal{T}_h)$, with the advantage that the second expression is still well-defined for general $H^1(\mathcal{T}_h)$ arguments.

The notion of weak gradient will also be useful. Specifically, we set

$$\mathfrak{G}(\phi) := \boldsymbol{\nabla}\phi - \boldsymbol{\ell}_h([\![\phi]\!]) \in \boldsymbol{L}^2(\mathcal{T}_h)$$

for all $\phi \in H^1(\mathcal{T}_h)$. Importantly, we have $\mathfrak{G}(\phi) = \boldsymbol{\nabla}\phi$ whenever $\phi \in H^1_{\Gamma_{\mathrm{D}}}(\Omega)$, and

$$(\mathfrak{G}(\phi), \boldsymbol{w}_h)_{\mathcal{T}_h} + (\phi, \boldsymbol{\nabla} \cdot \boldsymbol{w}_h)_{\mathcal{T}_h} = (\phi, \boldsymbol{w}_h \cdot \boldsymbol{n})_{\Gamma_{\mathrm{R}}}$$

for all $\phi \in H^1(\mathcal{T}_h)$ and $\boldsymbol{w}_h \in \boldsymbol{\mathcal{P}}_p(\mathcal{T}_h) \cap \boldsymbol{X}_{\Gamma_{\mathrm{N}}}(\mathrm{div}, \Omega)$.

2.8. Broken norms

For $v \in H^1(\mathcal{T}_h)$, the broken counterpart of the energy norm introduced in (2.1) is defined by

$$\|v\|_{\omega,\mathcal{T}_h}^2 := \omega^2\|v\|_{\mu,\mathcal{T}_h}^2 + \omega\|v\|_{\gamma,\mathcal{F}_h^{\mathrm{R}}}^2 + \|\mathfrak{G}(v)\|_{\boldsymbol{A},\mathcal{T}_h}^2.$$

Notice that it is equal to the $\|\cdot\|_{\omega,\Omega}$ norm if $v \in H^1_{\Gamma_{\mathrm{D}}}(\Omega)$. We will also employ the mesh-dependent norms

$$\|v\|_{\dagger,1,\mathcal{T}_h}^2 := \sum_{K\in\mathcal{T}_h}\left\{\max\left(1,\frac{\omega^2h_K^2}{\vartheta_K^2}\right)\frac{\alpha_K}{h_K^2}\|v\|_K^2 + \|\mathfrak{G}(v)\|_{\boldsymbol{A},K}^2\right\} + \sum_{F\in\mathcal{F}_h^{\mathrm{R}}}\max\left(1,\frac{\omega h_F}{\vartheta_F}\right)\frac{\alpha_F}{h_F}\|v\|_F^2$$

and

$$\|\boldsymbol{w}\|_{\dagger,\mathrm{div},\mathcal{T}_h}^2 := \sum_{K\in\mathcal{T}_h}\left\{\|\boldsymbol{w}\|_{\boldsymbol{A}^{-1},K}^2 + \frac{h_K^2}{\alpha_K}\|\boldsymbol{\nabla}\cdot\boldsymbol{w}\|_K^2\right\} + \sum_{F\in\mathcal{F}_h^{\mathrm{R}}}\frac{h_F}{\alpha_F}\|\boldsymbol{w}\cdot\boldsymbol{n}\|_F^2$$

for $\boldsymbol{w} \in \boldsymbol{X}_{\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$. These norms are "dual" to each other in the sense that

$$|(\mathfrak{G}(\phi),\boldsymbol{w})_{\mathcal{T}_h} + (\phi,\boldsymbol{\nabla}\cdot\boldsymbol{w})_{\mathcal{T}_h} - (\phi,\boldsymbol{w}\cdot\boldsymbol{n})_{\mathcal{F}_h^{\mathrm{R}}}| \leqslant \|\phi\|_{\dagger,1,\mathcal{T}_h}\|\boldsymbol{w}\|_{\dagger,\mathrm{div},\mathcal{T}_h} \tag{2.5}$$

for all $\phi \in H^1(\mathcal{T}_h)$ and $\boldsymbol{w} \in \boldsymbol{X}_{\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$. For future references, we notice that the "max" in the definition of the $\|\cdot\|_{\dagger,1,\mathcal{T}_h}$-norm chosen so that

$$\|v\|_{\omega,\mathcal{T}_h} \leqslant \|v\|_{\dagger,1,\mathcal{T}_h} \qquad \forall v \in H^1(\mathcal{T}_h). \tag{2.6}$$

2.9. IPDG form

Following Section 4.3.3 from [19], our DG approximation of the $(\boldsymbol{A}\boldsymbol{\nabla}\cdot,\boldsymbol{\nabla}\cdot)_\Omega$ form is given by

$$a_h(\phi,v) := (\boldsymbol{A}\mathfrak{G}(\phi),\mathfrak{G}(v))_{\mathcal{T}_h} + s_h(\phi,v) \qquad \forall \phi,v \in H^1(\mathcal{T}_h) \tag{2.7}$$

where $s_h : H^1(\mathcal{T}_h)\times H^1(\mathcal{T}_h) \to \mathbb{C}$ is a sesquilinear form satisfying $s_h(\phi,v) = 0$ whenever $\phi \in H^1_{\Gamma_{\mathrm{D}}}(\Omega)$ or $v \in H^1_{\Gamma_{\mathrm{D}}}(\Omega)$. Importantly, we have

$$a_h(\phi,v) = (\boldsymbol{A}\boldsymbol{\nabla}\phi,\boldsymbol{\nabla}v)_\Omega \tag{2.8}$$

when $\phi,v \in H^1_{\Gamma_{\mathrm{D}}}(\Omega)$.

It is worthy to note that the IPDG form is usually not presented (and not implemented) as it is presented in (2.7). Instead [2, 26, 31, 40], the method is usually written as

$$a_h(\phi,v) = (\boldsymbol{A}\boldsymbol{\nabla}\phi,\boldsymbol{\nabla}v)_{\mathcal{T}_h} - (\{\!\!\{\boldsymbol{A}\boldsymbol{\nabla}\phi\}\!\!\}\cdot\boldsymbol{n}_F,[\![v]\!])_{\mathcal{F}_h} - ([\![\phi]\!],\{\!\!\{\boldsymbol{A}\boldsymbol{\nabla}v\}\!\!\}\cdot\boldsymbol{n}_F)_{\mathcal{F}_h} + \sum_{F\in\mathcal{F}_h^{\mathrm{i}}\cup\mathcal{F}_h^{\mathrm{D}}}\frac{\beta_F}{h_F}([\![\phi]\!],[\![v]\!])_F \tag{2.9}$$

where $\beta_F \sim p^2$ is a penalty parameter chosen to be large enough [2]. However, as shown in Section 4.3.3 from [19], the formulation of (2.9) can be recast in the framework of (2.7) by setting

$$s_h(\phi,v) := \sum_{F\in\mathcal{F}_h^{\mathrm{i}}\cup\mathcal{F}_h^{\mathrm{D}}}\frac{\beta_F}{h_F}([\![\phi]\!],[\![v]\!])_F - (\boldsymbol{A}\boldsymbol{\ell}_h([\![\phi]\!]),\boldsymbol{\ell}_h([\![v]\!]))_\Omega. \tag{2.10}$$

2.10. The discrete Helmholtz problem

Consider the sesquilinear form

$$b_h(\phi,v) := -\omega^2(\mu\phi,v)_{\mathcal{T}_h} - i\omega(\gamma\phi,v)_{\mathcal{F}_h^{\mathrm{R}}} + a_h(\phi,v) \qquad \forall \phi,v \in H^1(\mathcal{T}_h).$$

Then, the discrete problem consists in finding $u_h \in \mathcal{P}_p(\mathcal{T}_h)$ such that

$$b_h(u_h, v_h) = (f, v_h)_{\mathcal{T}_h} \qquad \forall v_h \in \mathcal{P}_p(\mathcal{T}_h). \tag{2.11}$$

For any u_h satisfying (2.11), recalling (2.4) and due to (2.8), the Galerkin orthogonality property

$$b_h(u - u_h, w_h) = 0 \tag{2.12}$$

holds true for all discrete conforming test functions $w_h \in \mathcal{P}_p(\mathcal{T}_h) \cap H^1_{\Gamma_{\mathrm{D}}}(\Omega)$. Similarly, we have

$$|b_h(\phi, v)| \leqslant \|\phi\|_{\omega,\mathcal{T}_h} \|v\|_{\omega,\mathcal{T}_h} \tag{2.13}$$

for $\phi, v \in H^1(\mathcal{T}_h)$ whenever ϕ or v belongs to $H^1_{\Gamma_{\mathrm{D}}}(\Omega)$.

2.11. Conforming subspaces and projections

The sets $\mathcal{P}_p(\mathcal{T}_h) \cap H^1_{\Gamma_{\mathrm{D}}}(\Omega)$ and $\boldsymbol{\mathcal{P}}_p(\mathcal{T}_h) \cap \boldsymbol{X}_{\Gamma_{\mathrm{N}}}(\mathrm{div}, \Omega)$ are the usual Lagrange and BDM finite element spaces (see, *e.g.*, [24], Sects. 8.5.1 and 11.5). Although they are not needed to implement the IPDG discretization, they will be extremely useful for the analysis. The projections

$$\pi_h^{\mathrm{g}} v := \arg \min_{v_h \in \mathcal{P}_p(\mathcal{T}_h) \cap H^1_{\Gamma_{\mathrm{D}}}(\Omega)} \|v - v_h\|_{\dagger,1,\mathcal{T}_h}$$

and

$$\pi_h^{\mathrm{d}} \boldsymbol{w} := \arg \min_{\boldsymbol{w}_h \in \boldsymbol{\mathcal{P}}_p(\mathcal{T}_h) \cap \boldsymbol{X}_{\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)} \|\boldsymbol{w} - \boldsymbol{w}_h\|_{\dagger,\mathrm{div},\mathcal{T}_h}$$

are well-defined for all $v \in H^1_{\Gamma_{\mathrm{D}}}(\Omega)$ and $\boldsymbol{w} \in \boldsymbol{X}_{\Gamma_{\mathrm{N}}}(\mathrm{div}, \Omega)$, since the mesh-dependent norms appearing in the right-hand sides are naturally associated with inner-product. We will also need the conforming projector

$$\pi^{\mathrm{g}} v := \arg \min_{s \in H^1_{\Gamma_{\mathrm{D}}}(\Omega)} \|v - s\|_{\dagger,1,\mathcal{T}_h}$$

defined for all $v \in H^1(\mathcal{T}_h)$.

2.12. Approximation factors

Following [14, 36, 37, 40], our analysis will rely on duality arguments and approximation factors will be a central concept. For $\psi \in L^2(\Omega)$ and $\Psi \in L^2(\Gamma_{\mathrm{R}})$, let $u^\star_\psi, U^\star_\Psi \in H^1_{\Gamma_{\mathrm{D}}}(\Omega)$ solve

$$b(w, u^\star_\psi) = \omega(\mu w, \psi)_\Omega, \qquad b(w, U^\star_\Psi) = \omega^{1/2}(\gamma w, \Psi)_{\Gamma_{\mathrm{R}}} \qquad \forall w \in H^1_{\Gamma_{\mathrm{D}}}(\Omega).$$

The approximation factors

$$\breve{\gamma}_{\mathrm{ba,g}} := \max_{\substack{\psi \in L^2(\Omega) \\ \|\psi\|_{\mu,\Omega}=1}} \|u^\star_\psi - \pi_h^{\mathrm{g}} u^\star_\psi\|_{\dagger,1,\mathcal{T}_h}, \quad \widetilde{\gamma}_{\mathrm{ba,g}} := \max_{\substack{\Psi \in L^2(\Gamma_{\mathrm{R}}) \\ \|\Psi\|_{\gamma,\Gamma_{\mathrm{R}}}=1}} \|U^\star_\Psi - \pi_h^{\mathrm{g}} U^\star_\Psi\|_{\dagger,1,\mathcal{T}_h}, \tag{2.14a}$$

where previously employed in [14] for the *a posteriori* error analysis of conforming discretizations. Here, we will additionally need divergence-conforming approximation factors

$$\begin{aligned} \breve{\gamma}_{\mathrm{ba,d}} &:= \max_{\substack{\psi \in L^2(\Omega) \\ \|\psi\|_{\mu,\Omega}=1}} \|\boldsymbol{A}\boldsymbol{\nabla} u^\star_\psi - \pi_h^{\mathrm{d}}(\boldsymbol{A}\boldsymbol{\nabla} u^\star_\psi)\|_{\dagger,\mathrm{div},\mathcal{T}_h}, \\ \widetilde{\gamma}_{\mathrm{ba,d}} &:= \max_{\substack{\Psi \in L^2(\Gamma_{\mathrm{R}}) \\ \|\Psi\|_{\gamma,\Gamma_{\mathrm{R}}}=1}} \|\boldsymbol{A}\boldsymbol{\nabla} U^\star_\Psi - \pi_h^{\mathrm{d}}(\boldsymbol{A}\boldsymbol{\nabla} U^\star_\Psi)\|_{\dagger,\mathrm{div},\mathcal{T}_h}, \end{aligned} \tag{2.14b}$$

to deal with the non-conformity of the IPDG approximation under minimal regularity. We finally introduce the following short-hand notations

$$\gamma_{\mathrm{ba,g}}^2 := 4\check{\gamma}_{\mathrm{ba,g}}^2 + 2\widetilde{\gamma}_{\mathrm{ba,g}}^2, \qquad \gamma_{\mathrm{ba,d}}^2 := 4\check{\gamma}_{\mathrm{ba,d}}^2 + 2\widetilde{\gamma}_{\mathrm{ba,d}}^2, \tag{2.14c}$$

$$\check{\gamma}_{\mathrm{ba}}^2 := \check{\gamma}_{\mathrm{ba,g}}^2 + \check{\gamma}_{\mathrm{ba,d}}^2, \qquad \widetilde{\gamma}_{\mathrm{ba}}^2 := \widetilde{\gamma}_{\mathrm{ba,g}}^2 + \widetilde{\gamma}_{\mathrm{ba,d}}^2, \tag{2.14d}$$

and

$$\gamma_{\mathrm{ba}}^2 := \gamma_{\mathrm{ba,g}}^2 + \gamma_{\mathrm{ba,d}}^2 = 4\check{\gamma}_{\mathrm{ba}}^2 + 2\widetilde{\gamma}_{\mathrm{ba}}^2. \tag{2.14e}$$

By combining elliptic regularity and approximation properties of finite element spaces, it is easy to see that $\gamma_{\mathrm{ba}} \to 0$ as $h/p \to 0$. However, the dependency on the PDE coefficients and on the frequency may be hard to track. In particular, the stability constant γ_{st} typically plays a central role in estimating γ_{ba}. Qualitative estimates for γ_{ba} can be found in [11,36], and explicit estimates are available in specific situations [14]. We also provide an estimate in Appendix A below.

3. Duality under minimal regularity

Here, we state in a separate section the key technical result that enables us to perform duality techniques under minimal regularity assumptions. We believe it will be important in other contexts, so that we make it easy to reference. Although the arguments are not complicated, we were not able to find a proof in the literature.

Lemma 3.1 (Lifting error)**.** *For all $\phi \in H^1(\mathcal{T}_h)$ and $\boldsymbol{\sigma} \in \boldsymbol{X}_{\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$, we have*

$$(\mathfrak{G}(\phi),\boldsymbol{\sigma})_{\mathcal{T}_h} + (\phi,\boldsymbol{\nabla}\cdot\boldsymbol{\sigma})_{\mathcal{T}_h} - (\phi,\boldsymbol{\sigma}\cdot\boldsymbol{n})_{\mathcal{F}_h^{\mathrm{R}}} = (\mathfrak{G}(\phi-\widetilde{\phi}),\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)_{\mathcal{T}_h} + (\phi-\widetilde{\phi},\boldsymbol{\nabla}\cdot(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h))_{\mathcal{T}_h} - (\phi-\widetilde{\phi},(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\cdot\boldsymbol{n})_{\mathcal{F}_h^{\mathrm{R}}} \tag{3.1}$$

and in particular

$$|(\mathfrak{G}(\phi),\boldsymbol{\sigma})_{\mathcal{T}_h} + (\phi,\boldsymbol{\nabla}\cdot\boldsymbol{\sigma})_{\mathcal{T}_h} - (\phi,\boldsymbol{\sigma}\cdot\boldsymbol{n})_{\mathcal{F}_h^{\mathrm{R}}}| \leqslant \|\phi-\widetilde{\phi}\|_{\dagger,1,\mathcal{T}_h}\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{\dagger,\mathrm{div},\mathcal{T}_h} \tag{3.2}$$

for all $\widetilde{\phi} \in H^1_{\Gamma_{\mathrm{D}}}(\Omega)$ and $\boldsymbol{\sigma}_h \in \boldsymbol{\mathcal{P}}_p(\mathcal{T}_h) \cap \boldsymbol{X}_{\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$.

Proof. We first observe that if $\widetilde{\phi} \in H^1_{\Gamma_{\mathrm{D}}}(\Omega)$, integration by parts gives

$$(\mathfrak{G}(\widetilde{\phi}),\boldsymbol{\tau})_{\mathcal{T}_h} + (\widetilde{\phi},\boldsymbol{\nabla}\cdot\boldsymbol{\tau}) - (\widetilde{\phi},\boldsymbol{\tau}\cdot\boldsymbol{n})_{\Gamma_{\mathrm{R}}} = (\boldsymbol{\nabla}\widetilde{\phi},\boldsymbol{\tau})_{\mathcal{T}_h} + (\widetilde{\phi},\boldsymbol{\nabla}\cdot\boldsymbol{\tau}) - (\widetilde{\phi},\boldsymbol{\tau}\cdot\boldsymbol{n})_{\Gamma_{\mathrm{R}}} = 0$$

for all $\boldsymbol{\tau} \in \boldsymbol{X}_{\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$ due to the essential boundary conditions on Γ_{D} and Γ_{N}. Thus,

$$(\mathfrak{G}(\phi),\boldsymbol{\sigma})_{\mathcal{T}_h} + (\phi,\boldsymbol{\nabla}\cdot\boldsymbol{\sigma})_{\mathcal{T}_h} - (\phi,\boldsymbol{\sigma}\cdot\boldsymbol{n})_{\mathcal{F}_h^{\mathrm{R}}} = (\mathfrak{G}(\phi-\widetilde{\phi}),\boldsymbol{\sigma})_{\mathcal{T}_h} + (\phi-\widetilde{\phi},\boldsymbol{\nabla}\cdot\boldsymbol{\sigma})_{\mathcal{T}_h} - (\phi-\widetilde{\phi},\boldsymbol{\sigma}\cdot\boldsymbol{n})_{\mathcal{F}_h^{\mathrm{R}}}.$$

If $\boldsymbol{\sigma}_h \in \boldsymbol{\mathcal{P}}_p(\mathcal{T}_h) \cap \boldsymbol{X}_{\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$, we also have

$$\begin{aligned} 0 &= (\mathfrak{G}(\phi),\boldsymbol{\sigma}_h)_{\mathcal{T}_h} + (\phi,\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}_h)_{\mathcal{T}_h} - (\phi,\boldsymbol{\sigma}_h\cdot\boldsymbol{n})_{\mathcal{F}_h^{\mathrm{R}}} \\ &= (\mathfrak{G}(\phi)-\boldsymbol{\nabla}\widetilde{\phi},\boldsymbol{\sigma}_h)_{\mathcal{T}_h} + (\boldsymbol{\nabla}\widetilde{\phi},\boldsymbol{\sigma}_h)_{\mathcal{T}_h} + (\phi,\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}_h)_{\mathcal{T}_h} - (\phi,\boldsymbol{\sigma}_h\cdot\boldsymbol{n})_{\mathcal{F}_h^{\mathrm{R}}} \\ &= (\mathfrak{G}(\phi-\widetilde{\phi}),\boldsymbol{\sigma}_h)_{\mathcal{T}_h} + (\phi-\widetilde{\phi},\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}_h)_{\mathcal{T}_h} - (\phi-\widetilde{\phi},\boldsymbol{\sigma}_h\cdot\boldsymbol{n})_{\mathcal{F}_h^{\mathrm{R}}}, \end{aligned}$$

and (3.1) follows after summation. Then, (3.2) is a direct consequence of (2.5). □

A simple consequence of Lemma 3.1 is the following result, which is crucial to estimate the non-conformity error in the context of duality arguments.

Corollary 3.2 (Control of the non-confomity)**.** *For all $\phi \in H^1(\mathcal{T}_h)$ and $\xi \in H^1_{\Gamma_{\mathrm{D}}}(\Omega)$ with $\boldsymbol{A}\boldsymbol{\nabla}\xi \in \boldsymbol{X}_{\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$, we have*

$$\left|a_h(\phi,\xi) + (\phi,\boldsymbol{\nabla}\cdot(\boldsymbol{A}\boldsymbol{\nabla}\xi))_{\mathcal{T}_h} - (\phi,\boldsymbol{A}\boldsymbol{\nabla}\xi\cdot\boldsymbol{n})_{\mathcal{F}_h^{\mathrm{R}}}\right| \leqslant \|\phi-\widetilde{\phi}\|_{\dagger,1,\mathcal{T}_h}\|\boldsymbol{A}\boldsymbol{\nabla}\xi-\boldsymbol{\sigma}_h\|_{\dagger,\mathrm{div},\mathcal{T}_h} \tag{3.3}$$

for all $\widetilde{\phi} \in H^1_{\Gamma_{\mathrm{D}}}(\Omega)$ and $\boldsymbol{\sigma}_h \in \boldsymbol{\mathcal{P}}_p(\mathcal{T}_h) \cap \boldsymbol{X}_{\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$.

4. A PRIORI ANALYSIS

The purpose of this section is to establish the existence and uniqueness of a discrete solution $u_h \in \mathcal{P}_p(\mathcal{T}_h)$ to (2.11) and to derive *a priori* error estimates controlling $u - u_h$ in suitable norms. The proof follows the line of the Schatz argument [41], with suitable modifications to take into account the non-conforming nature of the discrete scheme.

In this section, we require that there exists $\rho > 0$ such that

$$\operatorname{Re} s_h(v_h, v_h) \geqslant \rho^2 \|v_h - \pi^{\mathrm{g}} v_h\|_{\dagger,1,\mathcal{T}_h}^2 \qquad \forall v_h \in \mathcal{P}_p(\mathcal{T}_h). \tag{4.1}$$

This assumption always holds true when $s_h(\cdot,\cdot) = \beta/h([\![\cdot]\!],[\![\cdot]\!])_{\mathcal{F}_h}$ with $\beta > 0$. It is also the case when s_h takes the form (2.10) if the penalization parameter is sufficiently large [2]. We note that this assumption rules out some choices of stabilization including complex-value stabilization parameters, as proposed *e.g.* in [26, 40]. For the sake of simplicity, we only consider stabilizations satisfying (4.1), but slight modifications of our proofs could handle more general situations.

We start with the main duality argument, which is a generalization of the Aubin-Nitsche trick.

Lemma 4.1 (Aubin-Nitsche)**.** *Assume that $u_h \in \mathcal{P}_p(\mathcal{T}_h)$ satisfies (2.11). Then, we have*

$$\omega \|u - u_h\|_{\mu,\mathcal{T}_h} \leqslant \check{\gamma}_{\mathrm{ba,g}} |\!|\!| u - u_h |\!|\!|_{\omega,\mathcal{T}_h} + \check{\gamma}_{\mathrm{ba,d}} \|u_h - \pi^{\mathrm{g}} u_h\|_{\dagger,1,\mathcal{T}_h} \tag{4.2}$$

and

$$\omega^{1/2} \|u - u_h\|_{\gamma,\mathcal{F}_h^{\mathrm{R}}} \leqslant \widetilde{\gamma}_{\mathrm{ba,g}} |\!|\!| u - u_h |\!|\!|_{\omega,\mathcal{T}_h} + \widetilde{\gamma}_{\mathrm{ba,d}} \|u_h - \pi^{\mathrm{g}} u_h\|_{\dagger,1,\mathcal{T}_h}. \tag{4.3}$$

Proof. For the sake of shortness, we set $e_h := u - u_h$. We first establish (4.2). To do so, we define ξ as the unique element of $H^1_{\Gamma_{\mathrm{D}}}(\Omega)$ such that $b(\phi,\xi) = \omega(\mu\phi, e_h)_\Omega$ for all $\phi \in H^1_{\Gamma_{\mathrm{D}}}(\Omega)$. Integrating by parts, we see that the identities

$$-\omega^2 \mu \xi - \boldsymbol{\nabla} \cdot (\boldsymbol{A} \boldsymbol{\nabla} \xi) = \omega \mu e_h \qquad \boldsymbol{A}\boldsymbol{\nabla}\xi \cdot \boldsymbol{n} + i\omega\xi = 0$$

respectively hold in $L^2(\Omega)$ and $L^2(\Gamma_{\mathrm{R}})$. As a result,

$$\omega \|e_h\|_{\mu,\mathcal{T}_h}^2 = -\omega^2 (\mu e_h, \xi)_{\mathcal{T}_h} - i\omega(\gamma e_h, \xi)_{\mathcal{F}_h^{\mathrm{R}}} - \{(e_h, \boldsymbol{\nabla}\cdot(\boldsymbol{A}\boldsymbol{\nabla}\xi))_{\mathcal{T}_h} - (e_h, \boldsymbol{A}\boldsymbol{\nabla}\xi\cdot\boldsymbol{n})_{\mathcal{F}^{\mathrm{R}}}\}$$

and we have from Corollary 3.2 that

$$\omega \|e_h\|_{\mu,\mathcal{T}_h}^2 \leqslant |b_h(e_h,\xi)| + \|e_h - \widetilde{\phi}\|_{\dagger,1,\mathcal{T}_h} \|\boldsymbol{A}\boldsymbol{\nabla}\xi - \boldsymbol{\sigma}_h\|_{\dagger,\mathrm{div},\mathcal{T}_h}$$

for all $\widetilde{\phi} \in H^1_{\Gamma_{\mathrm{D}}}(\Omega)$ and $\boldsymbol{\sigma}_h \in \boldsymbol{\mathcal{P}}_p(\mathcal{T}_h) \cap \boldsymbol{X}_{\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$. On the one hand, we select $\widetilde{\phi} = u - \pi^{\mathrm{g}} u_h$ and $\boldsymbol{\sigma}_h = \pi_h^{\mathrm{d}}(\boldsymbol{A}\boldsymbol{\nabla}\xi)$, so that using (2.14b), we have

$$\omega \|e_h\|_{\mu,\mathcal{T}_h}^2 \leqslant |b_h(e_h,\xi)| + \check{\gamma}_{\mathrm{ba,d}} \|u_h - \pi^{\mathrm{g}} u_h\|_{\dagger,1,\mathcal{T}_h} \|e_h\|_{\mu,\mathcal{T}_h}.$$

On the other hand, using (2.12) and (2.13), we have

$$|b_h(e_h,\xi)| = |b_h(e_h, \xi - \pi_h^{\mathrm{g}}\xi)| \leqslant |\!|\!| e_h |\!|\!|_{\omega,\mathcal{T}_h} |\!|\!| \xi - \pi_h^{\mathrm{g}}\xi |\!|\!|_{\omega,\Omega} \leqslant \check{\gamma}_{\mathrm{ba,g}} |\!|\!| e_h |\!|\!|_{\omega,\mathcal{T}_h} \|e_h\|_{\mu,\mathcal{T}_h}.$$

We then prove (4.3). Instead of ξ, we define χ as the unique element of $H^1_{\Gamma_{\mathrm{D}}}(\Omega)$ such that $b(\phi,\chi) = \omega^{1/2}(\gamma\phi, e_h)_{\Gamma_{\mathrm{R}}}$ for all $\phi \in H^1_{\Gamma_{\mathrm{D}}}(\Omega)$. We then have

$$-\omega^2 \mu \chi - \boldsymbol{\nabla}\cdot(\boldsymbol{A}\boldsymbol{\nabla}\chi) = 0 \qquad \boldsymbol{A}\boldsymbol{\nabla}\chi\cdot\boldsymbol{n} + i\omega\chi = \omega^{1/2}\gamma e_h$$

in $L^2(\Omega)$ and $L^2(\Gamma_{\mathrm{R}})$, respectively. We thus have

$$\omega^{1/2} \|e_h\|_{\gamma,\mathcal{F}_h^{\mathrm{R}}}^2 = -\omega^2 (\mu e_h, \chi)_{\mathcal{T}_h} - i\omega(e_h, \phi)_{\mathcal{F}_h^{\mathrm{R}}} - \{(e_h, \boldsymbol{\nabla}\cdot(\boldsymbol{A}\boldsymbol{\nabla}\chi))_{\mathcal{T}_h} - (e_h, \boldsymbol{A}\boldsymbol{\nabla}\chi\cdot\boldsymbol{n})_{\mathcal{F}^{\mathrm{R}}}\},$$

and applying again Corollary 3.2 we arrive at

$$\omega^{1/2}\|e_h\|^2_{\gamma,\mathcal{F}_h^{\mathrm{R}}} \leqslant |b_h(e_h,\chi-\pi_h^{\mathrm{g}}\chi)| + \|u_h-\pi^{\mathrm{g}}u_h\|_{\dagger,1,\mathcal{T}_h}\|\boldsymbol{A}\boldsymbol{\nabla}\chi-\pi_h^{\mathrm{d}}(\boldsymbol{A}\boldsymbol{\nabla}\chi)\|_{\dagger,\mathrm{div},\mathcal{T}_h},$$

and the result follows from the definitions of $\widetilde{\gamma}_{\mathrm{ba,g}}$ and $\widetilde{\gamma}_{\mathrm{ba,d}}$ given in (2.14c). □

We can now complete the Schatz argument, leading to quasi-optimality of the discrete solution for sufficiently refined meshes.

Theorem 4.2 (*A priori* estimate)**.** *If $\gamma_{\mathrm{ba,g}} < 1$ and $\sqrt{2}\gamma_{\mathrm{ba,d}} \leqslant \rho$, then there exists a unique solution $u_h \in \mathcal{P}_p(\mathcal{T}_h)$ to (2.11) and the estimate*

$$\|u-u_h\|_{\omega,\mathcal{T}_h} \leqslant \frac{1}{1-\gamma^2_{\mathrm{ba,g}}} \left(\min_{v_h\in\mathcal{P}_p(\mathcal{T}_h)\cap H^1_{\Gamma_{\mathrm{D}}}(\Omega)}\|u-v_h\|^2_{\omega,\Omega} + \frac{1}{\rho^2}\min_{\boldsymbol{\sigma}_h\in\boldsymbol{\mathcal{P}}_p(\mathcal{T}_h)\cap\boldsymbol{X}_{\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega}\|\boldsymbol{A}\boldsymbol{\nabla}u-\boldsymbol{\sigma}_h\|^2_{\dagger,\mathrm{div},\mathcal{T}_h}\right)^{1/2} \tag{4.4}$$

holds true. In addition, if $\sqrt{2}\gamma_{\mathrm{ba,d}} < \rho$, we also have

$$\|u_h-\pi^{\mathrm{g}}u_h\|_{\dagger,1,\mathcal{T}_h} \leqslant \left(\frac{1}{(1-\gamma^2_{\mathrm{ba,g}})(\rho^2-2\gamma^2_{\mathrm{ba,d}})}\right)^{1/2} \left(\min_{v_h\in\mathcal{P}_p(\mathcal{T}_h)\cap H^1_{\Gamma_{\mathrm{D}}}(\Omega)}\|u-v_h\|^2_{\omega,\Omega} + \frac{1}{\rho^2}\min_{\boldsymbol{\sigma}_h\in\boldsymbol{\mathcal{P}}_p(\mathcal{T}_h)\cap\boldsymbol{X}_{\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega}\|\boldsymbol{A}\boldsymbol{\nabla}u-\boldsymbol{\sigma}_h\|^2_{\dagger,\mathrm{div},\mathcal{T}_h}\right)^{1/2}. \tag{4.5}$$

Remark 4.3 (Higher-order term)**.** The last term appearing in the right-hand side (4.4) is unpleasant, since in general the finite element error is not solely controlled by the best-approximation error. Nevertheless, this term exhibits the right convergence rate. Besides, it can be made higher-order for sufficiently (piecewise) smooth solutions u. Indeed, the polynomial degree p appearing in the approximation of $\boldsymbol{A}\boldsymbol{\nabla}u$ only depends on the polynomial degree chosen for the gradient reconstruction $\mathfrak{G}(u_h)$, not the polynomial degree of the discretization space $\mathcal{P}_p(\mathcal{T}_h)$. It is therefore possible to reconstruct $\mathfrak{G}(u_h)$ in, *e.g.*, $\boldsymbol{\mathcal{P}}_{p+1}(\mathcal{T}_h)$ to make the undesired term in (4.4) negligible. Notice that this argument works if the stabilization parameter is sufficiently large according to (4.1).

Proof. We first observe that

$$\begin{aligned}\operatorname{Re} b_h(e_h,e_h) + \omega\|e_h\|^2_{\gamma,\mathcal{F}_h^{\mathrm{R}}} + 2\omega^2\|e_h\|^2_{\mu,\mathcal{T}_h} &= \|e_h\|^2_{\omega,\mathcal{T}_h} + \operatorname{Re} s_h(u_h,u_h)\\ &\geqslant \|e_h\|^2_{\omega,\mathcal{T}_h} + \rho^2\|u_h-\pi^{\mathrm{g}}u_h\|^2_{\dagger,1,\mathcal{T}_h}.\end{aligned}$$

On the other hand, for $v_h \in \mathcal{P}_p(\mathcal{T}_h)\cap H^1_{\Gamma_{\mathrm{D}}}(\Omega)$, we have

$$b_h(e_h,e_h) = b_h(e_h,u-v_h) + b_h(e_h,v_h-u_h) = b_h(e,u-v_h) + b_h(u,v_h-u_h) - (f,v_h-u_h)_{\mathcal{T}_h}.$$

Since a_h is Hermitian, Corollary 3.2 ensures that

$$\begin{aligned}|b_h(u,v_h-u_h)-(f,v_h-u_h)_{\mathcal{T}_h}| &\leqslant |a_h(u,v_h-u_h) - (\boldsymbol{\nabla}\cdot(\boldsymbol{A}\boldsymbol{\nabla}u),v_h-u_h)_{\mathcal{T}_h}|\\ &\leqslant \|u_h-\pi_h^{\mathrm{g}}u_h\|_{\dagger,1,\mathcal{T}_h}\min_{\boldsymbol{\sigma}_h\in\boldsymbol{\mathcal{P}}_p(\mathcal{T}_h)\cap\boldsymbol{X}_{\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega}\|\boldsymbol{A}\boldsymbol{\nabla}u-\boldsymbol{\sigma}_h\|_{\dagger,\mathrm{div},\mathcal{T}_h}.\\ &\leqslant \frac{\rho^2}{2}\|u_h-\pi_h^{\mathrm{g}}u_h\|^2_{\dagger,1,\mathcal{T}_h} + \frac{1}{2\rho^2}\min_{\boldsymbol{\sigma}_h\in\boldsymbol{\mathcal{P}}_p(\mathcal{T}_h)\cap\boldsymbol{X}_{\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega}\|\boldsymbol{A}\boldsymbol{\nabla}u-\boldsymbol{\sigma}_h\|^2_{\dagger,\mathrm{div},\mathcal{T}_h}.\end{aligned}$$

As a result, we have

$$
\begin{aligned}
\|e_h\|_{\omega,\mathcal{T}_h}^2 + \frac{1}{2}\rho^2\|u_h - \pi^{\mathrm{g}} u_h\|_{\dagger,1,\mathcal{T}_h}^2 &\leqslant |b_h(e_h, u - v_h)| + 2\omega^2\|e_h\|_{\mu,\mathcal{T}_h}^2 + \omega\|e_h\|_{\gamma,\mathcal{F}_h^{\mathrm{R}}}^2 \\
&\quad + \frac{1}{2\rho^2} \min_{\boldsymbol{\sigma}_h \in \boldsymbol{\mathcal{P}}_p(\mathcal{T}_h)\cap \boldsymbol{X}_{\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega} \|\boldsymbol{A}\boldsymbol{\nabla} u - \boldsymbol{\sigma}_h\|_{\dagger,\mathrm{div},\mathcal{T}_h}^2 \\
&\leqslant \|e_h\|_{\omega,\mathcal{T}_h}\|u - v_h\|_{\omega,\Omega} \\
&\quad + 2\left(\check{\gamma}_{\mathrm{ba,g}}\|e_h\|_{\omega,\mathcal{T}_h} + \check{\gamma}_{\mathrm{ba,d}}\|u_h - \pi^{\mathrm{g}} u_h\|_{\dagger,1,\mathcal{T}_h}\right)^2 \\
&\quad + \left(\tilde{\gamma}_{\mathrm{ba,g}}\|e_h\|_{\omega,\mathcal{T}_h} + \tilde{\gamma}_{\mathrm{ba,d}}\|u_h - \pi^{\mathrm{g}} u_h\|_{\dagger,1,\mathcal{T}_h}\right)^2 \\
&\quad + \frac{1}{2\rho^2} \min_{\boldsymbol{\sigma}_h \in \boldsymbol{\mathcal{P}}_p(\mathcal{T}_h)\cap \boldsymbol{X}_{\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega} \|\boldsymbol{A}\boldsymbol{\nabla} u - \boldsymbol{\sigma}_h\|_{\dagger,\mathrm{div},\mathcal{T}_h}^2 \\
&\leqslant \|e_h\|_{\omega,\mathcal{T}_h}\|u - v_h\|_{\omega,\Omega} \\
&\quad + (4\check{\gamma}_{\mathrm{ba,g}} + 2\tilde{\gamma}_{\mathrm{ba,g}})\|e_h\|_{\omega,\mathcal{T}_h}^2 + (4\check{\gamma}_{\mathrm{ba,d}} + 2\tilde{\gamma}_{\mathrm{ba,d}})\|u_h - \pi^{\mathrm{g}} u_h\|_{\dagger,1,\mathcal{T}_h}^2 \\
&\quad + \frac{1}{2\rho^2} \min_{\boldsymbol{\sigma}_h \in \boldsymbol{\mathcal{P}}_p(\mathcal{T}_h)\cap \boldsymbol{X}_{\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega} \|\boldsymbol{A}\boldsymbol{\nabla} u - \boldsymbol{\sigma}_h\|_{\dagger,\mathrm{div},\mathcal{T}_h}^2
\end{aligned}
$$

so that

$$
\begin{aligned}
&(1 - 4\check{\gamma}_{\mathrm{ba,g}}^2 - 2\tilde{\gamma}_{\mathrm{ba,g}}^2)\|e_h\|_{\omega,\mathcal{T}_h}^2 + \frac{1}{2}(\rho^2 - 8\check{\gamma}_{\mathrm{ba,d}}^2 - 4\tilde{\gamma}_{\mathrm{ba,d}}^2)\|u_h - \pi_h^{\mathrm{g}} u_h\|_{\dagger,1,\mathcal{T}_h}^2 \\
&\qquad\qquad \leqslant \|e_h\|_{\omega,\mathcal{T}_h}\|u - v_h\|_{\omega,\Omega} + \frac{1}{2\rho^2} \min_{\boldsymbol{\sigma}_h \in \boldsymbol{\mathcal{P}}_p(\mathcal{T}_h)\cap \boldsymbol{X}_{\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega} \|\boldsymbol{A}\boldsymbol{\nabla} u - \boldsymbol{\sigma}_h\|_{\dagger,\mathrm{div},\mathcal{T}_h}^2
\end{aligned}
$$

and

$$
\begin{aligned}
&(1 - 4\check{\gamma}_{\mathrm{ba,g}}^2 - 2\tilde{\gamma}_{\mathrm{ba,g}}^2)\|e_h\|_{\omega,\mathcal{T}_h}^2 + (\rho^2 - 8\check{\gamma}_{\mathrm{ba,d}}^2 - 4\tilde{\gamma}_{\mathrm{ba,d}}^2)\|u_h - \pi_h^{\mathrm{g}} u_h\|_{\dagger,1,\mathcal{T}_h}^2 \\
&\qquad\qquad \leqslant \|e_h\|^2 + \frac{1}{\rho^2} \min_{\boldsymbol{\sigma}_h \in \boldsymbol{\mathcal{P}}_p(\mathcal{T}_h)\cap \boldsymbol{X}_{\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega} \|\boldsymbol{A}\boldsymbol{\nabla} u - \boldsymbol{\sigma}_h\|_{\dagger,\mathrm{div},\mathcal{T}_h}^2.
\end{aligned}
$$

At this point, (4.4) and (4.5) follow recalling the definitions of $\gamma_{\mathrm{ba,g}}$ and $\gamma_{\mathrm{ba,d}}$ in (2.14). □

5. *A POSTERIORI* ANALYSIS

We now provide *a posteriori* error estimates under minimal regularity assumptions. We first present an abstract framework, and then show how it can be applied to the particular case of residual-based estimators.

Throughout this section, we assume that u_h is a fixed function in $\mathcal{P}_p(\mathcal{T}_h)$ satisfying (2.11). Notice that unique solvability of (2.11) is not required. In particular, the proposed analysis applies without any restriction on the mesh size or polynomial degree. Also, in contrast to the *a priori* analysis, we do not need any specific positivity assumption on the stabilisation form s_h.

5.1. Abstract reliability analysis

Following Theorem 3.3 from [23], the discretisation error may be bounded using two different terms that respectively control the equation residual, and the non-conformity of the discrete solution. For the equation residual, we introduce the residual functional and its norm

$$
\langle \mathscr{R}_{\mathrm{r}}, v\rangle := b_h(e_h, v) \quad \forall v \in H^1_{\Gamma_{\mathrm{D}}}(\Omega), \qquad \mathrm{R}_{\mathrm{r}} := \sup_{\substack{v \in H^1_{\Gamma_{\mathrm{D}}}(\Omega) \\ \|\boldsymbol{\nabla} v\|_{\boldsymbol{A},\Omega} = 1}} |\langle \mathscr{R}_{\mathrm{r}}, v\rangle|. \tag{5.1a}
$$

Notice that this first term only involves conforming test functions. To account for the non-conformity of the discrete solution, we additionally introduce

$$\mathrm{R_c} := \|u_h - \pi^{\mathrm{g}} u_h\|_{\dagger,1,\mathcal{T}_h}. \tag{5.1b}$$

The total abstract estimator is then the Hilbertian sum of the two above components:

$$\mathrm{R}^2 := \mathrm{R}_\mathrm{r}^2 + \mathrm{R}_\mathrm{c}^2. \tag{5.1c}$$

Following [14, 40], our *a posteriori* analysis follows the lines of the Schatz argument [41]. Specifically, we start by controlling weak norms of the errors with the estimator and the approximation factor using a duality argument.

Lemma 5.1 (Aubin-Nitsche)**.** *The following estimates hold true:*

$$\omega\|u - u_h\|_{\mu,\mathcal{T}_h} \leqslant \breve{\gamma}_{\mathrm{ba}}\mathrm{R}, \qquad \omega^{1/2}\|u - u_h\|_{\gamma,\mathcal{F}_h^{\mathrm{R}}} \leqslant \widetilde{\gamma}_{\mathrm{ba}}\mathrm{R}. \tag{5.2}$$

Proof. Let us set $e_h := u - u_h$. We start with the first estimate in (5.2). We define ξ as the unique element of $H^1_{\Gamma_\mathrm{D}}(\Omega)$ such that $b(w,\xi) = \omega(\mu w, e_h)_{\mathcal{T}_h}$ for all $w \in H^1_{\Gamma_\mathrm{D}}(\Omega)$. Following the lines the proof of Lemma 4.1, we write

$$\begin{aligned} b_h(e_h,\xi) &= -\omega^2(\mu e_h,\xi)_{\mathcal{T}_h} - i\omega(\gamma e_h,\xi)_{\mathcal{F}_h^{\mathrm{R}}} + (\mathfrak{G}(e_h),\boldsymbol{A}\boldsymbol{\nabla}\xi)_{\mathcal{T}_h} \\ &= (e_h, -\omega^2\mu\xi - \boldsymbol{\nabla}\cdot(\boldsymbol{A}\boldsymbol{\nabla}\xi))_{\mathcal{T}_h} - (e_h,\boldsymbol{A}\boldsymbol{\nabla}\xi\cdot\boldsymbol{n})_{\mathcal{F}_h^{\mathrm{R}}} + (e_h,\boldsymbol{\nabla}\cdot(\boldsymbol{A}\boldsymbol{\nabla}\xi))_{\mathcal{T}_h} + (\mathfrak{G}(e_h),\boldsymbol{A}\boldsymbol{\nabla}\xi)_{\mathcal{T}_h} \\ &= \omega\|e_h\|^2_{\mu,\mathcal{T}_h} + (\mathfrak{G}(\pi^{\mathrm{g}}u_h - u_h),\boldsymbol{A}\boldsymbol{\nabla}\xi - \boldsymbol{\sigma}_h)_{\mathcal{T}_h} \\ &\quad + (\pi^{\mathrm{g}}u_h - u_h,\boldsymbol{\nabla}\cdot(\boldsymbol{A}\boldsymbol{\nabla}\xi - \boldsymbol{\sigma}_h))_{\mathcal{T}_h} - (\pi^{\mathrm{g}}u_h - u_h,(\boldsymbol{A}\boldsymbol{\nabla}\xi - \boldsymbol{\sigma}_h)\cdot\boldsymbol{n})_{\mathcal{F}_h^{\mathrm{R}}}, \end{aligned}$$

for all $\boldsymbol{\sigma}_h \in \boldsymbol{\mathcal{P}}_p(\mathcal{T}_h) \cap \boldsymbol{X}_{\Gamma_\mathrm{N}}(\mathrm{div},\Omega)$, where we employed (3.1) in the last identity. It follows that

$$\omega\|e_h\|^2_{\mu,\mathcal{T}_h} \leqslant |b_h(e_h,\xi)| + \|u_h - \pi^{\mathrm{g}}u_h\|_{\dagger,1,\mathcal{T}_h}\|\boldsymbol{A}\boldsymbol{\nabla}\xi - \boldsymbol{\sigma}_h\|_{\dagger,\mathrm{div},\mathcal{T}_h}.$$

On the one hand, recalling the Galerkin orthogonality property stated in (2.12) and the definition of $\mathrm{R_r}$ at (5.1a), we have

$$|b_h(e_h,\xi)| = |b_h(e_h,\xi - \pi_h^{\mathrm{g}}\xi)| \leqslant \mathrm{R_r}\|\boldsymbol{\nabla}(\xi - \pi_h^{\mathrm{g}}\xi)\|_{\boldsymbol{A},\Omega} \leqslant \breve{\gamma}_{\mathrm{ba,g}}\mathrm{R_r}\|e_h\|_{\mu,\mathcal{T}_h}.$$

On the other hand, recalling (2.14b) and (5.1b), we have

$$\|u_h - \pi^{\mathrm{g}}u_h\|_{\dagger,1,\mathcal{T}_h}\|\boldsymbol{A}\boldsymbol{\nabla}\xi - \pi_h^{\mathrm{d}}(\boldsymbol{A}\boldsymbol{\nabla}\xi)\|_{\dagger,\mathrm{div},\mathcal{T}_h} \leqslant \breve{\gamma}_{\mathrm{ba,d}}\mathrm{R_c}\|e_h\|_{\mu,\mathcal{T}_h}.$$

Then, the first estimate of (5.2) follows from (5.1c) since

$$\breve{\gamma}_{\mathrm{ba,g}}\mathrm{R_r} + \breve{\gamma}_{\mathrm{ba,d}}\mathrm{R_c} \leqslant \left(\breve{\gamma}^2_{\mathrm{ba,g}} + \breve{\gamma}^2_{\mathrm{ba,d}}\right)^{1/2}\left(\mathrm{R_r}^2 + \mathrm{R_c}^2\right)^{1/2} = \breve{\gamma}_{\mathrm{ba}}\mathrm{R}.$$

For the sake of shortness, we do not detail the proof of the second estimate in (5.2) as it follows the lines of the first estimate but using the function χ from the proof of Lemma 4.1 instead of ξ. □

We can now conclude the abstract reliability analysis. Following [23], our proof uses a Pythagorean identity to separate the conforming and non-conforming parts of the error.

Theorem 5.2 (Abstract reliability)**.** *We have*

$$|\!|\!| u - u_h |\!|\!|_{\omega,\mathcal{T}_h} \leqslant \sqrt{1 + \gamma^2_{\mathrm{ba}}}\,\mathrm{R}. \tag{5.3}$$

Proof. For the sake of simplicity, we introduce

$$b_h^+(\phi, v) := \omega^2(\mu\phi, v)_{\mathcal{T}_h} + \omega(\gamma\phi, v)_{\mathcal{F}_h^{\mathrm{R}}} + (\boldsymbol{A}\mathfrak{G}(\phi), \mathfrak{G}(v))_{\mathcal{T}_h} \qquad \forall \phi, v \in H^1(\mathcal{T}_h),$$

and we define $\widetilde{u}$ as the only element of $H^1_{\Gamma_{\mathrm{D}}}(\Omega)$ such that

$$b_h^+(\widetilde{u} - u_h, v) = 0 \qquad \forall v \in H^1_{\Gamma_{\mathrm{D}}}(\Omega).$$

Notice that this indeed a well-formed definition, since b_h^+ is equivalent to the usual inner-product of $H^1_{\Gamma_{\mathrm{D}}}(\Omega)$. Classically, we have the Pythagorean identity

$$\begin{aligned}\|u - u_h\|^2_{\omega,\mathcal{T}_h} &= \|\widetilde{u} - u_h\|^2_{\omega,\mathcal{T}_h} + 2\operatorname{Re} b^+(\widetilde{u} - u_h, u - \widetilde{u}) + \|u - \widetilde{u}\|^2_{\omega,\mathcal{T}_h} \\ &= \|u_h - \widetilde{u}\|^2_{\omega,\mathcal{T}_h} + \|u - \widetilde{u}\|^2_{\omega,\mathcal{T}_h}.\end{aligned} \tag{5.4}$$

Then, on the one hand, it is clear using (2.6) that

$$\|u_h - \widetilde{u}\|_{\omega,\mathcal{T}_h} \leqslant \|u_h - \pi^{\mathrm{g}} u_h\|_{\omega,\mathcal{T}_h} \leqslant \|u_h - \pi^{\mathrm{g}} u_h\|_{\dagger,1,\mathcal{T}_h} = \mathrm{R}_{\mathrm{c}}, \tag{5.5}$$

and on the other hand, we have

$$\begin{aligned}\|u - \widetilde{u}\|^2_{\omega,\Omega} &= b_h^+(u - \widetilde{u}, u - \widetilde{u}) = b_h^+(e_h, u - \widetilde{u}) \\ &= b_h(e_h, u - \widetilde{u}) + 2\omega^2(\mu e_h, u - \widetilde{u})_{\mathcal{T}_h} + (1 - i)\omega(\gamma e_h, u - \widetilde{u})_{\mathcal{F}_h^{\mathrm{R}}} \\ &\leqslant \mathrm{R}_{\mathrm{r}}\|\boldsymbol{\nabla}(u - \widetilde{u})\|_{\boldsymbol{A},\Omega} + 2\check{\gamma}_{\mathrm{ba}}\mathrm{R}\omega\|u - \widetilde{u}\|_{\mu,\mathcal{T}_h} + \sqrt{2}\widetilde{\gamma}_{\mathrm{ba}}\mathrm{R}\omega^{1/2}\|u - \widetilde{u}\|_{\gamma,\mathcal{F}_h^{\mathrm{R}}} \\ &\leqslant \left(\mathrm{R}_{\mathrm{r}}^2 + (4\check{\gamma}_{\mathrm{ba}}^2 + 2\widetilde{\gamma}_{\mathrm{ba}}^2)\mathrm{R}^2\right)^{1/2} \|u - \widetilde{u}\|_{\omega,\Omega},\end{aligned}$$

and recalling (2.14c), we arrive at

$$\|u - \widetilde{u}\|^2_{\omega,\Omega} \leqslant \mathrm{R}_{\mathrm{r}}^2 + (4\check{\gamma}_{\mathrm{ba}}^2 + 2\widetilde{\gamma}_{\mathrm{ba}}^2)\mathrm{R}^2 = \mathrm{R}_{\mathrm{r}}^2 + \gamma_{\mathrm{ba}}^2\mathrm{R}^2. \tag{5.6}$$

The estimate in (5.3) then follows from (5.4), (5.5) and (5.6). □

5.2. Residual-based *a posteriori* estimator

The residuals R_{r} and R_{c} can be straightforwardly controlled using flux and potential reconstructions. We refer the reader to [17,23] for details about these constructions. Here, we focus instead on residual-based estimators, for which the link may be less clear. In this section, the letter C refers to a generic constant that may change from one occurrence to the other, and that only depends on the contrasts of the coefficients μ, $\boldsymbol{A}$ and γ as well as the polynomial degree p and the shape-regularity parameter κ. The stability result for the lifting operator

$$\|\boldsymbol{\ell}_h([\![v_h]\!])\|^2_{\boldsymbol{A},\mathcal{T}_h} \leqslant C \sum_{K \in \mathcal{T}_h} \frac{\alpha_K}{h_K} \|[\![v_h]\!]\|^2_{\partial K} \quad \forall v_h \in \mathcal{P}_p(\mathcal{T}_h) \tag{5.7}$$

can be found, *e.g.*, in Section 4.3 from [19] or Lemma 4.1 from [33].

We first establish that R_{c} can be controlled by the jumps of u_h. The notations

$$\frac{\omega h_F^\star}{\vartheta_F^\star} := \max_{F \in \mathcal{F}_h^{\mathrm{R}}} \frac{\omega h_F}{\vartheta_F}, \qquad \frac{\omega h_K^\star}{\vartheta_K^\star} := \max_{K \in \mathcal{T}_h} \frac{\omega h_K}{\vartheta_K},$$

will be useful.

Lemma 5.3 (Control of the non-conformity)**.** *We have*

$$\mathrm{R}_{\mathrm{c}}^2 \leqslant C \max\left(1, \frac{\omega h_F^\star}{\vartheta_F^\star}, \left(\frac{\omega h_K^\star}{\vartheta_K^\star}\right)^2\right) \sum_{K \in \mathcal{T}_h} \frac{\alpha_K}{h_K} \|[\![u_h]\!]\|^2_{\partial K \setminus (\Gamma_{\mathrm{N}} \cup \Gamma_{\mathrm{R}})}. \tag{5.8}$$

Proof. We start with two estimates that are easily inferred from Lemma 4.3 from [22]. For all $v_h \in \mathcal{P}_p(\mathcal{T}_h)$, there exists a conforming approximation $\mathcal{J}v_h \in \mathcal{P}_p(\mathcal{T}_h) \cap H^1_{\Gamma_\mathrm{D}}(\Omega)$ such that and all $K \in \mathcal{T}_h$, we have

$$\|v_h - \mathcal{J}v_h\|_K^2 \leqslant C h_K \sum_{K' \in \widetilde{\mathcal{T}}_K} \|[\![v_h]\!]\|^2_{\partial K'}, \qquad \|\boldsymbol{\nabla}(v_h - \mathcal{J}v_h)\|_K^2 \leqslant C \frac{1}{h_K} \sum_{K' \in \widetilde{\mathcal{T}}_K} \|[\![v_h]\!]\|^2_{\partial K'},$$

where $\widetilde{\mathcal{T}}_K$ collects those elements $K' \in \mathcal{T}_h$ sharing at least one vertex with K. Employing the multiplicative trace inequality

$$\|\theta\|^2_{\partial K} \leqslant C \left\{ \frac{1}{h_K}\|\theta\|_K^2 + \|\theta\|_K \|\boldsymbol{\nabla}\theta\|_K \right\} \leqslant C \left\{ \frac{1}{h_K}\|\theta\|_K^2 + h_K \|\boldsymbol{\nabla}\theta\|_K^2 \right\} \qquad \forall \theta \in H^1(K),$$

we also have

$$\|v_h - \mathcal{J}v_h\|^2_{\partial K} \leqslant C \sum_{K' \in \widetilde{\mathcal{T}}_K} \|[\![v_h]\!]\|^2_{\partial K'}.$$

It is therefore clear that

$$\|u_h - \mathcal{J}u_h\|^2_{\dagger,1,\mathcal{T}_h} \leqslant C \max\left(1, \frac{\omega h_F^\star}{\vartheta_F^\star}, \left(\frac{\omega h_K^\star}{\vartheta_K^\star}\right)^2\right) \sum_{K \in \mathcal{T}_h} \frac{\alpha_K}{h_K} \|[\![u_h]\!]\|^2_{\partial K},$$

and (5.8) follows from the definition of $\mathrm{R_c}$ in (5.1b). □

We then control the conforming residual $\mathrm{R_r}$. This is classically done by combining element-wise integration by parts and a quasi-interpolation operator.

Lemma 5.4 (Control of the residual)**.** *We have*

$$\begin{aligned}\mathrm{R_r^2} \leqslant C\Bigg\{ \sum_{K \in \mathcal{T}_h} &\left(\frac{h_K^2}{\alpha_K} \|f + \omega^2 \mu u_h + \boldsymbol{\nabla}\cdot(\boldsymbol{A}\boldsymbol{\nabla}u_h)\|^2_{\mathcal{T}_h} + \frac{h_K}{\alpha_K} \|[\![\boldsymbol{A}\boldsymbol{\nabla}u_h]\!]\cdot\boldsymbol{n}_K\|^2_{\partial K \setminus \partial\Omega} + \frac{\alpha_K}{h_K}\|[\![u_h]\!]\|^2_{\partial K\setminus\partial\Omega} \right) \\ &+ \sum_{F \in \mathcal{F}_h^\mathrm{D}} \frac{\alpha_F}{h_F}\|u_h\|_F^2 + \sum_{F \in \mathcal{F}_h^\mathrm{N}} \frac{h_F}{\alpha_F}\|\boldsymbol{A}\boldsymbol{\nabla}u_h\cdot\boldsymbol{n}\|_F^2 + \sum_{F \in \mathcal{F}_h^\mathrm{R}} \frac{h_F}{\alpha_F}\|\boldsymbol{A}\boldsymbol{\nabla}u_h\cdot\boldsymbol{n} - i\omega\gamma u_h\|_F^2 \Bigg\}. \end{aligned} \tag{5.9}$$

Proof. We consider an arbitrary $v \in H^1_{\Gamma_\mathrm{D}}(\Omega)$, and let $\widetilde{v} := v - \mathcal{Q}_h v \in H^1_{\Gamma_\mathrm{D}}(\Omega)$, where $\mathcal{Q}_h$ is a quasi-interpolation operator satisfying

$$\frac{1}{h_K^2}\|v - \mathcal{Q}_h v\|_K^2 + \frac{1}{h_K}\|v - \mathcal{Q}_h v\|^2_{\partial K} + \|\boldsymbol{\nabla}(v - \mathcal{Q}_h v)\|_K^2 \leqslant C \frac{1}{\alpha_K}\|\boldsymbol{\nabla}v\|^2_{\boldsymbol{A},\widetilde{K}},$$

with $\widetilde{K}$ the domain corresponding to all the elements $K' \in \mathcal{T}_h$ that share at least one vertex with K. We refer the reader to, *e.g.*, Theorem 5.2 from [22], for the construction of $\mathcal{Q}_h$. Thanks to Galerkin orthogonality (2.12), we can write that

$$\begin{aligned}\langle \mathscr{R}_\mathrm{r}, v\rangle &= b_h(e_h, v) = b_h(e_h, \widetilde{v}) \\ &= (f, \widetilde{v})_\Omega + \omega^2(\mu u_h, \widetilde{v})_{\mathcal{T}_h} + i\omega(\gamma u_h, \widetilde{v})_{\mathcal{F}_h^\mathrm{R}} - (\boldsymbol{A}\mathfrak{G}(u_h), \boldsymbol{\nabla}\widetilde{v})_{\mathcal{T}_h} \\ &= \underbrace{(f, \widetilde{v})_\Omega + \omega^2(\mu u_h, \widetilde{v})_{\mathcal{T}_h} + i\omega(\gamma u_h, \widetilde{v})_{\mathcal{F}_h^\mathrm{R}} - (\boldsymbol{A}\boldsymbol{\nabla}u_h, \boldsymbol{\nabla}\widetilde{v})_{\mathcal{T}_h}}_{r_1} + \underbrace{(\boldsymbol{A}\boldsymbol{\ell}_h([\![u_h]\!]), \boldsymbol{\nabla}\widetilde{v})_{\mathcal{T}_h}}_{r_2}. \end{aligned} \tag{5.10}$$

Then, on the one hand, we see that

$$\begin{aligned} r_1 &= (f + \omega^2 \mu u_h + \boldsymbol{\nabla} \cdot (\boldsymbol{A} \boldsymbol{\nabla} u_h), \widetilde{v})_{\mathcal{T}_h} + i\omega(\gamma u_h, \widetilde{v})_{\mathcal{F}_h^{\mathrm{R}}} - \sum_{K \in \mathcal{T}_h} (\boldsymbol{A} \boldsymbol{\nabla} u_h \cdot \boldsymbol{n}_K, \widetilde{v})_{\partial K} \\ &= (f + \omega^2 \mu u_h + \boldsymbol{\nabla} \cdot (\boldsymbol{A} \boldsymbol{\nabla} u_h), \widetilde{v})_{\mathcal{T}_h} - \sum_{F \in \mathcal{F}_h^{\mathrm{i}}} ([\![\boldsymbol{A} \boldsymbol{\nabla} u_h]\!] \cdot \boldsymbol{n}_F, \widetilde{v})_F \\ &- (\boldsymbol{A} \boldsymbol{\nabla} u_h \cdot \boldsymbol{n}, \widetilde{v})_{\mathcal{F}_h^{\mathrm{N}}} - (\boldsymbol{A} \boldsymbol{\nabla} u_h \cdot \boldsymbol{n} - i\omega\gamma u_h, \widetilde{v})_{\mathcal{F}_h^{\mathrm{R}}}, \end{aligned}$$

and therefore,

$$\begin{aligned} |r_1| &\leqslant \sum_{K \in \mathcal{T}_h} \left(\|f + \omega^2 \mu u_h + \boldsymbol{\nabla} \cdot (\boldsymbol{A} \boldsymbol{\nabla} u_h)\|_K \|\widetilde{v}\|_K + \|[\![\boldsymbol{A} \boldsymbol{\nabla} u_h]\!] \cdot \boldsymbol{n}_K\|_{\partial K} \|\widetilde{v}\|_{\partial K} \right) \\ &+ \sum_{F \in \mathcal{F}_h^{\mathrm{N}}} \|\boldsymbol{A} \boldsymbol{\nabla} u_h \cdot \boldsymbol{n}\|_F \|\widetilde{v}\|_F + \sum_{F \in \mathcal{F}_h^{\mathrm{R}}} \|\boldsymbol{A} \boldsymbol{\nabla} u_h \cdot \boldsymbol{n} - i\omega\gamma u_h\|_F \|\widetilde{v}\|_F \\ &\leqslant C \left\{ \sum_{K \in \mathcal{T}_h} \left(\frac{h_K^2}{\alpha_K} \|f + \omega^2 \mu u_h + \boldsymbol{\nabla} \cdot (\boldsymbol{A} \boldsymbol{\nabla} u_h)\|_K^2 + \frac{h_K}{\alpha_K} \|[\![\boldsymbol{\nabla} u_h]\!] \cdot \boldsymbol{n}_K\|_{\partial K}^2 \right) \right. \\ &\quad \left. + \sum_{F \in \mathcal{F}_h^{\mathrm{N}}} \frac{h_F}{\alpha_F} \|\boldsymbol{A} \boldsymbol{\nabla} u_h \cdot \boldsymbol{n}\|_F^2 + \sum_{F \in \mathcal{F}_h^{\mathrm{R}}} \frac{h_F}{\alpha_F} \|\boldsymbol{A} \boldsymbol{\nabla} u_h \cdot \boldsymbol{n} - i\omega\gamma u_h\|_F^2 \right\}^{1/2} \|\boldsymbol{\nabla} v\|_{\boldsymbol{A},\Omega}. \end{aligned} \tag{5.11}$$

On the other hand, we have

$$|r_2| \leqslant \|\boldsymbol{\ell}_h([\![u_h]\!])\|_{\boldsymbol{A},\Omega} \|\boldsymbol{\nabla} \widetilde{v}\|_{\boldsymbol{A},\Omega} \leqslant C \left\{ \sum_{K \in \mathcal{T}_h} \frac{\alpha_K}{h_K} \|[\![u_h]\!]\|_{\partial K \setminus \partial\Omega}^2 + \sum_{F \in \mathcal{F}_h^{\mathrm{D}}} \frac{\alpha_F}{h_F} \|[\![u_h]\!]\|_F^2 \right\}^{1/2} \|\boldsymbol{\nabla} v\|_{\boldsymbol{A},\Omega}, \tag{5.12}$$

and (5.9) follows from (5.10), (5.11) and (5.12) recalling the definition of $\mathrm{R_r}$ in (5.1a). □

We therefore define a residual-based estimator by

$$\begin{aligned} \eta^2 := &\sum_{K \in \mathcal{T}_h} \left(\frac{h_K^2}{\alpha_K} \|f + \omega^2 \mu u_h + \boldsymbol{\nabla} \cdot (\boldsymbol{A} \boldsymbol{\nabla} u_h)\|_{\mathcal{T}_h}^2 + \frac{h_K}{\alpha_K} \|[\![\boldsymbol{A} \boldsymbol{\nabla} u_h]\!] \cdot \boldsymbol{n}_K\|_{\partial K \setminus \partial\Omega}^2 + \frac{\alpha_K}{h_K} \|[\![u_h]\!]\|_{\partial K \setminus \partial\Omega}^2 \right) \\ &+ \sum_{F \in \mathcal{F}_h^{\mathrm{D}}} \frac{\alpha_F}{h_F} \|u_h\|_F^2 + \sum_{F \in \mathcal{F}_h^{\mathrm{N}}} \frac{h_F}{\alpha_F} \|\boldsymbol{A} \boldsymbol{\nabla} u_h \cdot \boldsymbol{n}\|_F^2 + \sum_{F \in \mathcal{F}_h^{\mathrm{R}}} \frac{h_F}{\alpha_F} \|\boldsymbol{A} \boldsymbol{\nabla} u_h \cdot \boldsymbol{n} - i\omega\gamma u_h\|_F^2. \end{aligned}$$

As a direct consequence of (5.3), (5.8) and (5.9), we obtain a reliability estimate stated in Theorem 5.5. It is to be compared with Theorem 2.3 from [14] and Theorem 3.6 from [40].

Theorem 5.5 (Reliability of the residual estimator)**.** *We have*

$$\|e_h\|_{\omega,\mathcal{T}_h} \leqslant C \left(1 + \left(\frac{\omega h_F^\star}{\vartheta_F^\star} \right)^{1/2} + \frac{\omega h_K^\star}{\vartheta_K^\star} \right) (1 + \gamma_{\mathrm{ba}}) \eta.$$

5.3. Efficiency

For the sake of completeness, we state an efficiency estimate for the residual-based estimator. The proof can be found in Section 3.3 from [40], and we also refer the reader to Section 4 from [17] for the corresponding result for equilibrated estimators.

Considering an element $K \in \mathcal{T}_h$, we need a few additional notation. The component of the error estimator associated with K reads

$$\eta_K^2 := \frac{h_K^2}{\alpha_K}\|f + \omega^2 \mu u_h + \boldsymbol{\nabla} \cdot (\boldsymbol{A}\boldsymbol{\nabla} u_h)\|_{\mathcal{T}_h}^2 + \frac{h_K}{\alpha_K}\|[\![\boldsymbol{A}\boldsymbol{\nabla} u_h]\!] \cdot \boldsymbol{n}_K\|_{\partial K \setminus \partial\Omega}^2 + \frac{\alpha_K}{h_K}\|[\![u_h]\!]\|_{\partial K \setminus \partial\Omega}^2 + \sum_{\substack{F \in \mathcal{F}_h^{\mathrm{D}} \\ F \subset \partial K}} \frac{\alpha_F}{h_F}\|u_h\|_F^2 + \sum_{\substack{F \in \mathcal{F}_h^{\mathrm{N}} \\ F \subset \partial K}} \frac{h_F}{\alpha_F}\|\boldsymbol{A}\boldsymbol{\nabla} u_h \cdot \boldsymbol{n}\|_F^2 + \sum_{\substack{F \in \mathcal{F}_h^{\mathrm{R}} \\ F \subset \partial K}} \frac{h_F}{\alpha_F}\|\boldsymbol{A}\boldsymbol{\nabla} u_h \cdot \boldsymbol{n} - i\omega\gamma u_h\|_F^2.$$

Then, to measure the discretization error locally around K, we introduce

$$|\!|\!| u - u_h |\!|\!|_{\omega,\widetilde{K}}^2 = \sum_{\substack{K' \in \mathcal{T}_h \\ \partial K' \cap \partial K \neq \varnothing}} \left\{\omega^2\|u - u_h\|_{\mu,K'}^2 + \omega\|u - u_h\|_{\gamma,\Gamma_{\mathrm{R}} \cap \partial K'}^2 + \|\boldsymbol{\nabla}(u - u_h)\|_{\boldsymbol{A},K'}^2\right\}.$$

Finally, the quantity

$$\operatorname{osc}_{\widetilde{K}}^2 := \sum_{\substack{K' \in \mathcal{T}_h \\ \partial K' \cap \partial K \neq \varnothing}} h_{K'}^2 \min_{f_p \in \mathcal{P}_p(K')} \|f - f_p\|_{\mu,K'}^2$$

is usually called the "data-oscillation term". It is of higher-order if f is piecewise smooth.

Theorem 5.6 (Efficiency of the residual estimator)**.** *For all $K \in \mathcal{T}_h$, we have*

$$\eta_K \leqslant C \left\{\left(1 + \max_{\substack{F \in \mathcal{F}_h^{\mathrm{R}} \\ F \subset \partial K}} \left(\frac{\omega h_F}{\vartheta_F}\right)^{1/2} + \frac{\omega h_K}{\vartheta_K}\right) |\!|\!| u - u_h |\!|\!|_{\omega,\widetilde{K}} + \operatorname{osc}_{\widetilde{K}}\right\}.$$

Appendix A. Approximation factors

In this section, we provide an estimate for the approximation factor γ_{ba} valid under minimal regularity assumptions. For the sake of simplicity and shortness, we restrict our attention to the case where $\Gamma_{\mathrm{R}} = \varnothing$. The reason behind this choice is twofold. Obviously, we automatically have $\widetilde{\gamma}_{\mathrm{ba,g}} = \widetilde{\gamma}_{\mathrm{ba,d}} = 0$ in this case, since there are no boundary right-hand sides. But more importantly, the norm $\|\cdot\|_{\dagger,\mathrm{div},\mathcal{T}_h}$ does not contain a boundary term. This allows us to easily use the quasi-interpolation operators introduced in [25] to upper-bound $\gamma_{\mathrm{ba,d}}$ (see the estimates in (A.2) below). Notice that, in principle, there is no obstruction to the construction of quasi-interpolation operators with good interpolation properties on the boundary if they are defined on $\boldsymbol{X}_{\Gamma_{\mathrm{N}}}(\mathrm{div}, \Omega)$. However, the author is not aware of any convenient reference where such an operator can be found. Alternatively, the canonical Raviart–Thomas interpolation operator has optimal approximation properties on Γ_{R}, but it is not defined under minimal regularity (see *e.g.* [10]).

Classically, the analysis of this section relies on (arbitrarily low) elliptic regularity shifts. To properly state them, we need to introduce (broken) fractional Sobolev norms [1]. First, for $K \in \mathcal{T}_h$, $\boldsymbol{w} \in \boldsymbol{L}^2(\Omega)$, and $s \in (0, 1)$, we set

$$|\boldsymbol{w}|_{H^s(K)}^2 := \int_K \int_K \frac{|\boldsymbol{w}(\boldsymbol{x}) - \boldsymbol{w}(\boldsymbol{y})|^2}{|\boldsymbol{x} - \boldsymbol{y}|^{2s+d}} d\boldsymbol{y} d\boldsymbol{x}$$

and, for $s = 1$,

$$|\boldsymbol{w}|_{H^1(K)}^2 := \sum_{n=1}^{d} \|\boldsymbol{\nabla} \boldsymbol{w}_n\|_K^2.$$

Then, the corresponding global norm reads

$$|\boldsymbol{w}|_{H^s(\mathcal{T}_h)}^2 := \sum_{K \in \mathcal{T}_h} |\boldsymbol{w}|_{H^s(K)}^2.$$

There exists $\theta \in (0,1]$ and such that such that all $v \in H^1_{\Gamma_{\rm D}}(\Omega)$ with $\boldsymbol{A}\boldsymbol{\nabla} v \in \boldsymbol{X}_{\Gamma_{\rm N}}(\Omega)$, we have

$$|\boldsymbol{A}\boldsymbol{\nabla} v|_{H^\theta(\mathcal{T}_h)} + \alpha_{\max}|\boldsymbol{\nabla} v|_{H^\theta(\mathcal{T}_h)} \leqslant C\ell^{1-\theta}\|\boldsymbol{\nabla}\cdot(\boldsymbol{A}\boldsymbol{\nabla} v)\|_\Omega, \tag{A.1}$$

where ℓ is the diameter of Ω, and $\alpha_{\max}$ is the essential supremum of the maximal eigenvalue of $\boldsymbol{A}$ *e.g.* [34]. For generic values of $\boldsymbol{A}$, θ may be arbitrarily close to zero. On the other hand, if we assume that $\boldsymbol{A}$ is constant and that either $\Gamma_{\rm D} = \varnothing$ or $\Gamma_{\rm N} = \varnothing$, (A.1) holds true with $\theta > 1/2$. Finally, if we further assume that Ω is convex, (A.1) does hold true with $\theta = 1$. We refer the reader to [30] for these last two statements.

As respectively constructed for instance in [22, 25], there exist (quasi-)interpolation operators $\mathcal{I}_h : L^2(\Omega) \to \mathcal{P}_p(\mathcal{T}_h) \cap H^1_{\Gamma_{\rm D}}(\Omega)$ and $\mathcal{J}_h : \boldsymbol{L}^2(\Omega) \to \boldsymbol{\mathcal{P}}_p(\mathcal{T}_h) \cap \boldsymbol{X}_{\Gamma_{\rm N}}(\mathrm{div},\Omega)$ such that

$$\sum_{K\in\mathcal{T}_h} \left(h_K^{-2}\|v - \mathcal{I}_h v\|_K^2 + \|\boldsymbol{\nabla}(v - \mathcal{I}_h v)\|_K^2\right) \leqslant Ch^{2\theta}|\boldsymbol{\nabla} v|^2_{H^\theta(\mathcal{T}_h)} \tag{A.2a}$$

and

$$\sum_{K\in\mathcal{T}_h} \left(\|\boldsymbol{w} - \mathcal{J}_h\boldsymbol{w}\|_K^2 + h_K^2\|\boldsymbol{\nabla}\cdot(\boldsymbol{w} - \mathcal{I}_h\boldsymbol{w})\|_K^2\right) \leqslant C\left(h^{2\theta}|\boldsymbol{w}|^2_{H^\theta(\mathcal{T}_h)} + h^2\|\boldsymbol{\nabla}\cdot\boldsymbol{w}\|_\Omega\right) \tag{A.2b}$$

for all $v \in H^1_{\Gamma_{\rm D}}(\Omega)$ and $\boldsymbol{w} \in \boldsymbol{X}_{\Gamma_{\rm N}}(\mathrm{div},\Omega)$.

Theorem A.1 (Approximation factor)**.** *The following estimate holds true:*

$$\gamma_{\rm ba} \leqslant C\gamma_{\rm st}\left(\frac{\omega\ell}{\vartheta_{\min}}\right)^{1-\theta}\left(\frac{\omega h}{\vartheta_{\min}}\right)^\theta, \tag{A.3}$$

where $\vartheta_{\min} := \sqrt{\alpha_{\min}/\mu_{\max}}$ is the minimal wavespeed.

Proof. Fix $\psi \in L^2(\Omega)$. On the one hand, we have

$$\begin{aligned}
\|u^\star_\psi - \mathcal{I}_h u^\star_\psi\|^2_{\dagger,1,\mathcal{T}_h} &= \sum_{K\in\mathcal{T}_h}\left(\frac{\alpha_K}{h_K^2}\|u^\star_\psi - \mathcal{I}_h u^\star_\psi\|_K^2 + \|\boldsymbol{\nabla}(u^\star_\psi - \mathcal{I}_h u^\star_\psi)\|^2_{\boldsymbol{A},K}\right)\\
&\leqslant \alpha_{\max}\sum_{K\in\mathcal{T}_h}\left(h_K^{-2}\|u^\star_\psi - \mathcal{I}_h u^\star_\psi\|_K^2 + \|\boldsymbol{\nabla}(u^\star_\psi - \mathcal{I}_h u^\star_\psi)\|_K^2\right)\\
&\leqslant C\alpha_{\max}h^{2\theta}|u^\star_\psi|^2_{H^{1+\theta}(\mathcal{T}_h)} \leqslant C\frac{(\ell^{1-\theta}h^\theta)^2}{\alpha_{\min}}\|\boldsymbol{\nabla}\cdot(\boldsymbol{A}\boldsymbol{\nabla} u^\star_\psi)\|^2_\Omega
\end{aligned}$$

due to (A.2a) and (A.1), where $\alpha_{\min}$ is the essential infimum of the smallest eigenvalue of $\boldsymbol{A}$. Similarly, if we let $\boldsymbol{\sigma}^\star_\psi := \boldsymbol{A}\boldsymbol{\nabla} u^\star_\psi$, we have

$$\begin{aligned}
\|\boldsymbol{\sigma}^\star_\psi - \mathcal{J}_h\boldsymbol{\sigma}^\star_\psi\|^2_{\dagger,\mathrm{div},\mathcal{T}_h} &= \sum_{K\in\mathcal{T}_h}\left(\|\boldsymbol{\sigma}^\star_\psi - \mathcal{J}_h\boldsymbol{\sigma}^\star_\psi\|^2_{\boldsymbol{A}^{-1},K} + \frac{h_K^2}{\alpha_K}\|\boldsymbol{\nabla}\cdot(\boldsymbol{\sigma}^\star_\psi - \mathcal{J}_h\boldsymbol{\sigma}^\star_\psi)\|_K^2\right)\\
&\leqslant \frac{1}{\alpha_{\min}}\sum_{K\in\mathcal{T}_h}\left(\|\boldsymbol{\sigma}^\star_\psi - \mathcal{J}_h\boldsymbol{\sigma}^\star_\psi\|_K^2 + h_K^2\|\boldsymbol{\nabla}\cdot(\boldsymbol{\sigma}^\star_\psi - \mathcal{J}_h\boldsymbol{\sigma}^\star_\psi)\|_K^2\right)\\
&\leqslant C\frac{1}{\alpha_{\min}}\left(h^{2\theta}|\boldsymbol{A}\boldsymbol{\nabla} u^\star_\psi|^2_{H^s(\mathcal{T}_h)} + h^2\|\boldsymbol{\nabla}\cdot(\boldsymbol{A}\boldsymbol{\nabla} u^\star_\psi)|^2_\Omega\right)
\end{aligned}$$

where we employed (A.2b) and (A.1). By definition of $u^\star_\psi$, we have

$$\|\boldsymbol{\nabla}\cdot(\boldsymbol{A}\boldsymbol{\nabla} u^\star_\psi)\|^2_\Omega = \|\omega\mu\psi + \omega^2\mu u^\star_\psi\|^2_\Omega \leqslant C\mu_{\max}\omega^2\left(\|\psi\|^2_{\mu,\Omega} + \omega^2\|u^\star_\psi\|^2_{\mu,\Omega}\right) \leqslant C\gamma_{\rm st}^2\mu_{\max}\omega^2\|\psi\|^2_{\mu,\Omega},$$

with $\mu_{\max}$ the essential supremum of μ. This leads to

$$\|u^\star_\psi - \mathcal{I}_h u^\star_\psi\|_{\dagger,1,\mathcal{T}_h} \leqslant C\gamma_{\rm st}\left(\frac{\omega\ell}{\vartheta_{\min}}\right)^{1-\theta}\left(\frac{\omega h}{\vartheta_{\min}}\right)^\theta\|\psi\|_{\mu,\Omega}$$

and

$$\|\boldsymbol{\sigma}_\psi^\star - \mathcal{J}_h \boldsymbol{\sigma}_\psi^\star\|_{\dagger,\mathrm{div},\mathcal{T}_h} \leqslant C\gamma_{\mathrm{st}} \left(\frac{\omega\ell}{\vartheta_{\min}}\right)^{1-\theta} \left(\frac{\omega h}{\vartheta_{\min}}\right)^{\theta} \|\psi\|_{\mu,\Omega}.$$

Since they are valid for all $\psi \in L^2(\Omega)$, these two estimates respectively enable to control $\gamma_{\mathrm{ba,g}}$ and $\gamma_{\mathrm{ba,d}}$, leading to the bound on γ_{ba} in (A.3). □

Remark A.2 (Frequency scaling)**.** In the setting considered here, it can be shown that $\gamma_{\mathrm{st}} \sim \omega/\delta$, where δ is the distance between ω and the closest resonant frequency (the square-root of an eigenvalue). In this case, under the assumption of full elliptic regularity, we obtain

$$\gamma_{\mathrm{st}} \leqslant C\frac{\omega}{\delta}\frac{\omega h}{\vartheta_{\min}}.$$

This is consistent with previously obtained estimates, see *e.g.* [13].

Remark A.3 (Generalization to hanging nodes)**.** As can be seen by the proof of Theorem A.1, the approach proposed in this work does not need matching meshes, and allows for hanging nodes as long as (quasi-)interpolation operators satisfying (A.2) are available. For the full elliptic regularity case where we can take $\theta = 1$ in (A.1), such interpolation operators are available for some families of meshes with hanging nodes, see *e.g.* Lemma 12 from [3].

Remark A.4 (Comparison with earlier estimates)**.** It is instructive to compare our results under minimal regularity assumptions to the more standard approach that requires more regularity. In this case, an approximation factor similar to $\gamma_{\mathrm{ba,g}}$ is defined, but with an enhanced norm including the normal trace of the gradient. This can be seen, *e.g.*, in respectively Proposition 3.5 from [37] and Definition 2.7 from [40] for *a priori* and *a posteriori* error analysis. If $p = 1$ the estimate provided in Theorem 4.8 from [37] for such an approximation factor has the same scaling as the one derived here (notice that the result in [37] is obtained under the assumption that $\gamma_{\mathrm{st}} \sim k^{1+\vartheta}$ for some $\vartheta \geqslant 0$ and that we can take $\theta = 1$). The results obtained in [36] and [20] for conforming discretizations are also similar. We could also obtain sharp estimates for high-order polynomials by using regularity splitting resulting [11, 36]. We refrain from doing so for the sake of brevity though.

References

[1] R. Adams and J. Fournier, Sobolev Spaces. Academic Press (2003).

[2] C. Agut and J. Diaz, Stability analysis of the interior penalty discontinuous Galerkin method for the wave equation. *ESAIM Math. Model. Numer. Anal.* **47** (2013) 903–932.

[3] M. Ainsworth and K. Pinchedez, *hp*-approximation theory for BDFM and RT finite elements on quadrilaterals. *SIAM J. Numer. Anal.* **40** (2002) 2047–2068.

[4] D.N. Arnold, F. Brezzi, B. Cockburn and L.D. Marini, Unified analysis of discontinuous Galerkin, methods for elliptic problems. *SIAM J. Numer. Anal.* **39** (2002) 1749–1779.

[5] H. Barucq, T. Chaumont-Frelet and C. Gout, Stability analysis of heterogeneous Helmholtz problems and finite element solution based on propagation media approximation. *Math. Comput.* **86** (2017) 2129–2157.

[6] M. Bernkopf, S. Sauter, C. Torres and A. Veit, Solvability of discrete Helmholtz equations. Preprint: arXiv:2105.02273v2 (2022).

[7] A. Bonito and R.H. Nochetto, Quasi-optimal convergence rate of an adaptive discontinuous Galerkin method. *SIAM J. Numer. Anal.* **48** (2010) 734–771.

[8] J. Chand, Z. Wang, A. Modave, J.F. Remacle and T. Warburton, GPU-accelerated discontinuous Galerkin methods on hybrid meshes. *J. Comput. Phys.* **318** (2016) 142–168.

[9] S.N. Chandler-Wilde, E.A. Spence, A. Gibbs and V.P. Smyshlyaev, High-frequency bounds for the Helmholtz equation under parabolic trapping and applications in numerical analysis. *SIAM J. Math. Anal.* **52** (2020) 845–893.

[10] T. Chaumont-Frelet, Mixed finite element discretizations of acoustic Helmholtz problems with high wavenumbers. *Calcolo* **56** (2019).

[11] T. Chaumont-Frelet and S. Nicaise, Wavenumber explicit convergence analysis for finite element discretizations of general wave propagation problems. *IMA J. Numer. Anal.* **40** (2020) 1503–1543.

[12] T. Chaumont-Frelet and E.A. Spence, Scattering by finely-layered obstacles: frequency-explicit bounds and homogenization. *SIAM J. Math. Anal.* **55** (2023) 1319–1363.

[13] T. Chaumont-Frelet and P. Vega, Frequency-explicit approximability estimates for time-harmonic Maxwell's equations. *Calcolo* **59** (2022) 22.

[14] T. Chaumont-Frelet, A. Ern and M. Vohralík, On the derivation of guaranteed and p-robust a posteriori error estimates for the Helmholtz equation. *Numer. Math.* **148** (2021) 525–573.

[15] P.G. Ciarlet, The finite element method for elliptic problems. *SIAM* (2002).

[16] G. Cohen, P. Joly, J.E. Roberts and N. Tordjman, High order triangular finite element with mass lumping for the wave equation. *SIAM J. Numer. Anal.* **38** (2001) 2047–2078.

[17] S. Congreve, J. Gedicke and I. Perugia, Robust adaptive *hp* discontinuous Galerkin finite element methods for the Helmholtz equation. *SIAM J. Sci. Comput.* **41** (2019) A1121–A1147.

[18] M. Costabel, M. Dauge and S. Nicaise, Singularities of Maxwell interface problems. *ESAIM Math. Model. Numer. Anal.* **33** (1999) 627–649.

[19] D.A. Di Pietro and A. Ern, Mathematical Aspects of Discontinuous Galerkin Methods. Springer (2012).

[20] W. Dörfler and S. Sauter, A posteriori error estimation for highly indefinite Helmholtz problems. *Comput. Methods Appl. Math.* **13** (2013) 333–347.

[21] Y. Du and L. Zhu, Preasymptotic error analysis of high order interior penalty discontinuous Galerkin methods for the Helmholtz equation with high wave number. *J. Sci. Comput.* **67** (2016) 130–152.

[22] A. Ern and J.L. Guermond, Finite element quasi-interpolation and best approximation. *ESAIM Math. Model. Numer. Anal.* **51** (2017) 1367–1385.

[23] A. Ern and M. Vohralík, Polynomial-degree-robust a posteriori estimates in a unified setting for conforming, nonconforming, discontinuous Galerkin, and mixed discretizations. *SIAM J. Numer. Anal.* **53** (2015) 1058–1081.

[24] A. Ern and J.L. Guermond, Finite Elements I: Basic Theory and Practice. Springer Nature (2021).

[25] A. Ern, T. Gudi, I. Smears and M. Vohralík, Equivalence of local- and global-best approximations, a simple stable local commuting projector, and optimal *hp* approximation estimates in $\boldsymbol{H}(\mathrm{div})$. *IMA J. Numer. Anal.* **42** (2022) 1023–1049.

[26] X. Feng and H. Wu, Discontinuous Galerkin methods for the Helmholtz equation with large wave number. *SIAM J. Numer. Anal.* **47** (2009) 2872–2896.

[27] X. Feng and H. Wu, *hp*-discontinuous Galerkin methods for the Helmholtz equation with large wave number. *Math. Comput.* **80** (2011) 1997–2024.

[28] P. Fernandes and G. Gilardi, Magnetostatic and electrostatic problems in inhomogeneous anisotropic media with irregular boundary and mixed boundary conditions. *Math. Methods Appl. Sci.* **47** (1997) 2872–2896.

[29] V. Girault and P.A. Raviart, Finite Element Methods for Navier-Stokes Equations: Theory and Algorithms. Springer-Verlag (1986).

[30] P. Grisvard, Elliptic Problems in Nonsmooth Domains. Pitman (1985).

[31] M.J. Grote, A. Schneebeli and D. Schötzau, Discontinuous Galerkin finite element method for the wave equation. *SIAM J. Numer. Anal.* **44** (2006) 2408–2431.

[32] R.H.W. Hope and N. Sharma, Convergence analysis of an adaptive interior penalty discontinuous Galerkin method for the Helmholtz equation. *IMA J. Numer. Anal.* **33** (2013) 898–921.

[33] P. Houston, D. Schötzau and T.P. Wihler, Energy norm a posteriori error estimation of *hp*-adaptive discontinuous Galerkin methods for elliptic problems. *Math. Models Methods Appl. Sci.* **17** (2006) 33–62.

[34] F. Jochmann, An h^s-regularity result for the gradient of solutions to elliptic equations with mixed boundary conditions.
J. Math. Anal. Appl. **238** (1999) 459–450.

[35] D. Lafontaine, E.A. Spence and J. Wunsch, For most frequencies, strong tapping has a weak effect in frequency-domain scattering. *Commun. Pure Appl. Math.* **74** (2022) 2025–2063.

[36] J.M. Melenk and S. Sauter, Convergence analysis for finite element discretizations of the Helmholtz equation with Dirichlet-to-Neumann boundary conditions. *Math. Comput.* **79** (2010) 1871–1914.

[37] J.M. Melenk, A. Parsania and S. Sauter, General DG-methods for highly indefinite Helmholtz problems. *J. Sci. Comput.* **57** (2013) 536–581.

[38] A. Moiola and E.A Spence, Acoustic transmission problems: wavenumber-explicit bounds and resonance-free regions. *Math. Methods Appl. Sci.* **29** (2019) 317–354.

[39] P. Monk, Finite Element Methods for Maxwell's Equations. Oxford Science Publications (2003).

[40] S. Sauter and J. Zech, A posteriori error estimation of $hp-dg$ finite element methods for highly indefinite Helmholtz problems. *SIAM J. Numer. Anal.* **53** (2015) 2414–2440.

[41] A.H. Schatz, An observation concerning Ritz-Galerkin methods with indefinite bilinear forms. *Math. Comput.* **28** (1974) 959–962.