# Mixed finite element discretizations of acoustic Helmholtz problems with high wavenumbers

Théophile Chaumont-Frelet

## ▶ To cite this version:

Théophile Chaumont-Frelet. Mixed finite element discretizations of acoustic Helmholtz problems with high wavenumbers. Calcolo, Springer Verlag, In press, 10.1007/s10092-019-0346-z . hal-02197891

HAL Id: hal-02197891

https://hal.inria.fr/hal-02197891

Submitted on 30 Jul 2019

# MIXED FINITE ELEMENT DISCRETIZATIONS OF ACOUSTIC HELMHOLTZ PROBLEMS WITH HIGH WAVENUMBERS

T. CHAUMONT-FRELET

UNIVERSITÉ CÔTE D'AZUR, INRIA, CNRS, LABORATOIRE J.A. DIEUDONNÉ

ABSTRACT. We study the acoustic Helmholtz equation with impedance boundary conditions formulated in terms of velocity, and analyze the stability and convergence properties of lowest-order Raviart-Thomas finite element discretizations. We focus on the high-wavenumber regime, where such discretizations suffer from the so-called "pollution effect", and lack stability unless the mesh is sufficiently refined. We provide wavenumber-explicit mesh refinement conditions to ensure the well-posedness and stability of discrete scheme, as well as wavenumber-explicit error estimates. Our key result is that the condition "$k^2h$ is sufficiently small", where $k$ and $h$ respectively denote the wavenumber and the mesh size, is sufficient to ensure the stability of the scheme. We also present numerical experiments that illustrate the theory and show that the derived stability condition is actually necessary.

KEY WORDS. Helmholtz problems; Mixed finite elements; Pollution effect.

## 1. INTRODUCTION

We consider a convex polytopal domain $\Omega \subset \mathbb{R}^d$, with $d = 2$ or $3$. $\partial\Omega$ is the boundary of $\Omega$ and $\boldsymbol{n}$ is the unit vector normal to $\partial\Omega$ pointing outward $\Omega$. Given a wavenumber $k > 0$ and a load term $\boldsymbol{f} : \Omega \to \mathbb{C}^3$, we consider the problem to find $\boldsymbol{u} : \Omega \to \mathbb{C}^3$ such that

$$(1.1) \qquad \left\{ \begin{array}{rcll} -k^2\boldsymbol{u} - \boldsymbol{\nabla}\boldsymbol{\nabla} \cdot \boldsymbol{u} & = & \boldsymbol{f} & \text{in } \Omega, \\ \boldsymbol{\nabla} \cdot \boldsymbol{u} - ik\boldsymbol{u} \cdot \boldsymbol{n} & = & 0 & \text{on } \partial\Omega. \end{array} \right.$$

Here, $\Omega$ represents a volume occupied by a fluid, and boundary value problem (1.1) modelizes the propagation of (small) acoustic waves in the fluid (see [7] for instance). Specifically, (1.1) is usually called the velocity formulation of the acoustic Helmholtz equation [11, 12], since $\boldsymbol{u}$ represents the velocity of the fluid particles. The "impedance" (also known as "absorbing" in this setting) boundary condition is an approximation of the Sommerfeld radiation condition and prevent the propagation of incoming waves from the exterior to the interior of the computational domain [12, 13].

When the right-hand side is irrotational and vanishes close to $\partial\Omega$, we have $\boldsymbol{f} = \boldsymbol{\nabla}f$ for some $f : \Omega \to \mathbb{C}$, and the pressure formulation, that consists in finding $p : \Omega \to \mathbb{C}$ such that

$$\left\{ \begin{array}{rcll} -k^2p - \Delta p & = & f & \text{in } \Omega, \\ \boldsymbol{\nabla}p \cdot \boldsymbol{n} - ikp & = & 0 & \text{on } \partial\Omega, \end{array} \right.$$

can be employed instead of (1.1). In this case, $p$ represents the fluid pressure, and we have $\boldsymbol{u} = \boldsymbol{\nabla}p$.

The pressure formulation is usually preferred to the velocity formulation, since it only involves a scalar unknown, that is easier to discretize. However, such formulation is not always available. This is for instance the case if the load-term exhibits a non-vanishing rotation, or if more elaborated physical phenomena are modelized into the boundary value problem [11, 16, 26, 31]. Another attractive feature of the velocity formulation is that it generally delivers a divergence-conforming discrete field $\boldsymbol{u}_h$, without resorting to any post-processing [23, 25].

Although alternative technologies are available [8, 29], we focus on finite element approximations in this work [6, 22]. The discretization of velocity formulation (1.1) requires the use of Raviart-Thomas finite elements [23, 25], while the pressure formulation is usually approximated using Lagrange finite elements (see [6, 22] for instance). Although approximations of the pressure formulation have been thoroughly analyzed in the literature in the high-wavenumber regime

[2, 3, 4, 20, 21], much less is known for the discretization of the velocity formulation. For the pressure formulation, it is now well-established that finite element schemes are affected by the "pollution effect": the finite element solution is not quasi-optimal, unless the mesh is sufficiently refined. This effect is more important when the wavenumber is large, leading to strong mesh size restrictions in the high-wavenumber regime. Specifically [20, 21], the quantity $k^2h$ must be sufficiently small to ensure the stability of the discrete scheme.

The purpose of the present work is to extend the stability condition "$k^2h$ is sufficiently small" to the velocity formulation discretized with Raviart-Thomas elements. Surprisingly, consequent complications arise compared to the pressure formulation. On the one hand, the analysis of the continuous problem is tedious, partly because mixed finite elements require a sharp regularity analysis, but mostly because the impedance boundary condition in (1.1), is not "natural" for the Sobolev space $\boldsymbol{H}(\mathrm{div}, \Omega)$ (we also refer to [22, Chapter 3.8] for a similar situation in $\boldsymbol{H}(\mathrm{curl}, \Omega)$). On the other hand, the stability proof employs a duality technique known as the "Aubin-Nitsche" trick [20, 21, 27]. This technique is well established for Lagrange finite elements, and while its application to mixed finite elements is possible, it presents some difficulties [5, 15, 32].

Here, we carefully establish that the condition "$k^2h$ is small" is sufficient to ensure the well-posedness and stability of the discrete scheme. Furthermore, we present numerical experiments showing that this condition is actually necessary, thereby indicating that our stability condition is sharp.

We also provide a comparison of the pressure and velocity formulations on a two-dimensional scattering benchmark. This comparison suggests that the velocity formulation is a competitive alternative to the pressure formulation in terms of computational costs, with the advantage of producing a divergence-conforming velocity field.

Our work is organized as follows: in Section 2 we introduce the notation we employ. Section 3 is devoted to the analysis of (1.1) at the continuous level. We consider finite element discretizations in Section 4 where we established our key stability and convergence results. Numerical experiments are presented in Section 5, and we draw our conclusions in Section 6.

## 2. Notation

As we are especially interested in the discretization of (1.1) when the wavenumber is high, we assume that $k \geq k_0 > 0$ for given fixed minimal value $k_0$. As detailed hereafter, the constants in the estimates are allowed to depend on $k_0$. It means that our results are meaningful for the high wavenumber cases, but that the constants can blow up in the low wavenumber limit as $k$ tends to 0.

In Section 4, we will introduce a small parameter $h > 0$ as well as a number $\beta \in (0, 1)$ that describe the size and the regularity of the finite element meshes (see (4.1) for more details). Then, we employ the notation $A \lesssim B$ if $A, B \in \mathbb{R}$ are two numbers such that there exists a constant $C > 0$ that is independent of $A, B, k, h$ such that $A \leq CB$. We mention that the constant $C$ is allowed to depend on $\Omega$, $k_0$ and $\beta$. We also note $A \gtrsim B$ if $B \lesssim A$.

We denote by $\{\mathcal{F}_j\}_{j=1}^{N_\mathcal{F}}$ the set of faces of $\partial\Omega$. For each $\mathcal{F}_j$ we employ the notation $\mathcal{H}_j$ to denote the hyperplane containing $\mathcal{F}_j$. We note that $\mathcal{H}_j$ can clearly be identified with $\mathbb{R}^{d-1}$.

We close this section with the notation we employ for functional spaces, we shall only give a brief description and refer the reader to [1, 17, 30] for precise definitions.

For $m \in \mathbb{N}$ and $b = d - 1$ or $d$, we denote by $C^m(\mathbb{R}^b)$, the space of $u : \mathbb{R}^b \to \mathbb{C}$ that admits continuous classical derivatives up to order $m$. If $\mathcal{U} \subset \mathbb{R}^b$, $C^m(\overline{\mathcal{U}})$ denotes the restriction to $\mathcal{U}$ of functions in $C^m(\mathbb{R}^b)$. $C^m(\overline{\mathcal{U}})$ is equipped with its usual norm $\|\cdot\|_{m,\infty,\mathcal{U}}$. In addition we employ that notation $C^\infty(\overline{\mathcal{U}}) = \bigcap_{m \in \mathbb{N}} C^m(\overline{\mathcal{U}})$. $\mathcal{D}(\mathcal{U})$ is the space of function $C^\infty(\overline{\mathcal{U}})$ that have compact support in $\mathcal{U}$. It is equipped with its usual topology. The space of distributions $\mathcal{D}'(\mathcal{U})$ is the topological dual of $\mathcal{D}(\mathcal{U})$. We denote by $\langle \cdot, \cdot \rangle_\mathcal{U}$ the duality pairing between $\mathcal{D}'(\mathcal{U})$ and $\mathcal{D}(\mathcal{U})$.

We further denote by $L^2(\mathcal{U})$ the space of (equivalence classes of) functions $u : \mathcal{U} \to \mathbb{C}$ that are square integrable with respect to the Lebesgue measure of $\mathbb{R}^b$. We denote by $\|\cdot\|_{0,\mathcal{U}}$ and $(\cdot, \cdot)_\mathcal{U}$ the usual norm and inner products of $L^2(\mathcal{U})$. Since $\Omega$ is a Lipschitz domain, we can also define the space $L^2(\partial\Omega)$ of functions $u : \partial\Omega \to \mathbb{C}$ that are square-integrable for the surface measure of $\Omega$. For

$s > 0$, we denote by $H^s(\mathcal{U})$ the standard Sobolev space of fractional order $s$. It is equipped with its standard norm and semi-norm $\|\cdot\|_{s,\mathcal{U}}$ and $|\cdot|_{s,\mathcal{U}}$. Since $\Omega$ has a Lipschitz boundary (see [17, Definition 3.4] with $k = 0$), we can also define the space $H^s(\partial\Omega)$ for $0 < s < 1$. We also point out that since each face $\mathcal{F}_j$ can be seen as a subset of $\mathbb{R}^{d-1}$, the Sobolev spaces $H^s(\mathcal{F}_j)$ can be defined for all $s > 0$. A similar remark holds for the planes $\mathcal{H}_j$.

We further point out (see [30, Chapter 36] and [17, Theorem 1.5]) that since $\Omega$ has a Lipschitz boundary, for $0 < s < 1$, the space $H^{m+s}(\Omega)$ is equivalently defined by real interpolation between $H^m(\Omega)$ and $H^{m+1}(\Omega)$. As a result, for all $v \in H^{m+s}(\Omega)$, we have

$$\tag{2.1} \|v\|^2_{m+s,\Omega} \lesssim \|v\|^{1-s}_{m,\Omega}\|v\|^s_{m+1,\Omega}.$$

If $G(\Omega)$ is any of the aforementioned spaces, $\boldsymbol{G}(\Omega)$ denotes the space of vector-valued (with values in $\mathbb{C}^d$) functions that are component-wide $G(\Omega)$. We use the same notations for the norm and inner product of $G(\Omega)$ and $\boldsymbol{G}(\Omega)$.

We further require Sobolev spaces of vectorial functions that are related with the divergence and rotation operators, as discussed in [17]. If $\boldsymbol{v} \in \boldsymbol{L}^2(\Omega)$ we denote by $\boldsymbol{\nabla} \cdot \boldsymbol{f}$ and $\boldsymbol{\nabla} \times \boldsymbol{f}$ its divergence and rotational defined in the distributional sense. We employ the notations $\boldsymbol{H}(\mathrm{div}, \Omega)$ and $\boldsymbol{H}(\mathrm{curl}, \Omega)$ for the usual Sobolev space of $\boldsymbol{L}^2(\Omega)$ with square integrable divergence and rotational. Finally, $\boldsymbol{H}_0(\mathrm{div}, \Omega)$ and $\boldsymbol{H}_0(\mathrm{curl}, \Omega)$ denote the closure of $\boldsymbol{\mathcal{D}}(\Omega)$ in $\boldsymbol{H}(\mathrm{div}, \Omega)$ and $\boldsymbol{H}(\mathrm{curl}, \Omega)$.

## 3. Analysis of the continuous problem

We first set a convenient functional setting to recast (1.1) in a variational form. If $\boldsymbol{u} \in \boldsymbol{C}^2(\overline{\Omega})$ is a classical solution to (1.1), integration by parts shows that for any $\boldsymbol{v} \in \boldsymbol{C}^1(\overline{\Omega})$, we have

$$\tag{3.1} -k^2(\boldsymbol{u}, \boldsymbol{v})_\Omega - ik(\boldsymbol{u} \cdot \boldsymbol{n}, \boldsymbol{v} \cdot \boldsymbol{n})_{\partial\Omega} + (\boldsymbol{\nabla} \cdot \boldsymbol{u}, \boldsymbol{\nabla} \cdot \boldsymbol{v})_\Omega = (\boldsymbol{f}, \boldsymbol{v})_\Omega.$$

Reciprocally, standard arguments based on integration by parts show that if $\boldsymbol{u} \in \boldsymbol{C}^2(\Omega)$ satisfies (3.1) for all $\boldsymbol{v} \in \boldsymbol{C}^1(\Omega)$, then $\boldsymbol{u}$ is a classical solution to problem (1.1).

The boundary integral in (3.1) motivates the introduction of a dedicated functional space as the normal traces of functions in $\boldsymbol{H}(\mathrm{div}, \Omega)$ are not sufficiently regular. A similar situation for the $\boldsymbol{H}(\mathrm{curl}, \Omega)$ is described in [22, Chapter 3.8], and we follow the same approach. We thus introduce the norm

$$\|\boldsymbol{v}\|^2_{\boldsymbol{\mathcal{X}}} = k^2\|\boldsymbol{v}\|^2_{0,\Omega} + k\|\boldsymbol{v} \cdot \boldsymbol{n}\|^2_{0,\partial\Omega} + \|\boldsymbol{\nabla} \cdot \boldsymbol{v}\|^2_{0,\Omega},$$

for all $\boldsymbol{v} \in \boldsymbol{C}^1(\overline{\Omega})$, and define the Hilbert space $\boldsymbol{\mathcal{X}}$ as the closure of $\boldsymbol{C}^1(\overline{\Omega})$ in $\boldsymbol{L}^2(\Omega)$ with respect to the $\|\cdot\|_{\boldsymbol{\mathcal{X}}}$-norm. We use the notation $\boldsymbol{\mathcal{X}}'$ to denote the topological dual of $\boldsymbol{\mathcal{X}}$. Classically, by Riesz representation theorem, we identify $\boldsymbol{L}^2(\Omega)$ with its dual, and obtain the "Gelfand triplet" $\boldsymbol{\mathcal{X}} \subset \boldsymbol{L}^2(\Omega) \subset \boldsymbol{\mathcal{X}}'$. We refer the reader to [22, Chapter 2] for more informations.

We are now able to define the sesquilinear form

$$b(\boldsymbol{u}, \boldsymbol{v}) = -k^2(\boldsymbol{u}, \boldsymbol{v})_\Omega - ik(\boldsymbol{u} \cdot \boldsymbol{n}, \boldsymbol{v} \cdot \boldsymbol{n})_{\partial\Omega} + (\boldsymbol{\nabla} \cdot \boldsymbol{u}, \boldsymbol{\nabla} \cdot \boldsymbol{v})_\Omega, \quad \forall \boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{\mathcal{X}}.$$

Elementary computations show that $b$ is continuous with

$$\tag{3.2} |b(\boldsymbol{u}, \boldsymbol{v})| \lesssim \|\boldsymbol{u}\|_{\boldsymbol{\mathcal{X}}}\|\boldsymbol{v}\|_{\boldsymbol{\mathcal{X}}}.$$

We first focus on proving an inf-sup condition for the sesquilinear form $b$, that we establish in Theorem 3.1. As a direct consequence of this inf-sup condition, we obtain the existence and uniqueness of a weak solution to (3.1) in the space $\boldsymbol{\mathcal{X}}$ for every (generalized) right-hand side in $\boldsymbol{\mathcal{X}}'$. Such solution is also a generalized solution to (1.1). We start by establishing a useful result in Lemma 3.1.

**Lemma 3.1.** *For every $\boldsymbol{u} \in \boldsymbol{\mathcal{X}}$, there exists a unique element $p \in H^1(\Omega)$ such that*

$$\tag{3.3} \begin{cases} -k^2p - \Delta p & = & -\boldsymbol{\nabla} \cdot \boldsymbol{u} & in\ \Omega, \\ \boldsymbol{\nabla}p \cdot \boldsymbol{n} + ikp & = & 0 & on\ \partial\Omega. \end{cases}$$

*Furthermore, we have $\boldsymbol{\nabla}p \in \boldsymbol{\mathcal{X}}$ with*

$$\tag{3.4} \|\boldsymbol{\nabla}p\|_{\boldsymbol{\mathcal{X}}} \lesssim k\|\boldsymbol{\nabla} \cdot \boldsymbol{u}\|_{0,\Omega},$$

*and it holds that*

$$(3.5) \qquad b(\boldsymbol{u}, \boldsymbol{\nabla} p) = \|\boldsymbol{\nabla} \cdot \boldsymbol{u}\|_{0,\Omega}^2.$$

*Proof.* Since $-\boldsymbol{\nabla} \cdot \boldsymbol{u} \in L^2(\Omega)$ the existence and uniqueness of $p \in H^1(\Omega)$ follows from the standard theory of the Helmholtz equation [19]. In addition, as $\Omega$ is convex, Proposition 3.3 of [19] shows that

$$(3.6) \qquad k\|p\|_{0,\Omega} + |p|_{1,\Omega} \lesssim \|\boldsymbol{\nabla} \cdot \boldsymbol{u}\|_{0,\Omega}.$$

In addition, since $\Delta p = \boldsymbol{\nabla} \cdot \boldsymbol{u} - k^2 p$, we also have $\boldsymbol{\nabla} p \in \boldsymbol{H}(\mathrm{div}, \Omega)$ with

$$(3.7) \qquad \|\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} p)\|_{1,\Omega} \lesssim k\|\boldsymbol{\nabla} \cdot \boldsymbol{u}\|_{0,\Omega}.$$

Finally, since $\boldsymbol{\nabla} p \cdot \boldsymbol{n} = -ikp$, we have $\boldsymbol{\nabla} p \cdot \boldsymbol{n} \in H^{1/2}(\partial\Omega) \subset L^2(\partial\Omega)$ and $\boldsymbol{\nabla} p \in \boldsymbol{\mathcal{X}}$. Furthermore, applying trace Theorem 1.5.1.10 from [18], we have

$$\|p\|_{0,\partial\Omega}^2 \lesssim k\|p\|_{0,\Omega}^2 + k^{-1}|p|_{1,\Omega}^2 \lesssim k^{-1}\|\boldsymbol{\nabla} \cdot \boldsymbol{u}\|_{0,\Omega}^2$$

and therefore

$$(3.8) \qquad k^{1/2}\|\boldsymbol{\nabla} p \cdot \boldsymbol{n}\|_{0,\partial\Omega} \lesssim k^{3/2}\|p\|_{0,\partial\Omega} \lesssim k^{1/2}\|\boldsymbol{\nabla} \cdot \boldsymbol{u}\|_{0,\Omega} \lesssim k\|\boldsymbol{\nabla} \cdot \boldsymbol{u}\|_{0,\Omega}.$$

Then, (3.4) follows from (3.6), (3.7) and (3.8).

On the other hand, since $\boldsymbol{\nabla} p \in \boldsymbol{\mathcal{X}}$, we can prove (3.5) by integration by parts. Indeed, recalling that $p$ is solution to (3.3), we have

$$\begin{aligned}
b(\boldsymbol{u}, \boldsymbol{\nabla} p) &= -k^2(\boldsymbol{u}, \boldsymbol{\nabla} p)_\Omega - ik(\boldsymbol{u} \cdot \boldsymbol{n}, \boldsymbol{\nabla} p \cdot \boldsymbol{n})_{\partial\Omega} + (\boldsymbol{\nabla} \cdot \boldsymbol{u}, \Delta p)_\Omega \\
&= -k^2(\boldsymbol{u} \cdot \boldsymbol{n}, p)_{\partial\Omega} + k^2(\boldsymbol{\nabla} \cdot \boldsymbol{u}, p) - ik(\boldsymbol{u} \cdot \boldsymbol{n}, \boldsymbol{\nabla} p \cdot \boldsymbol{n})_{\partial\Omega} + (\boldsymbol{\nabla} \cdot \boldsymbol{u}, \Delta p)_\Omega \\
&= (\boldsymbol{\nabla} \cdot \boldsymbol{u}, k^2 p + \Delta p)_\Omega - ik(\boldsymbol{u} \cdot \boldsymbol{n}, \boldsymbol{\nabla} p \cdot \boldsymbol{n} + ikp)_{\partial\Omega} \\
&= \|\boldsymbol{\nabla} \cdot \boldsymbol{u}\|_{0,\Omega}^2.
\end{aligned}$$

$\square$

We are now ready to establish our first main result.

**Theorem 3.1.** *For $\boldsymbol{u} \in \boldsymbol{\mathcal{X}}$, we introduce $\boldsymbol{u}^\star = 2\boldsymbol{\nabla} p - (1+i)\boldsymbol{u} \in \boldsymbol{\mathcal{X}}$, where $\boldsymbol{\nabla} p \in \boldsymbol{\mathcal{X}}$ is defined following Lemma 3.1. Then, we have*

$$(3.9) \qquad \mathrm{Re}\, b(\boldsymbol{u}, \boldsymbol{u}^\star) = \|\boldsymbol{u}\|_{\boldsymbol{\mathcal{X}}}^2,$$

*and*

$$(3.10) \qquad \|\boldsymbol{u}^\star\|_{\boldsymbol{\mathcal{X}}} \lesssim k\|\boldsymbol{u}\|_{\boldsymbol{\mathcal{X}}}.$$

*In particular, it follows that*

$$(3.11) \qquad \inf_{\boldsymbol{u} \in \boldsymbol{\mathcal{X}} \setminus \{\boldsymbol{0}\}} \sup_{\boldsymbol{v} \in \boldsymbol{\mathcal{X}} \setminus \{\boldsymbol{0}\}} \frac{\mathrm{Re}\, b(\boldsymbol{u}, \boldsymbol{v})}{\|\boldsymbol{u}\|_{\boldsymbol{\mathcal{X}}} \|\boldsymbol{v}\|_{\boldsymbol{\mathcal{X}}}} \gtrsim \frac{1}{k}.$$

*Therefore, for every $\boldsymbol{\psi} \in \boldsymbol{\mathcal{X}}'$, there exists a unique element $\boldsymbol{u} \in \boldsymbol{\mathcal{X}}$ such that $b(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{\psi}(\boldsymbol{v})$ for all $\boldsymbol{v} \in \boldsymbol{\mathcal{X}}$, and we have*

$$(3.12) \qquad \|\boldsymbol{u}\|_{\boldsymbol{\mathcal{X}}} \lesssim k\|\boldsymbol{\psi}\|_{\boldsymbol{\mathcal{X}}'}.$$

*Proof.* Let $\boldsymbol{u} \in \boldsymbol{\mathcal{X}}$ and define $\boldsymbol{u}^\star \in \boldsymbol{\mathcal{X}}$ as above. Equality (3.9) follows from simple computations. Indeed, we have

$$\begin{aligned}
\mathrm{Re}\, b(\boldsymbol{u}, \boldsymbol{u}^\star) &= 2\,\mathrm{Re}\, b(\boldsymbol{u}, \boldsymbol{\nabla} p) - \mathrm{Re}\,\{(1-i)b(\boldsymbol{u}, \boldsymbol{u})\} \\
&= 2\,\mathrm{Re}\, b(\boldsymbol{u}, \boldsymbol{\nabla} p) - \mathrm{Re}\, b(\boldsymbol{u}, \boldsymbol{u}) - \mathrm{Im}\, b(\boldsymbol{u}, \boldsymbol{u}) \\
&= 2\|\boldsymbol{\nabla} \cdot \boldsymbol{u}\|_{0,\Omega}^2 + k^2\|\boldsymbol{u}\|_{0,\Omega}^2 - \|\boldsymbol{\nabla} \cdot \boldsymbol{u}\|_{0,\Omega}^2 + k\|\boldsymbol{u} \cdot \boldsymbol{n}\|_{0,\partial\Omega}^2 \\
&= \|\boldsymbol{u}\|_{\boldsymbol{\mathcal{X}}}^2.
\end{aligned}$$

In addition, (3.10) is a direct consequence of (3.4). To prove (3.11), we observe that (3.9) and (3.10) imply that

$$\mathrm{Re}\, b(\boldsymbol{u}, \boldsymbol{u}^\star) \gtrsim \frac{1}{k}\|\boldsymbol{u}\|_{\boldsymbol{\mathcal{X}}}\|\boldsymbol{u}^\star\|_{\boldsymbol{\mathcal{X}}}.$$

Assuming that $\boldsymbol{u} \neq \boldsymbol{0}$, (3.9) ensures that $\boldsymbol{u}^\star \neq 0$, so that for all $\boldsymbol{u} \in \boldsymbol{\mathcal{X}} \setminus \{\boldsymbol{0}\}$, we have

$$\sup_{\boldsymbol{v} \in \boldsymbol{\mathcal{X}} \setminus \{\boldsymbol{0}\}} \frac{\operatorname{Re} b(\boldsymbol{u}, \boldsymbol{u}^\star)}{\|\boldsymbol{u}\|_{\boldsymbol{\mathcal{X}}} \|\boldsymbol{v}\|_{\boldsymbol{\mathcal{X}}}} \geq \frac{\operatorname{Re} b(\boldsymbol{u}, \boldsymbol{u}^\star)}{\|\boldsymbol{u}\|_{\boldsymbol{\mathcal{X}}} \|\boldsymbol{u}^\star\|_{\boldsymbol{\mathcal{X}}}} \gtrsim \frac{1}{k},$$

and we obtain (3.11) by taking the infimum over $\boldsymbol{u} \in \boldsymbol{\mathcal{X}} \setminus \{\boldsymbol{0}\}$.

Let us now assume that $\boldsymbol{v} \in \boldsymbol{\mathcal{X}}$ satisfies $b(\boldsymbol{u}, \boldsymbol{v}) = 0$ for all $\boldsymbol{u} \in \boldsymbol{\mathcal{X}}$. Then, in particular, we have $b(\overline{\boldsymbol{u}}, \boldsymbol{v}) = 0$ for all $\boldsymbol{u} \in \boldsymbol{\mathcal{X}}$. Thus, we have

$$b(\overline{\boldsymbol{v}}, \boldsymbol{u}) = b(\overline{\boldsymbol{u}}, \boldsymbol{v}) = 0, \quad \forall \boldsymbol{u} \in \boldsymbol{\mathcal{X}},$$

and inf-sup condition (3.11) ensures that $\boldsymbol{v} = \overline{\boldsymbol{v}} = \boldsymbol{0}$. Recalling (3.2), $b$ is continuous with a continuity constant that is independent of $k$. As a result, the Babuška-Brezzi Theorem (see [22, Theorem 2.22] for instance) implies that for each $\boldsymbol{\psi} \in \boldsymbol{\mathcal{X}}'$, there exists a unique $\boldsymbol{u} \in \boldsymbol{\mathcal{X}}$ such that $b(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{\psi}(\boldsymbol{v})$ for all $\boldsymbol{v} \in \boldsymbol{\mathcal{X}}$, and that $\boldsymbol{u}$ satisfies (3.12). $\qquad \square$

In the remaining of this section, we refine Theorem 3.1 by assuming more regularity on the right-hand side $\boldsymbol{\psi} \in \boldsymbol{\mathcal{X}}'$. We start in Theorem 3.2 with improved regularity results for $\boldsymbol{u}$ in $\boldsymbol{\nabla} \cdot \boldsymbol{u}$ in $\Omega$.

**Theorem 3.2.** *For all $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$, there exists a unique $\boldsymbol{u} \in \boldsymbol{\mathcal{X}}$ such that $b(\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{f}, \boldsymbol{v})_\Omega$ for all $\boldsymbol{v} \in \boldsymbol{\mathcal{X}}$, and we have*

$$\tag{3.13} \|\boldsymbol{u}\|_{\boldsymbol{\mathcal{X}}} \lesssim \|\boldsymbol{f}\|_{0,\Omega}.$$

*Furthermore, $\boldsymbol{\nabla} \cdot \boldsymbol{u} \in H^1(\Omega)$, and we have*

$$\tag{3.14} |\boldsymbol{\nabla} \cdot \boldsymbol{u}|_{1,\Omega} \lesssim k \|\boldsymbol{f}\|_{0,\Omega}.$$

*In addition, if we assume that $\boldsymbol{f} \in \boldsymbol{H}(\operatorname{curl}, \Omega)$, then $\boldsymbol{u} \in \boldsymbol{H}^1(\Omega)$ with*

$$\tag{3.15} |\boldsymbol{u}|_{1,\Omega} \lesssim \|\boldsymbol{f}\|_{0,\Omega} + k^{-2} \|\boldsymbol{\nabla} \times \boldsymbol{f}\|_{0,\Omega}.$$

*Proof.* We define the anti-linear form $\boldsymbol{\psi} : \boldsymbol{\mathcal{X}} \ni \boldsymbol{v} \to (\boldsymbol{f}, \boldsymbol{v})_\Omega \in \mathbb{C}$. Since we have

$$|\boldsymbol{\psi}(\boldsymbol{v})| = |(\boldsymbol{f}, \boldsymbol{v})| \leq \|\boldsymbol{f}\|_{0,\Omega} \|\boldsymbol{v}\|_{0,\Omega} \leq k^{-1} \|\boldsymbol{f}\|_{0,\Omega} \|\boldsymbol{v}\|_{\boldsymbol{\mathcal{X}}},$$

for every $\boldsymbol{v} \in \boldsymbol{\mathcal{X}}$, $\boldsymbol{\psi} \in \boldsymbol{\mathcal{X}}'$ with $\|\boldsymbol{\psi}\|_{\boldsymbol{\mathcal{X}}'} \leq k^{-1} \|\boldsymbol{f}\|_{0,\Omega}$. Thus, $\boldsymbol{u}$ is uniquely defined as the element of $\boldsymbol{\mathcal{X}}$ satisfying $b(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{\psi}(\boldsymbol{v})$, and (3.13) follows from (3.12).

Consider a test function $\boldsymbol{\phi} \in \boldsymbol{\mathcal{D}}(\Omega)$. We have

$$-\langle \boldsymbol{\nabla} \boldsymbol{\nabla} \cdot \boldsymbol{u}, \boldsymbol{\phi} \rangle_\Omega = (\boldsymbol{\nabla} \cdot \boldsymbol{u}, \boldsymbol{\nabla} \cdot \boldsymbol{\phi})_\Omega = b(\boldsymbol{u}, \boldsymbol{\phi}) + k^2 (\boldsymbol{u}, \boldsymbol{\phi})_\Omega = (\boldsymbol{f} + k^2 \boldsymbol{u}, \boldsymbol{\phi})_\Omega,$$

so that $\boldsymbol{\nabla} \boldsymbol{\nabla} \cdot \boldsymbol{u} = -\boldsymbol{f} - k^2 \boldsymbol{u} \in \boldsymbol{L}^2(\Omega)$. Hence, $\boldsymbol{\nabla} \cdot \boldsymbol{u} \in H^1(\Omega)$ and (3.14) follows from (3.13) since

$$|\boldsymbol{\nabla} \cdot \boldsymbol{u}|_{1,\Omega} = \|\boldsymbol{\nabla} \boldsymbol{\nabla} \cdot \boldsymbol{u}\|_{0,\Omega} \lesssim \|\boldsymbol{f}\|_{0,\Omega} + k^2 \|\boldsymbol{u}\|_{0,\Omega} \lesssim k \|\boldsymbol{f}\|_{0,\Omega}.$$

On the other hand, for every $\boldsymbol{\phi} \in \boldsymbol{\mathcal{D}}(\Omega)$, we have

$$-k^2 \langle \boldsymbol{\nabla} \times \boldsymbol{u}, \boldsymbol{\phi} \rangle_\Omega = -k^2 (\boldsymbol{u}, \boldsymbol{\nabla} \times \boldsymbol{\phi})_\Omega = b(\boldsymbol{u}, \boldsymbol{\nabla} \times \boldsymbol{\phi}) = (\boldsymbol{f}, \boldsymbol{\nabla} \times \boldsymbol{\phi})_\Omega = (\boldsymbol{\nabla} \times \boldsymbol{f}, \boldsymbol{\phi})_\Omega,$$

so that $\boldsymbol{\nabla} \times \boldsymbol{u} = -k^{-2} \boldsymbol{\nabla} \times \boldsymbol{f}$. In addition, if $\boldsymbol{v} \in \boldsymbol{C}^\infty(\overline{\Omega})$, we have

$$(\boldsymbol{f}, \boldsymbol{v})_\Omega = b(\boldsymbol{u}, \boldsymbol{v}) = -k^2 (\boldsymbol{u}, \boldsymbol{v})_\Omega - ik(\boldsymbol{u} \cdot \boldsymbol{n}, \boldsymbol{v} \cdot \boldsymbol{n})_{\partial\Omega} + (\boldsymbol{\nabla} \cdot \boldsymbol{u}, \boldsymbol{\nabla} \cdot \boldsymbol{v})_\Omega,$$

$$= (-k^2 \boldsymbol{u} - \boldsymbol{\nabla} \boldsymbol{\nabla} \cdot \boldsymbol{u}, \boldsymbol{v})_\Omega + (\boldsymbol{\nabla} \cdot \boldsymbol{u} - ik\boldsymbol{u} \cdot \boldsymbol{n}, \boldsymbol{v} \cdot \boldsymbol{n})_{\partial\Omega}$$

$$= (\boldsymbol{f}, \boldsymbol{v})_\Omega + (\boldsymbol{\nabla} \cdot \boldsymbol{u} - ik\boldsymbol{u} \cdot \boldsymbol{n}, \boldsymbol{v} \cdot \boldsymbol{n})_{\partial\Omega},$$

so that $ik\boldsymbol{u} \cdot \boldsymbol{n} = \boldsymbol{\nabla} \cdot \boldsymbol{u}$ on $\partial\Omega$. Since $\boldsymbol{\nabla} \cdot \boldsymbol{u} \in H^1(\Omega)$ it follows that $\boldsymbol{u} \cdot \boldsymbol{n} \in H^{1/2}(\partial\Omega)$ with

$$\|\boldsymbol{u} \cdot \boldsymbol{n}\|_{1/2,\partial\Omega} = k^{-1} \|\boldsymbol{\nabla} \cdot \boldsymbol{u}\|_{1/2,\partial\Omega} \lesssim k^{-1} \|\boldsymbol{\nabla} \cdot \boldsymbol{u}\|_{1,\Omega} \lesssim \|\boldsymbol{f}\|_{0,\Omega}.$$

In addition, we can define $\eta \in H^1(\Omega)$ as the unique solution to

$$\begin{cases} \eta - \Delta\eta &= 0 & \text{in } \Omega, \\ \boldsymbol{\nabla}\eta \cdot \boldsymbol{n} &= \boldsymbol{u} \cdot \boldsymbol{n} & \text{on } \partial\Omega. \end{cases}$$

Since $\Omega$ is a convex polytop and $\boldsymbol{u} \cdot \boldsymbol{n} \in H^{1/2}(\partial\Omega)$, according to Theorem 23.3 of [10], we have $\eta \in H^2(\Omega)$ with $\|\eta\|_{2,\Omega} \lesssim \|\boldsymbol{u} \cdot \boldsymbol{n}\|_{1/2,\partial\Omega} \lesssim \|\boldsymbol{f}\|_{0,\Omega}$. Let $\boldsymbol{v} = \boldsymbol{u} - \boldsymbol{\nabla}\eta$, since $\boldsymbol{\nabla}\eta \in \boldsymbol{H}^1(\Omega)$ and

$\boldsymbol{v} \cdot \boldsymbol{n} = 0$ on $\partial\Omega$, we have $\boldsymbol{v} \in \boldsymbol{H}_0(\mathrm{div}, \Omega) \cap \boldsymbol{H}(\mathrm{curl}, \Omega)$. Then, Theorem 3.9 of [17] states that $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega)$, with

$$
\begin{aligned}
|\boldsymbol{v}|_{1,\Omega} &\lesssim \|\boldsymbol{\nabla} \cdot \boldsymbol{v}\|_{0,\Omega} + \|\boldsymbol{\nabla} \times \boldsymbol{v}\|_{0,\Omega} \\
&\lesssim \|\boldsymbol{\nabla} \cdot \boldsymbol{u} - \Delta\eta\|_{0,\Omega} + \|\boldsymbol{\nabla} \times \boldsymbol{u}\|_{0,\Omega} \\
&\lesssim \|\boldsymbol{\nabla} \cdot \boldsymbol{u}\|_{0,\Omega} + \|\eta\|_{2,\Omega} + \|\boldsymbol{\nabla} \times \boldsymbol{u}\|_{0,\Omega} \\
(3.16) \qquad &\lesssim \|\boldsymbol{f}\|_{0,\Omega} + k^{-2}\|\boldsymbol{\nabla} \times \boldsymbol{f}\|_{0,\Omega}.
\end{aligned}
$$

Recalling that $\boldsymbol{u} = \boldsymbol{v} + \boldsymbol{\nabla}\eta$, since $\boldsymbol{\nabla}\eta \in \boldsymbol{H}^1(\Omega)$ with $\|\boldsymbol{\nabla}\eta\|_{1,\Omega} \lesssim \|\boldsymbol{f}\|_{0,\Omega}$, estimate (3.15) follows from (3.16). □

We close this section with two Theorems that are concerned with improved regularity of $\boldsymbol{u} \cdot \boldsymbol{n}$ on $\partial\Omega$. Specifically, we establish in both cases that $\boldsymbol{u} \cdot \boldsymbol{n} \in H^1(\mathcal{F}_j)$ for every face $\mathcal{F}_j \subset \partial\Omega$. We remark that this does not happen under the sole assumption that $\boldsymbol{f} \in \boldsymbol{H}(\mathrm{curl}, \Omega)$. In Theorem 3.3, we make the assumption that $\boldsymbol{f} \in \boldsymbol{H}^{1/2}(\Omega)$ with $\boldsymbol{\nabla} \times \boldsymbol{f} = \boldsymbol{0}$. This assumption is rather technical, but is crucial in the error analysis when using duality arguments. On the other hand, Theorem 3.4 is established under the more practical assumption that $\boldsymbol{f} \in \boldsymbol{H}(\mathrm{curl}, \Omega)$ with $\mathrm{supp}\,\boldsymbol{f} \subset\subset \Omega$. This is a reasonable assumption from the application point of view, where the boundary $\partial\Omega$ is artificial and used to approximate the Sommerfeld radiation condition. As a result, $\partial\Omega$ is usually placed "sufficiently" far away from the source term to deliver an accurate approximation of the radiation condition [12, 13].

**Theorem 3.3.** *Let $\boldsymbol{f} \in \boldsymbol{H}^{1/2}(\Omega)$ with $\boldsymbol{\nabla} \times \boldsymbol{f} = \boldsymbol{0}$. There exists a unique $\boldsymbol{u} \in \mathcal{X}$ such that $b(\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{f}, \boldsymbol{v})$ for all $\boldsymbol{v} \in \mathcal{X}$. In addition, for all faces $\mathcal{F}_j$ we have $\boldsymbol{u} \cdot \boldsymbol{n} \in H^1(\mathcal{F}_j)$ with*

$$
(3.17) \qquad |\boldsymbol{u} \cdot \boldsymbol{n}|_{1,\mathcal{F}_j} \lesssim k^{-1}\|\boldsymbol{f}\|_{1/2,\Omega} + k^{1/2}\|\boldsymbol{f}\|_{0,\Omega}.
$$

*Proof.* The existence and uniqueness of $\boldsymbol{u} \in \mathcal{X}$ follows from Theorem 3.2. In addition, since $\boldsymbol{\nabla} \times \boldsymbol{f} = \boldsymbol{0}$, estimates (3.13) and (3.15) show that

$$
\|\boldsymbol{u}\|_{0,\Omega} \lesssim k^{-1}\|\boldsymbol{f}\|_{0,\Omega}, \quad \|\boldsymbol{u}\|_{1,\Omega} \lesssim \|\boldsymbol{f}\|_{0,\Omega}.
$$

Recalling property (2.1) of the interpolation norm $\|\cdot\|_{1/2,\Omega}$, we have

$$
\|\boldsymbol{u}\|_{1/2,\Omega} \lesssim k^{-1/2}\|\boldsymbol{f}\|_{0,\Omega}.
$$

On the other hand, we recall that $\boldsymbol{\nabla}\boldsymbol{\nabla} \cdot \boldsymbol{u} = -\boldsymbol{f} - k^2\boldsymbol{u} \in \boldsymbol{H}^{1/2}(\Omega)$ in the sense of distributions, so that $\boldsymbol{\nabla} \cdot \boldsymbol{u} \in H^{3/2}(\Omega)$ with

$$
|\boldsymbol{\nabla} \cdot \boldsymbol{u}|_{3/2,\Omega} \lesssim \|\boldsymbol{f}\|_{1/2,\Omega} + k^2\|\boldsymbol{u}\|_{1/2,\Omega} \lesssim \|\boldsymbol{f}\|_{1/2,\Omega} + k^{3/2}\|\boldsymbol{f}\|_{0,\Omega}.
$$

As a result, recalling (3.13), we have

$$
(3.18) \qquad \|\boldsymbol{\nabla} \cdot \boldsymbol{u}\|_{3/2,\Omega} \lesssim \|\boldsymbol{f}\|_{1/2,\Omega} + k^{3/2}\|\boldsymbol{f}\|_{0,\Omega}.
$$

Since $\Omega$ is a convex polytop, its boundary is Lipschitz, and we can use Theorem 1.4.3.1 of [18], which states that that $\boldsymbol{\nabla} \cdot \boldsymbol{u}$ admits an extension $P(\boldsymbol{\nabla} \cdot \boldsymbol{u}) \in H^{3/2}(\mathbb{R}^d)$ with

$$
(3.19) \qquad \|P(\boldsymbol{\nabla} \cdot \boldsymbol{u})\|_{3/2,\mathbb{R}^d} \lesssim \|\boldsymbol{\nabla} \cdot \boldsymbol{u}\|_{3/2,\Omega}.
$$

But then, Lemma 16.1 of [30] states that for all $j \in \{1, \dots, N_{\mathcal{F}}\}$, $P(\boldsymbol{\nabla} \cdot \boldsymbol{u}) \in H^1(\mathcal{H}_j)$ with

$$
(3.20) \qquad \|P(\boldsymbol{\nabla} \cdot \boldsymbol{u})\|_{1,\mathcal{H}_j} \lesssim \|P(\boldsymbol{\nabla} \cdot \boldsymbol{u})\|_{3/2,\mathbb{R}^d}.
$$

Finally, since $P(\boldsymbol{\nabla} \cdot \boldsymbol{u}) = \boldsymbol{\nabla} \cdot \boldsymbol{u}$ on $\mathcal{F}_j$, we have $\boldsymbol{\nabla} \cdot \boldsymbol{u} \in H^1(\mathcal{F}_j)$, and it follows from (3.18), (3.19) and (3.20) that

$$
\|\boldsymbol{\nabla} \cdot \boldsymbol{u}\|_{1,\mathcal{F}_j} \lesssim \|\boldsymbol{f}\|_{1/2,\Omega} + k^{3/2}\|\boldsymbol{f}\|_{0,\Omega}
$$

Then, (3.17) follows since $ik\boldsymbol{u} \cdot \boldsymbol{n} = \boldsymbol{\nabla} \cdot \boldsymbol{u}$ on $\partial\Omega$. □

**Theorem 3.4.** *Assume that $\boldsymbol{f} \in \boldsymbol{H}(\mathrm{curl}, \Omega)$ with $\mathrm{supp}\,\boldsymbol{f} \subset\subset \Omega$. There exists a unique $\boldsymbol{u} \in \boldsymbol{\mathcal{X}}$ such that $b(\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{f}, \boldsymbol{v})$ for all $\boldsymbol{v} \in \boldsymbol{\mathcal{X}}$. In addition, for all faces $\mathcal{F}_j$ of $\partial\Omega$, we have $\boldsymbol{u} \cdot \boldsymbol{n} \in H^1(\mathcal{F}_j)$ with*

$$(3.21) \qquad |\boldsymbol{u} \cdot \boldsymbol{n}|_{1,\mathcal{F}_j} \lesssim \ell^{-3/2}\left(k^{1/2}\|\boldsymbol{f}\|_{0,\Omega} + k^{-3/2}\|\boldsymbol{\nabla} \times \boldsymbol{f}\|_{0,\Omega}\right),$$

*where $\ell = \mathrm{dist}(\partial\Omega, \mathrm{supp}\,\boldsymbol{f})$.*

*Proof.* Since $\mathrm{supp}\,\boldsymbol{f} \subset\subset \Omega$, there exists a set $U \subset \mathbb{R}^d$ with a $C^\infty$ boundary $\Gamma$ such that $\mathrm{supp}\,\boldsymbol{f} \subset\subset U \subset\subset \Omega$. We introduce the notation $\widetilde{\Omega} = \Omega \setminus U$. We observe that $\partial\widetilde{\Omega} = \partial\Omega \cup \Gamma$. Since $\boldsymbol{f}$ vanishes on $\widetilde{\Omega}$, we have $\boldsymbol{\nabla}\boldsymbol{\nabla} \cdot \boldsymbol{u} = -k^2\boldsymbol{u}$ on $\widetilde{\Omega}$. Recalling Theorem 3.2 and in particular (3.15), we have $\boldsymbol{u} \in \boldsymbol{H}^1(\Omega)$ and it follows that $\boldsymbol{\nabla} \cdot \boldsymbol{u} \in H^2(\widetilde{\Omega})$ with

$$|\boldsymbol{\nabla} \cdot \boldsymbol{u}|_{2,\widetilde{\Omega}} = k^2|\boldsymbol{u}|_{1,\widetilde{\Omega}} \lesssim k^2|\boldsymbol{u}|_{1,\Omega} \lesssim k^2\|\boldsymbol{f}\|_{0,\Omega} + \|\boldsymbol{\nabla} \times \boldsymbol{f}\|_{0,\Omega}.$$

Also, recalling (3.13) and (3.14), we have

$$\|\boldsymbol{\nabla} \cdot \boldsymbol{u}\|_{1,\widetilde{\Omega}} \lesssim k\|\boldsymbol{f}\|_{0,\Omega},$$

and

$$\|\boldsymbol{\nabla} \cdot \boldsymbol{u}\|_{2,\widetilde{\Omega}} \lesssim \|\boldsymbol{\nabla} \cdot \boldsymbol{u}\|_{1,\widetilde{\Omega}} + |\boldsymbol{\nabla} \cdot \boldsymbol{u}|_{2,\widetilde{\Omega}} \lesssim k^2\|\boldsymbol{f}\|_{0,\Omega} + \|\boldsymbol{\nabla} \times \boldsymbol{f}\|_{0,\Omega}.$$

There exists a cutoff function $\chi \in C^\infty(\overline{\Omega})$ such that $\chi = 1$ on $\partial\Omega$ and $\chi = 0$ in $U$. Furthermore, we can choose $\chi$ such that $\|\chi\|_{j,\infty,\Omega} \lesssim \ell^{-j}$. We introduce the function $v = \chi\boldsymbol{\nabla} \cdot \boldsymbol{u}$. Since $\boldsymbol{\nabla} \cdot \boldsymbol{u} \in H^2(\widetilde{\Omega})$ and $\chi = 0$ on $U = \Omega \setminus \widetilde{\Omega}$, we see that $v \in H^2(\Omega)$ with

$$\|v\|_{1,\Omega} \lesssim \|\chi\|_{1,\infty,\Omega}\|\boldsymbol{\nabla} \cdot \boldsymbol{u}\|_{1,\widetilde{\Omega}} \lesssim \ell^{-1}k\|\boldsymbol{f}\|_{0,\Omega},$$

and

$$\|v\|_{2,\Omega} \lesssim \|\chi\|_{2,\infty,\Omega}\|\boldsymbol{\nabla} \cdot \boldsymbol{u}\|_{2,\widetilde{\Omega}} \lesssim \ell^{-2}\left(k^2\|\boldsymbol{f}\|_{0,\Omega} + \|\boldsymbol{\nabla} \times \boldsymbol{f}\|_{0,\Omega}\right).$$

Recalling that $H^{3/2}(\Omega)$ can be defined by interpolation between $H^1(\Omega)$ and $H^2(\Omega)$, (2.1) yields

$$\begin{aligned}
\|v\|_{3/2,\Omega}^2 &\lesssim \|v\|_{1,\Omega}\|v\|_{2,\Omega} \\
&\lesssim \ell^{-1}k\|\boldsymbol{f}\|_{0,\Omega}\ell^{-2}\left(k^2\|\boldsymbol{f}\|_{0,\Omega} + \|\boldsymbol{\nabla} \times \boldsymbol{f}\|_{0,\Omega}\right) \\
&\lesssim \ell^{-3}\left(k^3\|\boldsymbol{f}\|_{0,\Omega}^2 + k\|\boldsymbol{f}\|_{0,\Omega}\|\boldsymbol{\nabla} \times \boldsymbol{f}\|_{0,\Omega}\right) \\
&\lesssim \ell^{-3}\left(k^3\|\boldsymbol{f}\|_{0,\Omega}^2 + k^{-1}\|\boldsymbol{\nabla} \times \boldsymbol{f}\|_{0,\Omega}^2\right),
\end{aligned}$$

so that

$$(3.22) \qquad \|v\|_{3/2,\widetilde{\Omega}} \lesssim \ell^{-3/2}\left(k^{3/2}\|\boldsymbol{f}\|_{0,\Omega} + k^{-1/2}\|\boldsymbol{\nabla} \times \boldsymbol{f}\|_{0,\Omega}\right).$$

Then, we conclude as in Theorem 3.3. Since $\Omega$ has a Lipschitz boundary Theorem 1.4.3.1 of [18] states that that $v$ admits an extension $Pv \in H^{3/2}(\mathbb{R}^d)$ with

$$(3.23) \qquad \|Pv\|_{3/2,\mathbb{R}^d} \lesssim \|v\|_{3/2,\Omega},$$

and Lemma 16.1 of [30] shows that for all $j \in \{1, \ldots, N_{\mathcal{F}}\}$, $Pv \in H^1(\mathcal{H}_j)$ with

$$(3.24) \qquad \|Pv\|_{1,\mathcal{H}_j} \lesssim \|Pv\|_{3/2,\mathbb{R}^d}.$$

Finally, since $Pv = v = \boldsymbol{\nabla} \cdot \boldsymbol{u}$ on $\mathcal{F}_j$, we have $\boldsymbol{\nabla} \cdot \boldsymbol{u} \in H^1(\mathcal{F}_j)$, and it follows from (3.22), (3.23) and (3.24) that

$$\|\boldsymbol{\nabla} \cdot \boldsymbol{u}\|_{1,\mathcal{F}_j} \lesssim \ell^{-3/2}\left(k^{3/2}\|\boldsymbol{f}\|_{0,\Omega} + k^{-1/2}\|\boldsymbol{\nabla} \times \boldsymbol{f}\|_{0,\Omega}\right).$$

Then, we obtain (3.21) since $\boldsymbol{\nabla} \cdot \boldsymbol{u} = ik\boldsymbol{u} \cdot \boldsymbol{n}$ on $\partial\Omega$. $\square$

## 4. FINITE ELEMENT DISCRETIZATIONS

In this section, we consider a fixed right-hand side $\boldsymbol{\psi} \in \boldsymbol{\mathcal{X}}'$ and denote by $\boldsymbol{u}$ the unique element of $\boldsymbol{\mathcal{X}}$ such that $b(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{\psi}(\boldsymbol{v})$ for all $\boldsymbol{v} \in \boldsymbol{\mathcal{X}}$. We analyze the approximation of $\boldsymbol{u}$ by Raviart-Thomas finite elements [23, 25]. Since $\Omega$ is a polytop it is possible to mesh the domain with simplexes (triangle or tetrahedra depending on $d$). Thus, we consider a family of simplicial meshes $(\mathcal{T}_h)_{h>0}$ of $\Omega$ that is shape regular in the sense of [6]. Specifically, we assume that there exists a constant $\beta \in (0, 1)$ such that for all $h > 0$, and for all $K \in \mathcal{T}_h$, we have

$$
(4.1) \qquad \mathrm{diam}(K) \leq h, \quad \frac{\rho(K)}{\mathrm{diam}(K)} \geq \beta,
$$

where $\mathrm{diam}(K)$ is the diameter of $K$ and $\rho(K)$ is the diameter of the largest ball contained in $\overline{K}$. We also denote by $\mathcal{F}_h$ the set of faces of the mesh $\mathcal{T}_h$. We associate with each mesh $\mathcal{T}_h$ two finite element spaces. On the one hand, the $\boldsymbol{H}(\mathrm{div}, \Omega)$ conforming space of lowest order Raviart-Thomas elements defined in [23] as

$$
\boldsymbol{\mathcal{X}}_h = \{\boldsymbol{v}_h \in \boldsymbol{H}(\mathrm{div}, \Omega) \mid \boldsymbol{v}_h|_K \in R \ \forall K \in \mathcal{T}_h\}, \quad R = \{\boldsymbol{v} : \boldsymbol{x} \to \boldsymbol{a} + b\boldsymbol{x}, \quad \boldsymbol{a} \in \mathbb{C}^3, \ b \in \mathbb{C}\}.
$$

On the other hand, we also require the $\boldsymbol{H}_0(\mathrm{curl}, \Omega)$ conforming space of lowest order Nédélec edge elements that are defined in [23] as

$$
\boldsymbol{\mathcal{W}}_h = \{\boldsymbol{v}_h \in \boldsymbol{H}_0(\mathrm{curl}, \Omega) \mid \boldsymbol{v}_h|_K \in N \ \forall K \in \mathcal{T}_h\}, \quad N = \{\boldsymbol{v} : \boldsymbol{x} \to \boldsymbol{a} + \boldsymbol{b} \times \boldsymbol{x}, \quad \boldsymbol{a}, \boldsymbol{b} \in \mathbb{C}^3\}.
$$

We use the space $\boldsymbol{\mathcal{X}}_h$ to build the discretization to (1.1), while we employ the space $\boldsymbol{\mathcal{W}}_h$ as a tool in the stability and error analysis. We note that $\boldsymbol{\mathcal{X}}_h \subset \boldsymbol{\mathcal{X}}$ and that $\boldsymbol{\nabla} \times \boldsymbol{\mathcal{W}}_h \subset \boldsymbol{\mathcal{X}}_h$.

Our analysis requires two different interpolation operators. On the one hand, the standard Raviart-Thomas interpolant $\boldsymbol{\mathcal{I}}_h : \boldsymbol{H}^1(\Omega) \to \boldsymbol{\mathcal{X}}_h$ is defined by face-wise $L^2$ projection, and for $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega)$ we have

$$
(\boldsymbol{\mathcal{I}}_h\boldsymbol{v}) \cdot \boldsymbol{n}_F|_F = \frac{1}{|F|} \int_F \boldsymbol{v} \cdot \boldsymbol{n}_F, \quad \forall F \in \mathcal{F}_h.
$$

As shown in Theorem 5.25 of [22], we have the following standard interpolation error estimates. For all $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega)$, it holds that

$$
(4.2) \qquad \|\boldsymbol{v} - \boldsymbol{\mathcal{I}}_h\boldsymbol{v}\|_{0,\Omega} \lesssim h|\boldsymbol{v}|_{1,\Omega}.
$$

If in addition, $\boldsymbol{\nabla} \cdot \boldsymbol{v} \in H^1(\Omega)$, we also have

$$
(4.3) \qquad \|\boldsymbol{\nabla} \cdot (\boldsymbol{v} - \boldsymbol{\mathcal{I}}_h\boldsymbol{v})\|_{0,\Omega} \lesssim h|\boldsymbol{\nabla} \cdot \boldsymbol{v}|_{1,\Omega}.
$$

We also establish a useful result concerning the approximation properties of $\boldsymbol{\mathcal{I}}_h$ on the boundary $\partial\Omega$ in Lemma 4.1.

**Lemma 4.1.** *Let $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega)$. If $\boldsymbol{v} \cdot \boldsymbol{n} \in H^1(\mathcal{F}_j)$ for all $j \in \{1, \dots, N_{\mathcal{F}}\}$, we have*

$$
(4.4) \qquad \|(\boldsymbol{v} - \boldsymbol{\mathcal{I}}_h\boldsymbol{v}) \cdot \boldsymbol{n}\|_{0,\partial\Omega} \lesssim h \left( \sum_{j=1}^{N_{\mathcal{F}}} |\boldsymbol{v} \cdot \boldsymbol{n}|_{1,\mathcal{F}_j}^2 \right)^{1/2}.
$$

*Proof.* Let $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega)$ such that $\boldsymbol{v} \cdot \boldsymbol{n} \in H^1(\mathcal{F}_j)$ for all $j \in \{1, \dots, N_{\mathcal{F}}\}$. We denote by $\mathcal{F}_h^{\mathrm{ext}}$ the set of faces $F \in \mathcal{F}_h$ such that $F \subset \partial\Omega$. Note that since $F \subset \mathcal{F}_j$ for a single face $\mathcal{F}_j$ of $\partial\Omega$, we have $\boldsymbol{v} \cdot \boldsymbol{n} \in H^1(F)$. By definition of $\boldsymbol{\mathcal{I}}_h$, on each $F \in \mathcal{F}_h^{\mathrm{ext}}$, we have

$$
(\boldsymbol{\mathcal{I}}_h\boldsymbol{v}) \cdot \boldsymbol{n}|_F = \frac{1}{|F|} \int_F \boldsymbol{v} \cdot \boldsymbol{n} \in \mathcal{P}_0(F),
$$

where $\mathcal{P}_0(F)$ stands for the space of constant functions on $F$. Specifically, we see that $(\boldsymbol{\mathcal{I}}_h\boldsymbol{v}) \cdot \boldsymbol{n}$ is the orthogonal projection of $\boldsymbol{v} \cdot \boldsymbol{n}$ in $\mathcal{P}_0(F)$ for the $L^2(F)$ inner product. We recall that since $F$ is convex with $\mathrm{diam}\, F \leq h$. Then, as shown in [24], Poincaré-Wirtinger inequality yields that

$$
\|\boldsymbol{v} \cdot \boldsymbol{n} - (\boldsymbol{\mathcal{I}}_h\boldsymbol{v} \cdot \boldsymbol{n})\|_{0,F} \lesssim h|\boldsymbol{v} \cdot \boldsymbol{n}|_{1,F},
$$

and (4.4) follows by summation over all faces $F \in \mathcal{F}_h^{\mathrm{ext}}$. $\qquad\square$

We remark that the $L^2(F)$-projection of the normal trace of $\boldsymbol{v}$ over $F \in \mathcal{T}_h$ is involved in the definition of $\mathcal{I}_h\boldsymbol{v}$, so that this operator is still well-defined if we only have $\boldsymbol{v} \in \boldsymbol{H}^{1/2+\varepsilon}(\Omega)$ with $\varepsilon > 0$, but not if $\varepsilon = 0$. A key point of our analysis is the approximation of a function that precisely lies in $\boldsymbol{H}^{1/2}(\Omega)$. This motivate the introduction of a second "quasi-interpolation" operator, that is well-defined as soon as the interpolated function is in $\boldsymbol{L}^2(\Omega)$. We thus consider a quasi-interpolation operator $\boldsymbol{\pi}_h : \boldsymbol{L}^2(\Omega) \to \boldsymbol{\mathcal{X}}_h$ following [14, 28]. As a particular case of Theorem 5 and Corollary 6 of [28] if $\boldsymbol{v} \in \boldsymbol{H}^s(\Omega)$ for some $s \in (0,1]$, we have[1]

$$\|\boldsymbol{v} - \boldsymbol{\pi}_h\boldsymbol{v}\|_{0,\Omega} \lesssim h^s |\boldsymbol{v}|_{s,\Omega}. \tag{4.5}$$

Actually, we will also need a special quasi-interpolation operator for functions in $\boldsymbol{\mathcal{X}}$ of the form $\boldsymbol{\nabla} \times \boldsymbol{\omega}$ for some $\boldsymbol{\omega} \in \boldsymbol{H}_0(\mathrm{curl}, \Omega)$. As a result (recall that $\boldsymbol{\nabla} \times \boldsymbol{\mathcal{W}}_h \subset \boldsymbol{\mathcal{X}}_h$), following [14, 28], we introduce quasi-interpolation operator $\boldsymbol{\mathcal{P}}_h : \boldsymbol{L}^2(\Omega) \to \boldsymbol{\mathcal{W}}_h$. The operators $\boldsymbol{\pi}_h$ and $\boldsymbol{\mathcal{P}}_h$ commute, so that for all $\boldsymbol{v} \in \boldsymbol{H}_0(\mathrm{curl}, \Omega)$, we have

$$\boldsymbol{\pi}_h \left( \boldsymbol{\nabla} \times \boldsymbol{v} \right) = \boldsymbol{\nabla} \times \left( \boldsymbol{\mathcal{P}}_h \boldsymbol{v} \right). \tag{4.6}$$

To the best of the author's knowledge, the quasi-interpolation $\boldsymbol{\pi}_h$ does not exhibit special approximation properties like (4.4) on the boundary. This is the reason why we introduce two interpolation operators, as (4.4) also plays an important role in the upcoming analysis.

From now on, we assume that $\boldsymbol{u}_h$ is an element of $\boldsymbol{\mathcal{X}}_h$ satisfying $b(\boldsymbol{u}_h, \boldsymbol{v}_h) = \psi(\boldsymbol{v}_h)$ for all $\boldsymbol{v}_h \in \boldsymbol{\mathcal{X}}_h$. Note that we do not know a priori that such a $\boldsymbol{u}_h$ exists. Our analysis closely follows the "Schatz argument" employed for the pressure formulation [20, 21, 27]. However, special arguments are required to take advantage of this duality technique with mixed finite elements [5, 15, 32]. Also, in addition to the arguments from [5, 15, 32] that are established for "Dirichlet-like" boundary conditions, extra care is taken to analyze the impedance boundary conditions.

Following [5], the Helmholtz-Hodge decomposition of the error vector will be a central tool in the following. According to Corollary 3.4 of [17], there exist an element $\boldsymbol{q} \in \boldsymbol{H}(\mathrm{div}, \Omega)$ with $\boldsymbol{\nabla} \times \boldsymbol{q} = \boldsymbol{0}$ and a $\boldsymbol{\omega} \in \boldsymbol{H}_0(\mathrm{curl}, \Omega)$ such that

$$\boldsymbol{u} - \boldsymbol{u}_h = \boldsymbol{q} + \boldsymbol{\nabla} \times \boldsymbol{\omega}. \tag{4.7}$$

In addition, since $\boldsymbol{\pi}_h\boldsymbol{u}_h = \boldsymbol{u}_h$, commutativity property (4.6) shows that

$$\boldsymbol{\pi}_h\boldsymbol{u} - \boldsymbol{u}_h = \boldsymbol{\pi}_h\boldsymbol{q} + \boldsymbol{\nabla} \times (\boldsymbol{\mathcal{P}}_h\boldsymbol{\omega}). \tag{4.8}$$

We also note that Helmholtz-Hodge decomposition (4.7) is orthogonal. Specifically, we have

$$(\boldsymbol{q}, \boldsymbol{\nabla} \times \boldsymbol{\phi}) = 0, \quad (\boldsymbol{\nabla} \times \boldsymbol{\omega}, \boldsymbol{p}) = 0 \tag{4.9}$$

for all $\boldsymbol{\phi} \in \boldsymbol{H}_0(\mathrm{curl}, \Omega)$ and $\boldsymbol{p} \in \boldsymbol{H}(\mathrm{div}, \Omega)$ with $\boldsymbol{\nabla} \times \boldsymbol{p} = \boldsymbol{0}$.

We start our analysis with a result concerning the regularity of $\boldsymbol{q}$.

**Lemma 4.2.** *We have $\boldsymbol{q} \in \boldsymbol{H}^{1/2}(\Omega)$. In addition, it holds that*

$$\|\boldsymbol{q}\|_{1/2,\Omega} \lesssim k^{-1/2}\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{\mathcal{X}}}. \tag{4.10}$$

*Proof.* Let us first introduce $p \in H^1(\Omega)$ as the weak solution to

$$\begin{cases} -\Delta p &= \boldsymbol{\nabla} \cdot (\boldsymbol{u} - \boldsymbol{u}_h) & \text{in } \Omega, \\ p &= 0 & \text{on } \partial\Omega. \end{cases}$$

Since $\Omega$ is convex [18], standard regularity results for the Laplace operator imply that $p \in H^2(\Omega)$ and provide the estimate

$$\|\boldsymbol{\nabla}p\|_{1,\Omega} \le \|p\|_{2,\Omega} \lesssim \|\boldsymbol{\nabla} \cdot (\boldsymbol{u} - \boldsymbol{u}_h)\|_{0,\Omega}. \tag{4.11}$$

In addition, by integrating by parts with $p$, we see that

$$\|\boldsymbol{\nabla}p\|_{0,\Omega}^2 = (-\Delta p, p)_\Omega = (\boldsymbol{\nabla} \cdot (\boldsymbol{u} - \boldsymbol{u}_h), p)_\Omega = (\boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{\nabla}p)_\Omega,$$

so that

$$\|\boldsymbol{\nabla}p\|_{0,\Omega} \lesssim \|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,\Omega}. \tag{4.12}$$

---

[1]The result is actually stated for $s = 1$, but the general case easily follows by interpolation. See also [14].

As a result, recalling interpolation property (2.1) of the $\| \cdot \|_{1/2,\Omega}$-norm, a direct application of (4.11) and (4.12) shows that

$$(4.13) \quad \|\boldsymbol{\nabla} p\|^2_{1/2,\Omega} \lesssim \|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,\Omega} \|\boldsymbol{\nabla} \cdot (\boldsymbol{u} - \boldsymbol{u}_h)\|_{0,\Omega}$$
$$\lesssim k\|\boldsymbol{u} - \boldsymbol{u}_h\|^2_{0,\Omega} + k^{-1}\|\boldsymbol{\nabla} \cdot (\boldsymbol{u} - \boldsymbol{u}_h)\|^2_{0,\Omega} \lesssim k^{-1}\|\boldsymbol{u} - \boldsymbol{u}_h\|^2_{\boldsymbol{\mathcal{X}}}.$$

In addition, since $\boldsymbol{\nabla} p \in \boldsymbol{H}^1(\Omega)$ we have $\boldsymbol{\nabla} p \cdot \boldsymbol{n} \in L^2(\partial\Omega)$. By applying trace Theorem 1.5.1.10 of [18] to each component of $\boldsymbol{\nabla} p$, we see that

$$(4.14) \qquad \|\boldsymbol{\nabla} p \cdot \boldsymbol{n}\|^2_{0,\partial\Omega} \lesssim \|\boldsymbol{\nabla} p\|^2_{0,\partial\Omega} \lesssim k\|\boldsymbol{\nabla} p\|^2_{0,\Omega} + k^{-1}|\boldsymbol{\nabla} p|^2_{1,\Omega} \lesssim k^{-1}\|\boldsymbol{u} - \boldsymbol{u}_h\|^2_{\boldsymbol{\mathcal{X}}}.$$

Letting $\boldsymbol{\eta} = \boldsymbol{q} - \boldsymbol{\nabla} p \in \boldsymbol{L}^2(\Omega)$, we see that

$$\boldsymbol{\nabla} \cdot \boldsymbol{\eta} = \boldsymbol{\nabla} \cdot \boldsymbol{q} - \Delta p = 0, \quad \boldsymbol{\nabla} \times \boldsymbol{\eta} = \boldsymbol{\nabla} \times \boldsymbol{q} = \boldsymbol{0}.$$

Since $\boldsymbol{u}, \boldsymbol{u}_h \in \boldsymbol{\mathcal{X}}$ and $\boldsymbol{\nabla} \times \boldsymbol{\omega} \cdot \boldsymbol{n} = 0$ on $\partial\Omega$, we have $\boldsymbol{q} \cdot \boldsymbol{n} = (\boldsymbol{u} - \boldsymbol{u}_h) \cdot \boldsymbol{n} \in L^2(\partial\Omega)$. Thus, recalling (4.14), we have $\boldsymbol{\eta} \cdot \boldsymbol{n} \in L^2(\partial\Omega)$, and Theorem 2 of [9] states that $\boldsymbol{\eta} \in \boldsymbol{H}^{1/2}(\Omega)$ with

$$(4.15) \qquad \|\boldsymbol{\eta}\|_{1/2,\Omega} \lesssim \|\boldsymbol{\eta} \cdot \boldsymbol{n}\|_{0,\partial\Omega} \lesssim \|\boldsymbol{q} \cdot \boldsymbol{n}\|_{0,\Omega} + \|\boldsymbol{\nabla} p \cdot \boldsymbol{n}\|_{0,\Omega} \lesssim k^{-1/2}\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{\mathcal{X}}}.$$

Since $\boldsymbol{q} = \boldsymbol{\nabla} p + \boldsymbol{\eta}$, estimate (4.10) is a direct consequence of (4.13) and (4.15). $\qquad \square$

Having established Lemma 4.2, we are ready to derive an Aubin-Nitsche like estimate using a duality argument in Lemma 4.3.

**Lemma 4.3.** *We have*

$$(4.16) \qquad k\|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,\Omega} \lesssim ((kh)^{1/2} + k^2h)\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{\mathcal{X}}} + k\|\boldsymbol{u} - \boldsymbol{\pi}_h\boldsymbol{u}\|_{0,\Omega}.$$

*Proof.* We first establish that

$$(4.17) \qquad \|\boldsymbol{q}\|_{0,\Omega} \lesssim (kh + k^{-1/2}h^{1/2})\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{\mathcal{X}}}.$$

Since $\boldsymbol{q} \in \boldsymbol{L}^2(\Omega)$, following Theorem 3.2, there exists a unique $\boldsymbol{\xi} \in \boldsymbol{\mathcal{X}}$ such that $b(\boldsymbol{v}, \boldsymbol{\xi}) = (\boldsymbol{v}, \boldsymbol{q})$ for all $\boldsymbol{v} \in \boldsymbol{\mathcal{X}}$, with[2]

$$(4.18) \qquad |\boldsymbol{\nabla} \cdot \boldsymbol{\xi}|_{1,\Omega} \lesssim k\|\boldsymbol{q}\|_{0,\Omega}.$$

In addition, since $\boldsymbol{\nabla} \times \boldsymbol{q} = \boldsymbol{0}$, we have $\boldsymbol{\xi} \in \boldsymbol{H}^1(\Omega)$ with

$$(4.19) \qquad |\boldsymbol{\xi}|_{1,\Omega} \lesssim \|\boldsymbol{q}\|_{0,\Omega}.$$

Finally, recalling Lemma 4.2, we have $\boldsymbol{q} \in \boldsymbol{H}^{1/2}(\Omega)$ and therefore, $\boldsymbol{\xi} \cdot \boldsymbol{n} \in H^1(\partial\Omega)$ with

$$(4.20) \qquad |\boldsymbol{\xi} \cdot \boldsymbol{n}|_{1,\partial\Omega} \lesssim k^{1/2}\|\boldsymbol{q}\|_{0,\Omega} + k^{-1}|\boldsymbol{q}|_{1/2,\Omega} \lesssim k^{1/2}\|\boldsymbol{q}\|_{0,\Omega} + k^{-3/2}\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{\mathcal{X}}}.$$

As a direct consequence of (4.18), (4.19) and (4.20), we have

$$\|\boldsymbol{\xi} - \boldsymbol{\mathcal{I}}_h\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{X}}} \lesssim kh\|\boldsymbol{q}\|_{0,\Omega} + k^{-1}h\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{\mathcal{X}}}.$$

On the other hand, recalling (4.9), we have

$$\|\boldsymbol{q}\|^2_{0,\Omega} = (\boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{q}) = b(\boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{\xi}) = b(\boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{\xi} - \boldsymbol{\mathcal{I}}_h\boldsymbol{\xi})$$
$$\lesssim \|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{\mathcal{X}}}\|\boldsymbol{\xi} - \boldsymbol{\mathcal{I}}_h\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{X}}}$$
$$\lesssim kh\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{\mathcal{X}}}\|\boldsymbol{q}\|_{0,\Omega} + k^{-1}h\|\boldsymbol{u} - \boldsymbol{u}_h\|^2_{\boldsymbol{\mathcal{X}}}.$$

Using Young's inequality, we obtain

$$\|\boldsymbol{q}\|^2_{0,\Omega} \lesssim k^2h^2\|\boldsymbol{u} - \boldsymbol{u}_h\|^2_{\boldsymbol{\mathcal{X}}} + k^{-1}h\|\boldsymbol{u} - \boldsymbol{u}_h\|^2_{\boldsymbol{\mathcal{X}}},$$

and (4.17) follows by taking the square root.

We now establish that

$$(4.21) \qquad \|\boldsymbol{\nabla} \times \boldsymbol{\omega}\|_{0,\Omega} \lesssim \|\boldsymbol{u} - \boldsymbol{\pi}_h\boldsymbol{u}\|_{0,\Omega} + k^{-1/2}h^{1/2}\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{\mathcal{X}}}.$$

---

[2]Theorem 3.2 does not explicitly treat the adjoint problem, but one easily sees that $\boldsymbol{\xi}$ can be equivalently defined as the unique element of $\boldsymbol{\mathcal{X}}$ such that $b(\overline{\boldsymbol{\xi}}, \boldsymbol{v}) = (\overline{\boldsymbol{q}}, \boldsymbol{v})_\Omega$ for all $\boldsymbol{v} \in \boldsymbol{\mathcal{X}}$.

Let $\boldsymbol{\omega}_h \in \boldsymbol{\mathcal{W}}_h$. As in particular $\boldsymbol{\omega}_h \in \boldsymbol{H}_0(\operatorname{div}, \Omega)$, recalling orthogonality property (4.9) of the Hodge-Helmholtz decomposition, we have

$$-k^2(\boldsymbol{\nabla} \times \boldsymbol{\omega}, \boldsymbol{\nabla} \times \boldsymbol{\omega}_h) = -k^2(\boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{\nabla} \times \boldsymbol{\omega}_h) = b(\boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{\nabla} \times \boldsymbol{\omega}_h) = 0$$

since $\boldsymbol{\nabla} \times \boldsymbol{\omega}_h \in \boldsymbol{\mathcal{X}}_h$. Thus $(\boldsymbol{\nabla} \times \boldsymbol{\omega}_h, \boldsymbol{\nabla} \times \boldsymbol{\omega}) = 0$ for all $\boldsymbol{\omega}_h \in \boldsymbol{\mathcal{W}}_h$. Then, recalling (4.8), it holds that

$$
\begin{aligned}
\|\boldsymbol{\nabla} \times \boldsymbol{\omega}\|_{0,\Omega}^2 & = & (\boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{\nabla} \times \boldsymbol{\omega}) \\
& = & (\boldsymbol{u} - \boldsymbol{\pi}_h \boldsymbol{u}, \boldsymbol{\nabla} \times \boldsymbol{\omega}) + (\boldsymbol{\pi}_h \boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{\nabla} \times \boldsymbol{\omega}) \\
& = & (\boldsymbol{u} - \boldsymbol{\pi}_h \boldsymbol{u}, \boldsymbol{\nabla} \times \boldsymbol{\omega}) + (\boldsymbol{\pi}_h \boldsymbol{q} - \boldsymbol{\nabla} \times (\mathcal{P}_h \boldsymbol{\omega}), \boldsymbol{\nabla} \times \boldsymbol{\omega}) \\
& = & (\boldsymbol{u} - \boldsymbol{\pi}_h \boldsymbol{u}, \boldsymbol{\nabla} \times \boldsymbol{\omega}) + (\boldsymbol{\pi}_h \boldsymbol{q}, \boldsymbol{\nabla} \times \boldsymbol{\omega}) \\
& = & (\boldsymbol{u} - \boldsymbol{\pi}_h \boldsymbol{u}, \boldsymbol{\nabla} \times \boldsymbol{\omega}) + (\boldsymbol{\pi}_h \boldsymbol{q} - \boldsymbol{q}, \boldsymbol{\nabla} \times \boldsymbol{\omega}),
\end{aligned}
$$

where we have also used (4.9) in the first and last equalities (recall that $\boldsymbol{\nabla} \times \boldsymbol{q} = \boldsymbol{0}$). As a result, we have

$$\|\boldsymbol{\nabla} \times \boldsymbol{\omega}\|_{0,\Omega} \lesssim \|\boldsymbol{u} - \boldsymbol{\pi}_h \boldsymbol{u}\|_{0,\Omega} + \|\boldsymbol{q} - \boldsymbol{\pi}_h \boldsymbol{q}\|_{0,\Omega}.$$

Then, estimate (4.21) follows using (4.5) and (4.10), since it holds that

$$\|\boldsymbol{q} - \boldsymbol{\pi}_h \boldsymbol{q}\|_{0,\Omega} \lesssim h^{1/2} |\boldsymbol{q}|_{1/2,\Omega} \lesssim k^{-1/2} h^{1/2} \|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{\mathcal{X}}}.$$

Having proved (4.17) and (4.21), (4.16) immediately follows since

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,\Omega} \lesssim \|\boldsymbol{q}\|_{0,\Omega} + \|\boldsymbol{\nabla} \times \boldsymbol{\omega}\|_{0,\Omega}.$$

$\square$

We now deliver the main results of this section in Theorem 4.1. Specifically, we establish the existence and uniqueness of a finite element solution under the condition that $k^2 h$, and we show that in this case the finite element solution is "almost" quasi-optimal.

**Theorem 4.1.** *Under the assumption that $k^2 h$ is sufficiently small, there exists a unique $\boldsymbol{u}_h \in \boldsymbol{\mathcal{X}}_h$ such that $b(\boldsymbol{u}_h, \boldsymbol{v}_h) = \boldsymbol{\psi}(\boldsymbol{v}_h)$ for all $\boldsymbol{v}_h \in \boldsymbol{\mathcal{X}}_h$, and we have*

$$(4.22) \qquad \|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{\mathcal{X}}} \lesssim k \|\boldsymbol{u} - \boldsymbol{\pi}_h \boldsymbol{u}\|_{0,\Omega} + \inf_{\boldsymbol{v}_h \in \boldsymbol{\mathcal{X}}_h} \|\boldsymbol{u} - \boldsymbol{v}_h\|_{\boldsymbol{\mathcal{X}}}.$$

*Proof.* Before proving the existence and uniqueness of $\boldsymbol{u}_h \in \boldsymbol{\mathcal{X}}_h$, let us first establish (4.22) for any $\boldsymbol{u}_h \in \boldsymbol{\mathcal{X}}_h$ such that $b(\boldsymbol{u}_h, \boldsymbol{v}_h) = \boldsymbol{\psi}(\boldsymbol{v}_h)$ for all $\boldsymbol{v}_h \in \boldsymbol{\mathcal{X}}_h$.

Recalling (4.16) from Lemma 4.3, we see that we have

$$k^2 \|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,\Omega}^2 \leq C_0 (k^4 h^2 + kh) \|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{\mathcal{X}}}^2 + C_0 k^2 \|\boldsymbol{u} - \boldsymbol{\pi}_h \boldsymbol{u}\|_{0,\Omega}.$$

for some constant $C_0$ that is independent of $k$ and $h$. Then, we have

$$
\begin{aligned}
\operatorname{Re} b(\boldsymbol{u} - \boldsymbol{u}_h, (1+i)(\boldsymbol{u} - \boldsymbol{u}_h)) & \geq & \|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{\mathcal{X}}}^2 - 2k^2 \|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,\Omega}^2 \\
& \geq & \left\{ 1 - 2C_0(k^4 h^2 + kh) \right\} \|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{\mathcal{X}}}^2 - 2C_0 k^2 \|\boldsymbol{u} - \boldsymbol{\pi}_h \boldsymbol{u}\|_{0,\Omega}^2.
\end{aligned}
$$

Thus, assuming that $k^2 h$ is small enough, we have $1 - 2C_0(k^4 h^2 + kh) \geq 1/2$, and we have

$$\frac{1}{2} \|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{\mathcal{X}}}^2 \leq \operatorname{Re} b(\boldsymbol{u} - \boldsymbol{u}_h, (1+i)(\boldsymbol{u} - \boldsymbol{u}_h)) + 2C_0 k^2 \|\boldsymbol{u} - \boldsymbol{\pi}_h \boldsymbol{u}\|_{0,\Omega}^2,$$

that we rewrite

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{\mathcal{X}}}^2 \lesssim |b(\boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{u} - \boldsymbol{u}_h)| + k^2 \|\boldsymbol{u} - \boldsymbol{\pi}_h \boldsymbol{u}\|_{0,\Omega}^2.$$

Then, we using Galerkin orthogonality, we have

$$|b(\boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{u} - \boldsymbol{u}_h)| = |b(\boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{u} - \boldsymbol{v}_h)| \lesssim \|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{\mathcal{X}}}, \|\boldsymbol{u} - \boldsymbol{v}_h\|_{\boldsymbol{\mathcal{X}}}$$

and using Young's inequality, we obtain

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{\mathcal{X}}}^2 \lesssim k^2 \|\boldsymbol{u} - \boldsymbol{\pi}_h \boldsymbol{u}\|_{0,\Omega}^2 + \|\boldsymbol{u} - \boldsymbol{v}_h\|_{\boldsymbol{\mathcal{X}}}^2,$$

for all $\boldsymbol{v}_h \in \boldsymbol{\mathcal{X}}_h$. Error estimate (4.22) then follows by taking the infimum over $\boldsymbol{v}_h \in \boldsymbol{\mathcal{X}}_h$.

We now turn to the existence and uniqueness of $\boldsymbol{u}_h \in \boldsymbol{\mathcal{X}}_h$. If $\boldsymbol{\psi} = \boldsymbol{0} \in \boldsymbol{\mathcal{X}}'$, then $\boldsymbol{u} = \boldsymbol{0}$ and error estimate (4.22) shows that $\boldsymbol{u}_h = \boldsymbol{0}$. By linearity, we see that if there exists a $\boldsymbol{u}_h \in \boldsymbol{\mathcal{X}}_h$ such

that $b(\boldsymbol{u}_h, \boldsymbol{v}_h) = \boldsymbol{\psi}(\boldsymbol{v}_h)$ for all $\boldsymbol{v}_h \in \boldsymbol{\mathcal{X}}_h$, then $\boldsymbol{u}_h$ is unique. Furthermore, it is clear that $\boldsymbol{u}_h$ is defined through a finite dimensional square linear system. As a result, uniqueness of the solution implies existence, and we obtain that there exists a unique $\boldsymbol{u}_h \in \boldsymbol{\mathcal{X}}_h$ such that $b(\boldsymbol{u}_h, \boldsymbol{v}_h) = \boldsymbol{\psi}(\boldsymbol{v}_h)$ for all $\boldsymbol{v}_h \in \boldsymbol{\mathcal{X}}_h$. □

We conclude this section with an explicit error estimate, assuming some regularity on the right-hand side. As pointed out in Section 3 for Theorem 3.4, the assumption that $\boldsymbol{f} \in \boldsymbol{H}(\mathrm{curl}, \Omega)$ with $\mathrm{supp}\, \boldsymbol{f} \subset\subset \Omega$ is realistic from the application point of view.

**Corollary 4.1.** *Let* $\boldsymbol{f} \in \boldsymbol{H}(\mathrm{curl}, \Omega)$ *with* $\mathrm{supp}\, \boldsymbol{f} \subset\subset \Omega$. *Under the assumption that $k^2 h$ is sufficiently small, there exist a unique $\boldsymbol{u} \in \boldsymbol{\mathcal{X}}$ and a unique $\boldsymbol{u}_h \in \boldsymbol{\mathcal{X}}_h$ such that $b(\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{f}, \boldsymbol{v})$ for all $\boldsymbol{v} \in \boldsymbol{\mathcal{X}}$ and $b(\boldsymbol{u}_h, \boldsymbol{v}_h) = (\boldsymbol{f}, \boldsymbol{v}_h)$ for all $\boldsymbol{v}_h \in \boldsymbol{\mathcal{X}}_h$. In addition, we have*

$$(4.23) \qquad \|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{\mathcal{X}}} \lesssim \ell^{-3/2} kh \left( \|\boldsymbol{f}\|_{0,\Omega} + k^{-2} \|\boldsymbol{\nabla} \times \boldsymbol{f}\|_{0,\Omega} \right),$$

*where* $\ell = \mathrm{dist}(\partial\Omega, \mathrm{supp}\, \boldsymbol{f})$.

*Proof.* The existence and uniqueness of $\boldsymbol{u} \in \boldsymbol{\mathcal{X}}$ and $\boldsymbol{u}_h \in \boldsymbol{\mathcal{X}}_h$ follows from Theorems 3.2 and 4.1. In addition, (3.14) and (3.15) show that

$$|\boldsymbol{u}|_{1,\Omega} \lesssim \|\boldsymbol{f}\|_{0,\Omega} + k^{-2} \|\boldsymbol{\nabla} \times \boldsymbol{f}\|_{0,\Omega}, \quad |\boldsymbol{\nabla} \cdot \boldsymbol{u}|_{1,\Omega} \lesssim k \|\boldsymbol{f}\|_{0,\Omega}.$$

On the other hand, estimate (3.21) from Theorem 3.4 yields

$$\left( \sum_{j=1}^{N_{\mathcal{F}}} |\boldsymbol{u} \cdot \boldsymbol{n}|_{1,\mathcal{F}_j}^2 \right)^{1/2} \lesssim \ell^{-3/2} \left( k^{1/2} \|\boldsymbol{f}\|_{0,\Omega} + k^{-3/2} \|\boldsymbol{\nabla} \times \boldsymbol{f}\|_{0,\Omega} \right).$$

It thus follows from (4.2), (4.3) and (4.4) that

$$\begin{aligned}
\|\boldsymbol{u} - \boldsymbol{\mathcal{I}}_h \boldsymbol{u}\|_{\boldsymbol{\mathcal{X}}} &\lesssim & kh|\boldsymbol{u}|_{1,\Omega} + k^{1/2} h \left( \sum_{j=1}^{N_{\mathcal{F}}} |\boldsymbol{u} \cdot \boldsymbol{n}|_{1,\mathcal{F}_j}^2 \right)^{1/2} + h|\boldsymbol{\nabla} \cdot \boldsymbol{u}|_{1,\Omega} \\
&\lesssim & kh \left( \|\boldsymbol{f}\|_{0,\Omega} + k^{-2} \|\boldsymbol{\nabla} \times \boldsymbol{f}\|_{0,\Omega} \right) \\
&+& k^{1/2} h \ell^{-3/2} \left( k^{1/2} \|\boldsymbol{f}\|_{0,\Omega} + k^{-3/2} \|\boldsymbol{\nabla} \times \boldsymbol{f}\|_{0,\Omega} \right) \\
&+& h \left( k \|\boldsymbol{f}\|_{0,\Omega} \right) \\
&\lesssim & \ell^{-3/2} kh \left( \|\boldsymbol{f}\|_{0,\Omega} + k^{-2} \|\boldsymbol{\nabla} \times \boldsymbol{f}\|_{0,\Omega} \right).
\end{aligned}$$

On the other hand, (4.5) yields that

$$k \|\boldsymbol{u} - \boldsymbol{\pi}_h \boldsymbol{u}\|_{0,\Omega} \lesssim kh|\boldsymbol{u}|_{1,\Omega} \lesssim kh(\|\boldsymbol{f}\|_{0,\Omega} + k^{-2} \|\boldsymbol{\nabla} \times \boldsymbol{f}\|_{0,\Omega}).$$

and (4.23) directly follows from (4.22). □

## 5. NUMERICAL EXPERIMENTS

5.1. **An analytical solution.** This example is designed so that the solution is a superposition of three plane waves that exhibit different amplitudes and travelling directions inside the unit square $\Omega = (0,1)^2$. Specifically, the solution reads

$$\boldsymbol{u}(\boldsymbol{x}) = \sum_{j=1}^{3} j \boldsymbol{d}_j e^{ik\boldsymbol{d}_j \cdot \boldsymbol{x}},$$

where $\boldsymbol{d}_j = (\cos(j\pi/8), \sin(j\pi/8))$. We can easily see that

$$(5.1) \qquad \begin{cases} -k^2 \boldsymbol{u} - \boldsymbol{\nabla}\boldsymbol{\nabla} \cdot \boldsymbol{u} &= 0 \quad \text{in } \Omega, \\ \boldsymbol{\nabla} \cdot \boldsymbol{u} - ik\boldsymbol{u} \cdot \boldsymbol{n} &= g \quad \text{on } \partial\Omega, \end{cases}$$

where $g \in L^2(\partial\Omega)$ is defined by

$$g(\boldsymbol{x}) = i\omega \sum_{j=1}^{3} j(\boldsymbol{n} \cdot \boldsymbol{d}_j - 1) e^{ik\boldsymbol{d}_j \cdot \boldsymbol{x}}.$$
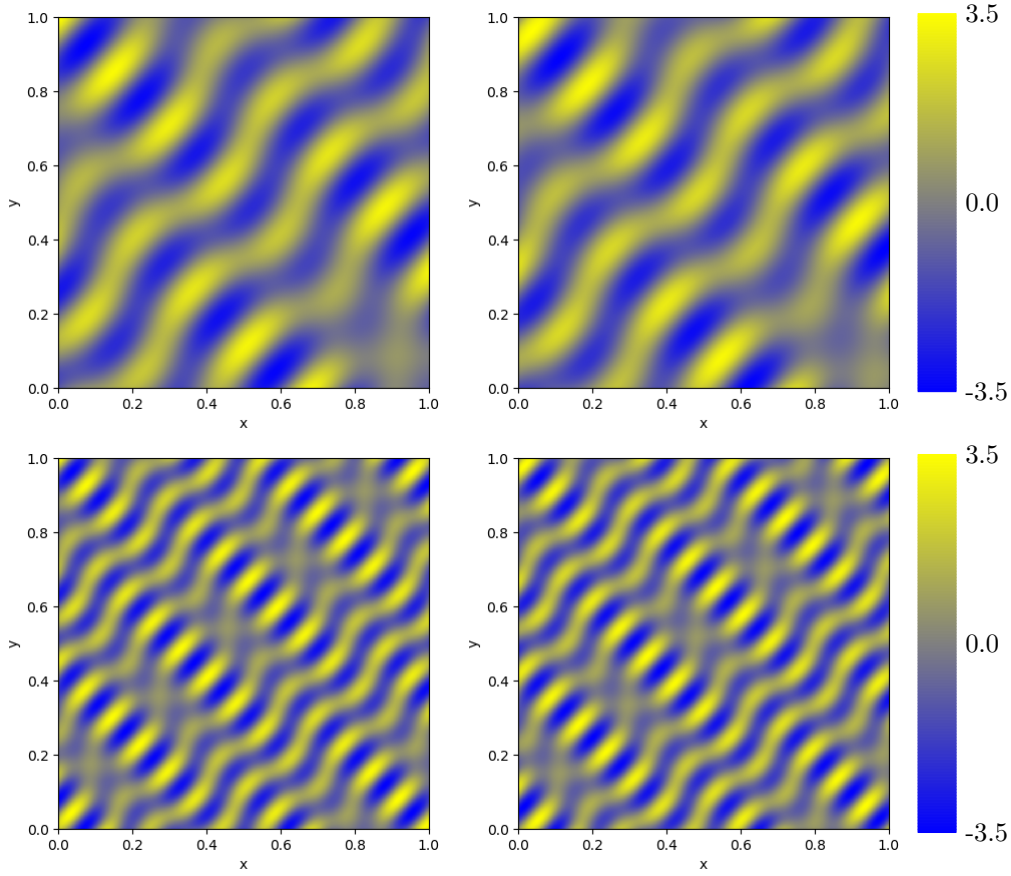
FIGURE 1. Real (left) and imaginary (right) parts of first component of $\boldsymbol{u}$ when $k = 10\pi$ (top) and $k = 20\pi$ (bottom)
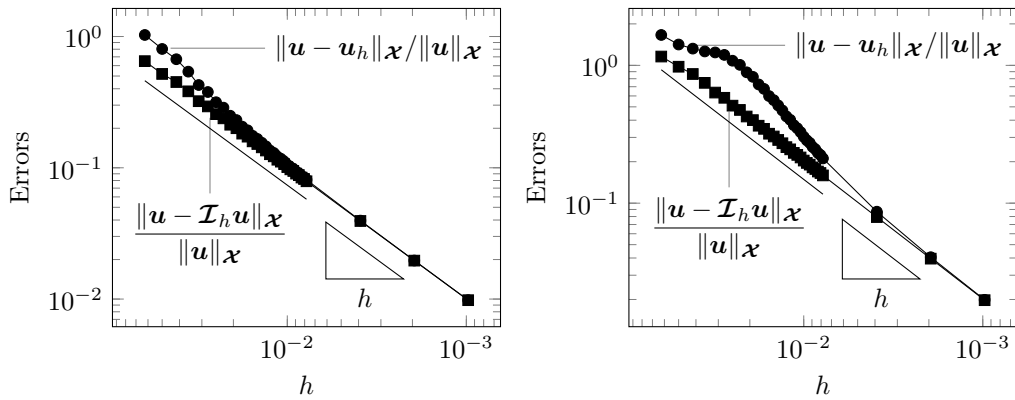


FIGURE 2. Convergence history for $k = 10\pi$ (left) and $k = 20\pi$ (right)

We thus see that $\boldsymbol{u}$ satisfies $b(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{\psi}(\boldsymbol{v})$ for all $\boldsymbol{v} \in \boldsymbol{\mathcal{X}}$, where $\boldsymbol{\psi} \in \boldsymbol{\mathcal{X}}'$ is the anti-linear mapping $\boldsymbol{\psi}(\boldsymbol{v}) = (g, \boldsymbol{v} \cdot \boldsymbol{n})_{\partial\Omega}$. Figure 1 represents the first component of $\boldsymbol{u}$ for two different wavenumbers.

We consider a set of wavenumbers $k$ ranging from $10\pi$ to $60\pi$. For each selected wavenumber, we solve (5.1) for different values of $h$, and record the convergence history of $\boldsymbol{u}_h$ and $\boldsymbol{\mathcal{I}}_h \boldsymbol{u}$ to $\boldsymbol{u}$ in the $\|\cdot\|_{\boldsymbol{\mathcal{X}}}$-norm. Figure 2 presents such convergence history for two wavenumbers.
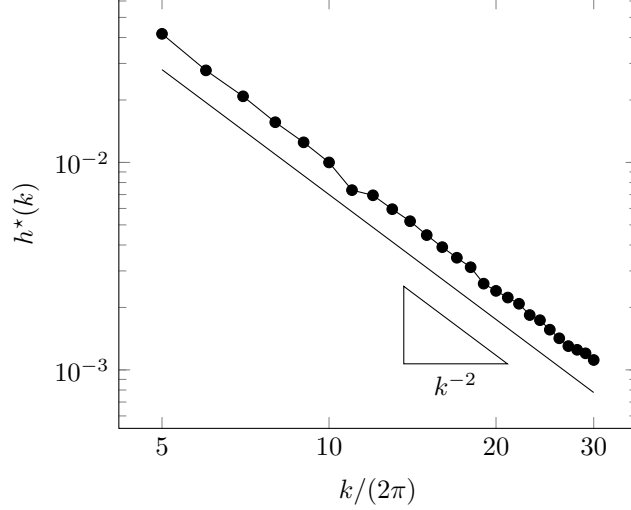
FIGURE 3. Asymptotic regime $h^\star(k)$

Figure 2 shows that the convergence is linear for both the interpolant and the finite element solution, as expected. In addition, the pollution effect is clearly visible. There is a gap between the interpolation error and the finite element error that decreases when $h$ decreases. This gap is more important for larger wavenumbers.

Next, we assess that the condition "$k^2 h$ is small enough" is necessary to achieve quasi-optimality. To do so, we browse the convergence history recorded for every $k$, and select $h^\star(k)$ as the largest mesh size such that

$$(5.2) \qquad \|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{\mathcal{X}}} \leq \frac{3}{2} \|\boldsymbol{u} - \boldsymbol{\mathcal{I}}_h \boldsymbol{u}\|_{\boldsymbol{\mathcal{X}}},$$

for all $h \leq h^\star(k)$. We arbitrarily selected the constant $3/2$ in (5.2), and another constant can be chosen, as long as it is independent of $h$ and $k$. The statement "$k^2 h$ is small enough" in Theorem 4.1 implies that $h^\star(k) \gtrsim k^{-2}$. As shown on Figure 3, we indeed observe that $h^\star(k) \simeq k^{-2}$, which shows that the stability condition of Theorem 4.1 is sharp.

5.2. **Scattering by a star-shaped obstacle.** Here, we consider a more realistic example that modelizes the scattering of a plane wave by a "sound hard" star-shaped obstacle [7]. Figure 4 describes the geometry of the problem. Specifically, we have

$$\Omega_0 = (-1, 1)^2, \quad \Gamma_A = \partial \Omega_0, \quad \Omega = \Omega_0 \setminus D,$$

and

$$D = \left\{ \boldsymbol{x} \in \mathbb{R}^2; \ 2|\boldsymbol{x}_1| - \frac{1}{2} < \boldsymbol{x}_2 < |\boldsymbol{x}_1| \right\}.$$

We modelize the following physical situation: a (known) incoming plane wave $p_i$ is diffracted by the obstacle $D$, generating a scattered field $p_s$. The total field $p = p_i + p_s$ satisfies the Helmholtz equation outside the obstacle

$$(5.3) \qquad \begin{cases} -k^2 p - \Delta p &= 0 \quad \text{in } \mathbb{R}^2 \setminus D, \\ \boldsymbol{\nabla} p \cdot \boldsymbol{n} &= 0 \quad \text{on } \partial D, \end{cases}$$

and the scattered field is subjected to the Sommerfeld radiation condition:

$$(5.4) \qquad \left| \boldsymbol{\nabla} p_s \cdot \frac{\boldsymbol{x}}{|\boldsymbol{x}|} - i k p_s \right| = \mathcal{O}(|\boldsymbol{x}|^{-1/2})$$

as $|\boldsymbol{x}| \to +\infty$.

We approximate the Sommerfeld radiation condition by a first-order absorbing boundary condition to simulate Problem (5.3)-(5.4) using finite element methods [12, 13]. Specifically, we replace
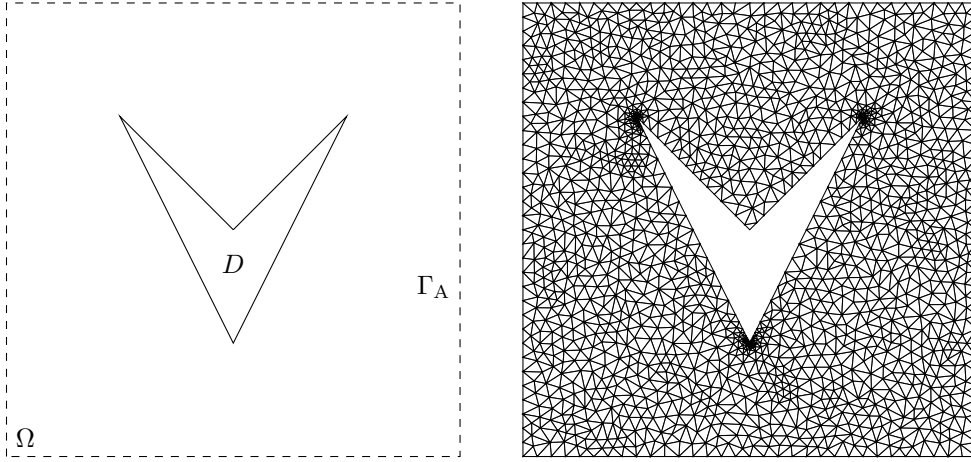
FIGURE 4. Geometry of the obstacle (left) and example of locally refined mesh (right)

(5.4) by the local boundary condition that $\boldsymbol{\nabla} p_s \cdot \boldsymbol{n} - ikp_s = 0$ on $\Gamma_{\mathrm{A}}$, where the boundary $\Gamma_{\mathrm{A}}$ is artificially introduced. As a result, we introduce the problem

$$(5.5) \qquad \begin{cases} -k^2 p - \Delta p &= 0 & \text{in } \Omega, \\ \boldsymbol{\nabla} p \cdot \boldsymbol{n} &= 0 & \text{on } \partial D, \\ \boldsymbol{\nabla} p \cdot \boldsymbol{n} - ikp &= -\boldsymbol{\nabla} p_i \cdot \boldsymbol{n} + ikp_i & \text{on } \Gamma_{\mathrm{A}}, \end{cases}$$

where $\Omega$ denotes the exterior region of $D$ enclosed by $\Gamma_{\mathrm{A}}$. We discretize (5.5) with Lagrange finite elements (see [6, 22] for instance).

Another approach to approximate (5.5) is to introduce the velocity $\boldsymbol{u} = \boldsymbol{\nabla} p$. Then, since $\boldsymbol{\nabla} \cdot \boldsymbol{u} = \Delta p = k^2 p$ in $\mathbb{R}^2 \setminus D$, we obtain the equivalent formulation

$$(5.6) \qquad \begin{cases} -k^2 \boldsymbol{u} - \boldsymbol{\nabla} \cdot \boldsymbol{\nabla} \boldsymbol{u} &= 0 & \text{in } \Omega, \\ \boldsymbol{u} \cdot \boldsymbol{n} &= 0 & \text{on } \partial D, \\ \boldsymbol{\nabla} \cdot \boldsymbol{u} - ik\boldsymbol{u} \cdot \boldsymbol{n} &= -\Delta p_i + ik\boldsymbol{\nabla} p_i \cdot \boldsymbol{n} & \text{on } \Gamma_{\mathrm{A}}. \end{cases}$$

Although formulation (5.6) does not exactly fits the framework analyzed in this work because of obstacle boundary condition, it is approximated here using Raviart-Thomas finite elements. In the pressure formulation, we immediately obtain an approximations $p_h$ and $\boldsymbol{\nabla} p_h$ to $p$ and $\boldsymbol{\nabla} p$ in a natural way. For the velocity formulation, we set $\boldsymbol{\nabla} p_h = \boldsymbol{u}_h$, and $p_h = -(1/k^2)\boldsymbol{\nabla} \cdot \boldsymbol{u}_h$.

To compare pressure formulation (5.5) and (5.6), we perform simulations for the incoming wave $p_i = \exp(ik\boldsymbol{d} \cdot \boldsymbol{x})$ with $\boldsymbol{d} = (\cos(\pi/12), \sin(\pi/12))$ for different wavenumbers $k$ and meshes. We employ meshes that are refined close to the corners of the obstacle, as the solution may exhibit a singular behaviour in these area [3, 18]. Figure 4 represents an example of such locally refined mesh. Figure 5 depicts the total fields (computed using a very fine mesh for comparison purposes) for different wavenumbers.

Figures 6 and 7 depict the convergence history of the finite element solutions to the reference solution, where we observe the expected converge rates. Figures 8 and 9 show a comparison of accuracy provided by the two discretization techniques in terms of the size of the corresponding linear systems $N_{dof}$ and the number of non-zero element in the matrix $N_{nz}$. In these figures, the "pressure" curves correspond to the results obtained with the pressure formulation, while the "velocity" curves are obtained using the velocity formulation.

Interestingly, in the high wavenumber regime, Figure 9 shows that if a relative accuracy of 10% is required, the velocity formulation provides an acceptable solution with sightly less degrees of freedom and a slightly sparser linear system than the pressure formulation.

Overall, if the quantity of interest is the velocity $\boldsymbol{u} = \boldsymbol{\nabla} p$, the velocity and pressure formulation provides a comparable accuracy (for comparable linear system characteristics) in the $\| \cdot \|_{0,\Omega}$-norm. The velocity formulation then presents the advantage that the discrete field is divergence-conforming, without requiring a post-processing step. This experiment thus indicates that the
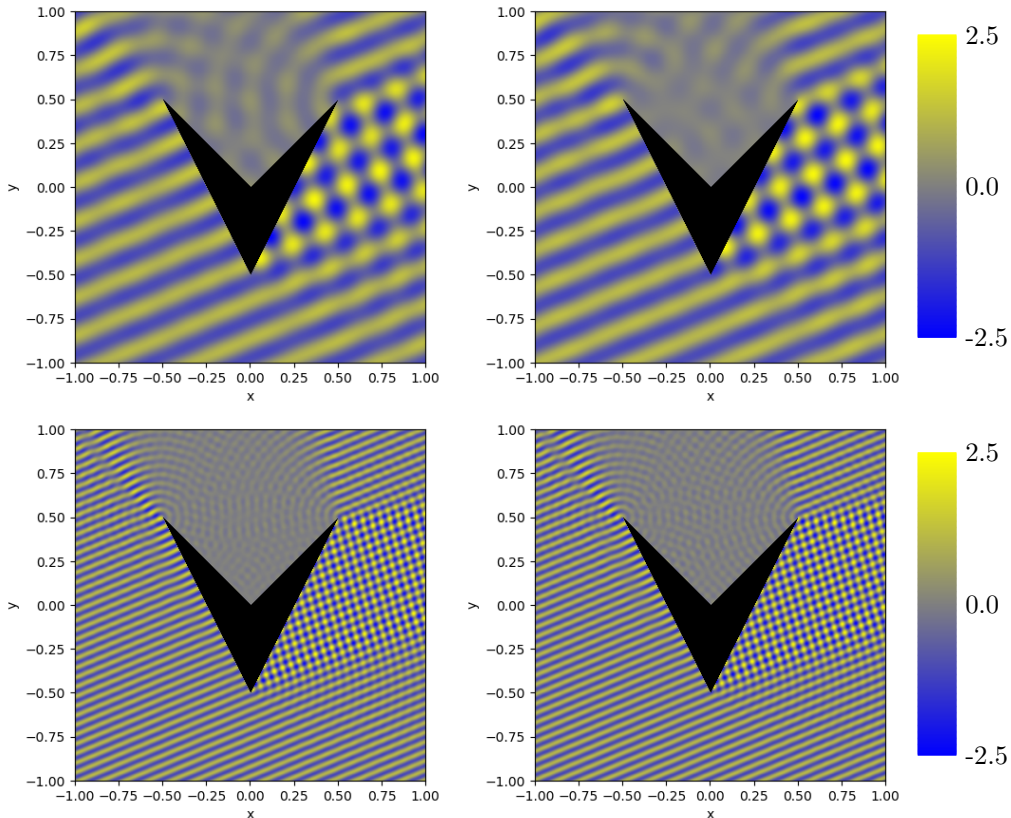
FIGURE 5. Real (left) and imaginary (right) parts of the total field $p$ when $k = 10\pi$ (top) and $k = 40\pi$ (bottom)

velocity formulation is viable alternative to the pressure formulation, even in simple configurations where the pressure formulation easily applies (irrotational right-hand sides, and so on).

## 6. CONCLUSIONS

We analyzed the "velocity formulation" of the acoustic Helmholtz equation set in a convex polytopal domain with impedance boundary conditions. This formulation requires the introduction of a tailored Sobolev space $\boldsymbol{\mathcal{X}}$ to properly handle the boundary conditions. We proved that the problem is well-posed and derived wavenumber-explicit stability results in the $\boldsymbol{\mathcal{X}}$-norm, that includes the $\boldsymbol{H}(\mathrm{div}, \Omega)$-norm and the $L^2(\partial\Omega)$-norm of the normal trace.

We considered a finite element discretization of the problem based on lowest-order Raviart-Thomas finite elements. We established that the scheme is well-posed and stable under the condition that $k^2h$ is sufficiently small, where $k$ and $h$ respectively stand for the wavenumber and the mesh size. Under this condition, we showed that the finite element error is "almost" quasi-optimal, and provided an error estimate in $\mathcal{O}(kh)$ assuming the right-hand side is sufficiently regular. These stability condition and error estimate are well-known for the pressure formulation discretized with first-order Lagrange finite elements, and we thus proved that they are also valid for the velocity formulation.

We presented numerical experiments to illustrate our findings. These examples show that indeed, the finite element scheme suffers from the pollution effect in the high-wavenumber regime. In addition, a thorough investigation on the stability of the numerical scheme revealed that our abstract analysis is sharp. Our key conclusion is thus that Raviart-Thomas discretizations are stable if and only if $k^2h$ is sufficiently small, which extend the well-known result for Lagrange discretizations. We also provided a quick comparison of the pressure and velocity formulations
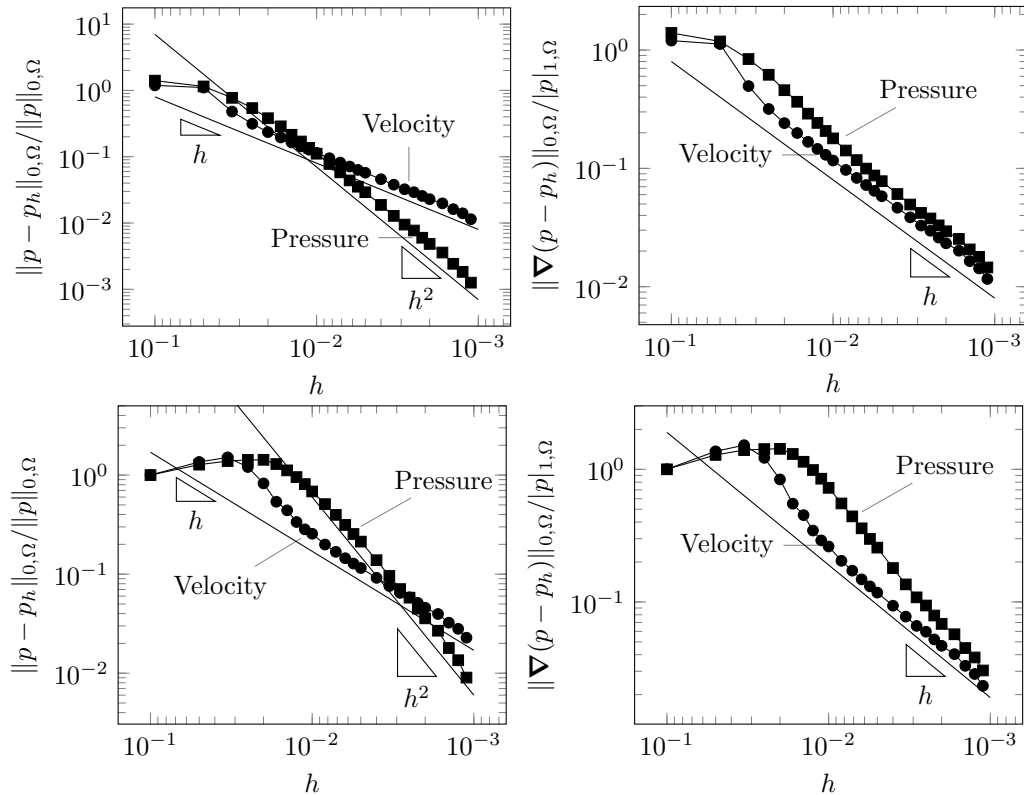
FIGURE 6. Convergence history in the scattering experiment for $k = 10\pi$ (top) and $k = 20\pi$ (bottom)

that indicates that the velocity formulation is a competitive alternative to the pressure formulation in terms of computational costs.

## REFERENCES

1. R.A. Adams and J.J. Fournier, *Sobolev spaces*, 2nd ed., Academic Press, 2003.
2. I. Babuška and S.A. Sauter, *Is the pollution effect of the FEM avoidable for the Helmholtz equation considering high wave numbers?*, SIAM J. Numer. Anal. **34** (1997), no. 6, 2392–2423.
3. T. Chaumont-Frelet and S. Nicaise, *High-frequency behaviour of corner singularities in Helmholtz problems*, ESAIM Math. Model. Numer. Anal. **5** (2018), 1803–1845.
4. _____, *Wavenumber explicit convergence analysis for finite element discretizations of general wave propagation problems*, IMA J. Numer. Anal., accepted (2019).
5. T. Chaumont-Frelet, S. Nicaise, and D. Pardo, *Finite element approximation of electromagnetic fields using nonfitting meshes for Geophysics*, SIAM J. Numer. Anal. **56** (2018), no. 4, 2288–2321.
6. P.G. Ciarlet, *The finite element method for elliptic problems*, SIAM, 2002.
7. D. Colton and R. Kress, *Inverse acoustic and electromagnetic scattering theory*, Springer, 2012.
8. _____, *Integral equation methods in scattering theory*, SIAM, 2013.
9. M. Costabel, *A remark on the regularity of solutions of Maxwell's equations on Lipschitz domains*, Math. Method Appl. Sci. **12** (1990), 365–368.
10. M. Dauge, *Elliptic bounday balue problems on corner domains*, Springer-Verlag, 1988.
11. A. Bonnet-Ben Dhia, E. Duclairoir, G. Legendre, and J. Mercier, *Time-harmonic acoustic propagation in the presence of a shear flow*, J. Comp. Appl. Math. **204** (2007), 428–439.
12. J. Diaz, *Approches analytiques et numériques de problèmes de transmission en propagation dondes en rgime transitoire. application au couplage fluide-structure et aux méthodes de couches parfaitement adaptées*, Ph.D. thesis, Paris 6, 2005.
13. B. Engquist and A. Majda, *Absorbing boundary conditions for the numerical simulation of waves*, Math. Comp. **31** (1977), no. 139, 629–651.
14. A. Ern and J.L. Guermond, *Mollification in strongly Lipshitz domains with application to continuous and discrete De Rham complexes*, Comput. Meth. Appl. Math. **16** (2016), no. 1, 51–75.
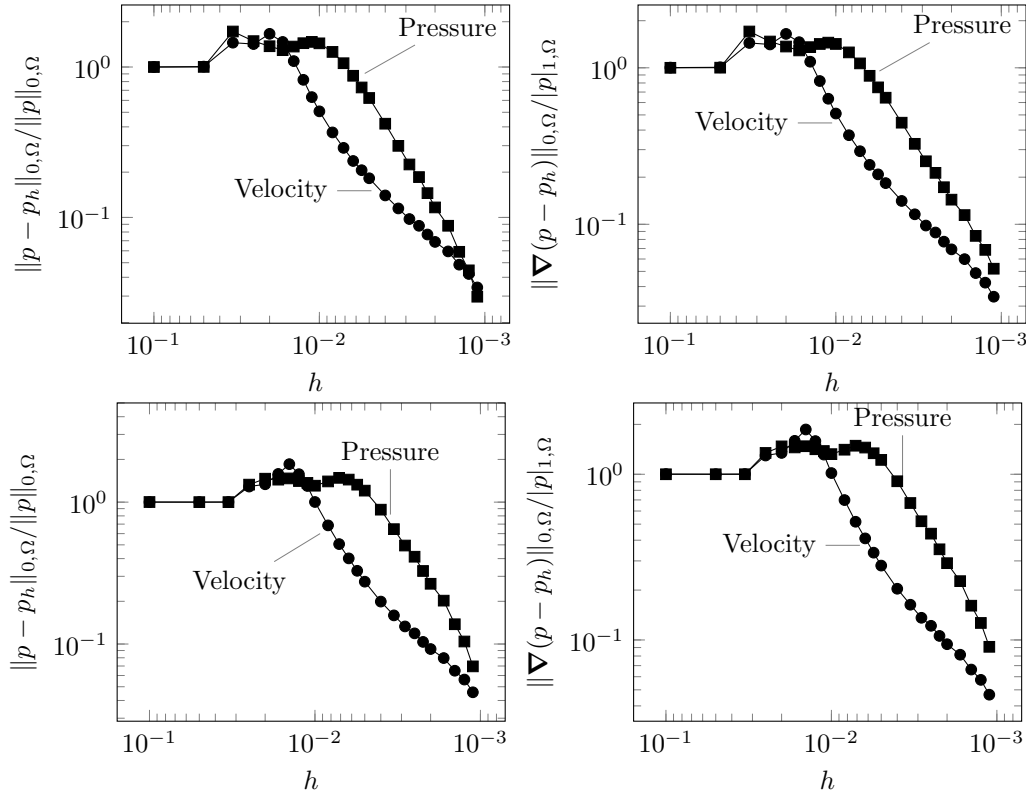
FIGURE 7. Convergence history in the scattering experiment for $k = 30\pi$ (top) and $k = 40\pi$ (bottom)

15. _____, *Analysis of the edge finite element approximation of the maxwell equations with low regularity solutions*, Comp. Math. Appl. **75** (2018), 918–932.

16. L. Gastaldi, *Mixed finite elemnet methods in fluid structure systems*, Numer. Math. **74** (1996), 153–176.

17. V. Girault and P.A. Raviart, *Finite element methods for Navier-Stokes equations*, Springer-Verlag, 1986.

18. P. Grisvard, *Elliptic problems in nonsmooth domains*, Pitman, 1985.

19. U. Hetmaniuk, *Stability estimates for a class of Helmholtz problems*, Commun. Math. Sci. **5** (2007), no. 3, 665–678.

20. F. Ihlenburg and I. Babuška, *Finite element solution of the Helmholtz equation with high wave number. Part I: The h-version of the FEM*, Comp. Math. Appl. **30** (1995), no. 9, 9–37.

21. J.M. Melenk and S. Sauter, *Wavenumber explicit convergence analysis for Galerkin discretizations of the Helmholtz equation*, SIAM J. Numer. Anal. **49** (2011), no. 3, 1210–1243.

22. P. Monk, *Finite element methods for Maxwell's equations*, Oxford science publications, 2003.

23. J.C. Nédélec, *Mixed finite elements in $\mathbb{R}^3$*, Numer. Math. **35** (1980), 315–341.

24. L.E. Payne and H.F. Weinberger, *An optimal Poincaré inequality for convex domains*, Arch. Ration. Mech. Anal. **5** (1960), no. 1, 286–292.

25. P.A. Raviart and J.M. Thomas, *A mixed finite element method for 2nd order elliptic problems*, Mathematical Aspect of Finite Element Methods, Springer-Verlag, 1977.

26. S. Retka and S. Marburg, *An infinite element for the solution of Galbrun equation*, Z. Angew. Math. Mech. **93** (2013), no. 2–3, 154–162.

27. A.H. Schatz, *An observation concerning Ritz-Galerkin methods with indefinite bilinear forms*, Math. Comp. **28** (1974), no. 128, 959–962.

28. J. Schoberl, *Commuting quasi-interpolation operators for mixed finite elements*, Tech. Report Preprint ISC-01-10-MATH, Texas A& M Univertsity, 2001.

29. I. Singer and E. Turkel, *High-order finite difference methods for the Helmholtz equation*, Comput. Meth. Appl. Mech. Engin. **163** (1998), 343–358.

30. L. Tartar, *An introduction to Sobolev spaces and interpolation spaces*, Springer, 2007.

31. X. Wang and K. Bathe, *On mixed elements for acoustic fluid-structure interactions*, Math. Models Meth. Appl. Sci. **7** (1997), no. 3, 329–343.
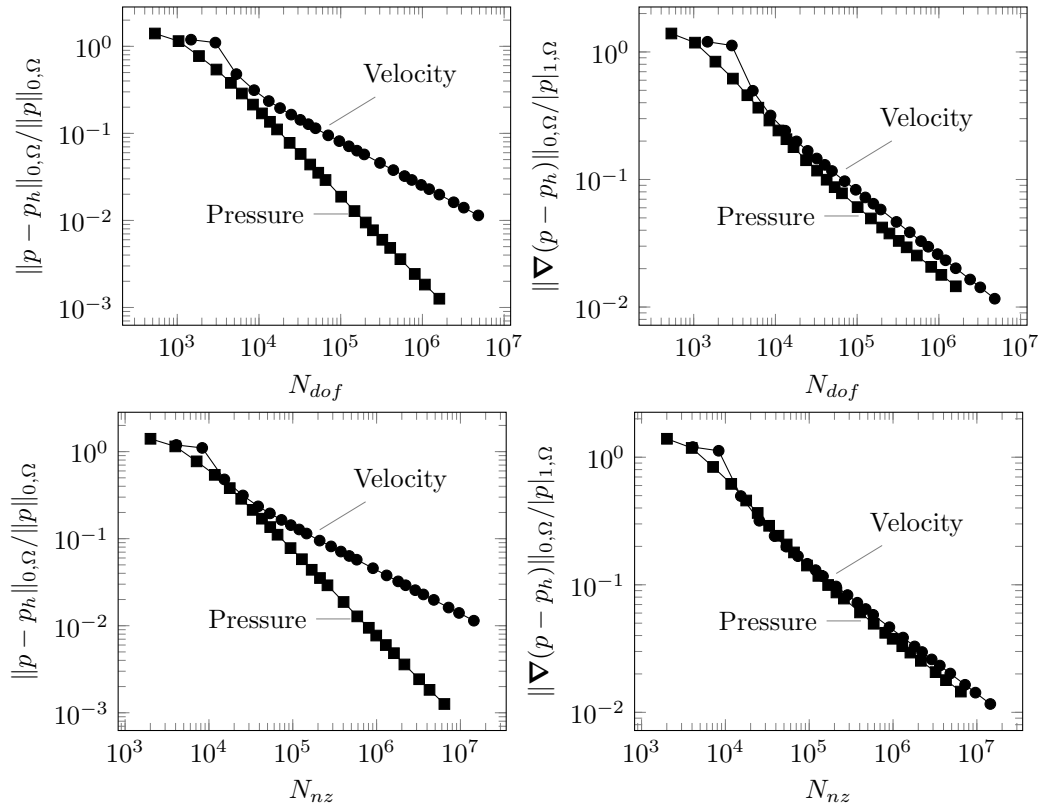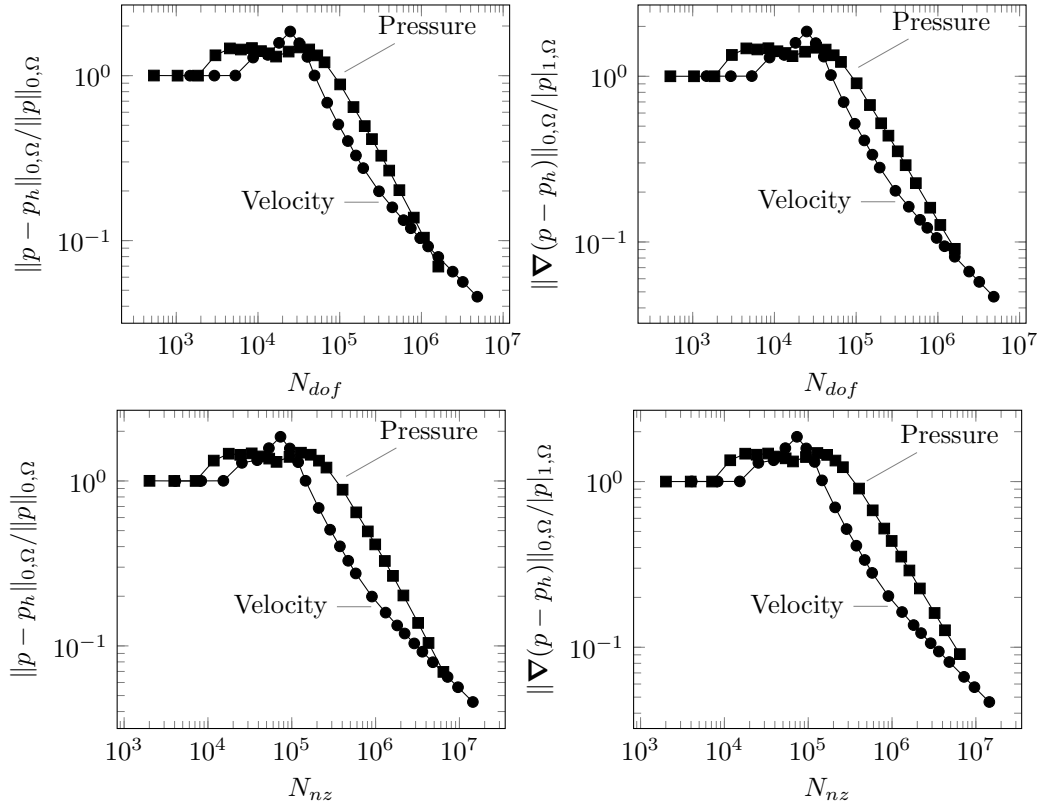
FIGURE 8. Characteristics of the finite element linear systems for $k = 10\pi$

32. L. Zhong, S. Shu, G. Wittum, and J. Xu, *Optimal error estimtaes for nédélec edge elements for time-harmonic Maxwell's equations*, J. Comp. Math. **27** (2009), no. 5, 563–572.

FIGURE 9. Characteristics of the finite element linear systems for $k = 40\pi$