Error analysis of the DtN-FEM for the scattering problem in acoustics via Fourier analysis

George C. Hsiao[a], Nilima Nigam[b], Joseph E. Pasciak[c], Liwei Xu[d,*]

[a]*Department of Mathematical Sciences, University of Delaware, Newark, DE 19716 - 2553, United States*
[b]*Department of Mathematics, Simon Fraser University, Burnaby, BC, V5A 1S6, Canada*
[c]*Department of Mathematics, Texas A&M University, College Station, TX 77843-3368, United States*
[d]*Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180-3590, United States*

Abstract

In this paper, we are concerned with the error analysis for the finite element solution of the two-dimensional exterior Neumann boundary value problem in acoustics. In particular, we establish an explicit priori error estimates in H^1 and L^2- norms including both the effect of the truncation of the DtN mapping and that of the numerical discretization. To apply the finite element method (FEM) to the exterior problem, the original boundary value problem is reduced to an equivalent nonlocal boundary value problem via a Dirichlet-to-Neumann (DtN) mapping represented in terms of the Fourier expansion series. We discuss essential features of the corresponding variational equation and its modification due to the truncation of the DtN mapping in appropriate function spaces. Numerical tests are presented to validate our theoretical results.

Keywords: Dirichlet-to-Neumann mapping, Finite element method, Acoustic scattering problem, Error analysis

---

[*]Corresponding author.
Email addresses: `hsiao@math.udel.edu` (George C. Hsiao), `nigam@math.sfu.ca` (Nilima Nigam), `pasciak@math.tamu.edu` (Joseph E. Pasciak), `xul3@rpi.edu` (Liwei Xu)

1. Introduction

Numerical solutions of the scattering of time-harmonic acoustic waves by an impenetrable bounded obstacle have been a subject of scientific investigation for many years. It entails considerable mathematical and computational challenges such as the oscillating character of solutions, and the unbounded domain to be considered. Among the most conventional numerical methods addressing the latter difficulty are the boundary integral equation methods and the coupled finite element methods (FEM). In the application of the coupled FEM, a popular way is to decompose the unbounded domain by introducing an artificial boundary enclosing the obstacle inside. Then, appropriate methods are used to solve the exterior problem outside the artificial boundary while finite element methods are employed for the solution of the Helmholtz equation on the bounded domain between the scatterer and the artificial boundary. There are several techniques ([1, 2, 3, 4, 5]) to realize such coupling procedure based on the above domain decomposition scheme, and one of them is to enforce a nonlocal boundary condition on the artificial boundary curve via deriving a Dirichlet-to-Neumann (DtN) mapping. Therefore, the exterior problem is reduced to a nonlocal boundary value problem, and accordingly such coupled FEM ([6]) is called the DtN finite element method (DtN-FEM). There are several methods used for the derivation of such DtN mapping. Defining the DtN mapping through basic boundary integral operators ([2]) gives the coupling of FEM and boundary element method (BEM), and representing the DtN mapping in terms of Fourier expansion series leads to the coupling of FEM and the method of separation of variables ([3, 4, 7, 8]). The present article is designed to make contributions to the error analysis in the latter application. As an extension of the standard DtN-FEM, authors of [9] developed a new method for the realization of exact non-reflecting boundary conditions without the restriction on the shape of artificial boundary, and carried out a sequence of works ([10, 11, 12]) on the numerical analysis and computation. Corresponding to the nonlocal boundary conditions, there are several type of local boundary conditions ([13, 14, 15]).

In essential, local boundary conditions are kinds of approximate boundary conditions, and the simplest local boundary condition can be obtained simply by employing the *Sommerfeld* condition on the artificial boundary.

In [16], the authors derived error estimates including effects of finite element discretization and series truncation for the exterior Laplace problem. Koyama ([17]) applied the analysis introduced in [18, 19] to consider both errors for the Helmholtz equation. In this work, we apply the strategy in [16], the analysis techniques ([20, 19]) originally developed for BEM, and the standard finite element analysis ([21]) to derive a more evident and concise priori error estimates. In addition, we perform a sequence of numerical tests to validate our theoretical results. To be more precise, we first write out explicitly a modified variational equation which is the result of replacing the exact DtN mapping with the truncated DtN mapping in the exact variational equation, and then show that such modified variational equation satisfies a Gårding's inequality and admits a unique weak solution. These two features allow us to establish the inf-sup condition ([20]) provided the finite element space satisfies the approximation property. Some analysis techniques in our presentation are closely related with the Schatz argument originally introduced in [19] and the later work [22] for the analysis of Ritz-Galerkin methods for indefinite bilinear forms. Starting from the inf-sup condition, we finally succeed in deriving a priori error estimates in H^1 and L^2-norms including the effects of both the discretization error and the truncation error. Finally, as a result of theoretical analysis and numerical tests, we report a chart reflecting the interaction of numerical parameters for the solution of exterior acoustic scattering problems.

The remainder of the paper is organized as follows. We first describe the classical Helmholtz exterior problem, and then reduce the exterior problem to a nonlocal boundary value problm in Section 3. In Section 4, we discuss essential mathematical features for the corresponding variational equation of the nonlocal boundary value problem, and its modification due to the truncation of the DtN mapping. In Section 5, we establish a priori error estimates for the Galerkin solution. Finally, Section 6 presents several numerical tests to confirm our

theoretical results.

2. Statement of the problem

Let Ω denote a bounded domain with smooth boundary Γ, and let $\Omega^c = \mathbb{R}^2 \backslash \overline{\Omega}$ be the unbounded exterior domain in $\mathbb{R}^2$ (see Figure 1 (left)). We consider the following problem in acoustics: *Given $\partial u^i / \partial n$, find $u(x) \in C^2(\Omega^c) \cap C^1(\overline{\Omega^c})$ satisfying*

$$\Delta u + k^2 u = 0 \quad \text{in} \quad \Omega^c, \tag{1}$$

$$\frac{\partial u}{\partial n} = -\frac{\partial u^i}{\partial n} \quad \text{on} \quad \Gamma. \tag{2}$$

In the above formulation, $k \neq 0$ is the wave number with $Im(k) \geq 0$, $u = u^s$ denotes the scattering field, and u^i the given incident field; $\partial/\partial n$ means the normal derivative on Γ (here and in the sequel, n is always the outer unit normal to the boundary). For the uniqueness, in addition, the scattering field u is required to satisfy the standard *Sommerfeld* radiation condition

$$\lim_{r \to \infty} r^{\frac{1}{2}} \left(\frac{\partial u}{\partial r} - iku \right) = 0, \tag{3}$$

where $i = \sqrt{-1}$, $r = |x|$ and $x = (x_1, x_2) \in \mathbb{R}^2$. We term the boundary-value problem (1)–(3) as the exterior Neumann problem in acoustics.

We state without proofs of the following uniqueness theorem, and a proof can be found in [23].

Theorem 2.1. *The exterior boundary value problem* (1)–(3) *has at most one solution.*

Prior to our discussion, we introduce the relevant *Sobolev* spaces ([20, 24]). Suppose Ω to be an open subset of $\mathbb{R}^2$ with smooth boundary Γ. Let $L^2(\Omega)$ be the function space consisting of all square integrable functions over Ω equipped with the norm

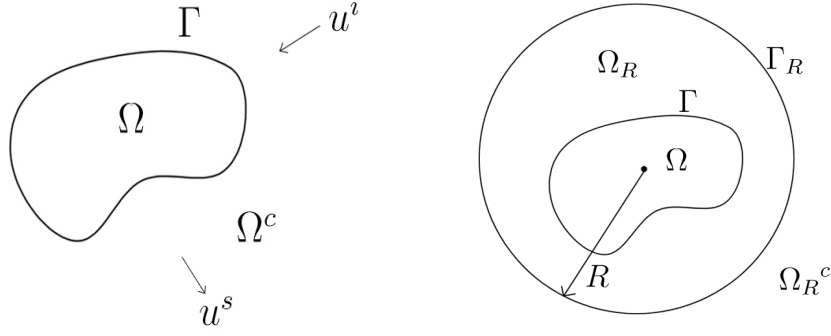$$\|v\|_{L^2(\Omega)} = \left(\int_\Omega |v(x)|^2 dx \right)^{1/2}.$$

4

Figure 1: Boundary value problem (1)–(3) (left); nonlocal boundary value problem (5)–(7)(right).

We denote by $H^1(\Omega)$ the *Sobolev* space

$$H^1(\Omega) = \{v \in L^2(\Omega) | \nabla v \in L^2(\Omega)\}$$

equipped with the norm

$$\|v\|_{H^1(\Omega)} = (\int_\Omega |v(x)|^2 + |\nabla v(x)|^2 dx)^{1/2}.$$

In particular, we have $H^0(\Omega) = L^2(\Omega)$. We denote by $(H^1(\Omega))'$ the dual space of $H^1(\Omega)$ equipped with the norm

$$\|f\|_{(H^1(\Omega))'} = \sup_{0 \neq v \in H^1(\Omega)} \frac{\langle f, v \rangle_\Omega}{\|v\|_{H^1(\Omega)}},$$

where $\langle \cdot, \cdot \rangle_\Omega$ stands for the standard L^2 duality pairing between $(H^1(\Omega))'$ and $H^1(\Omega)$. Let $L^2(\Gamma)$ be the space of all square integrable functions v on Γ equipped with the norm

$$\|v\|_{L^2(\Gamma)} = (\int_\Gamma |v(x)|^2 dx)^{1/2}.$$

We define by $H^s(\Gamma)$, $\forall s \in \mathbb{R}$, the *Sobolev* space

$$H^s(\Gamma) = \{v \in L^2(\Gamma) | \|v\|_{H^s(\Gamma)} < \infty\}$$

equipped with the norm

$$\|v\|_{H^s(\Gamma)}^2 = \frac{|a_0|^2}{2} + \sum_{n=1}^{\infty} (1 + n^2)^s (|a_n|^2 + |b_n|^2), \tag{4}$$

5

where a_n and b_n are Fourier coefficients of v.

3. Nonlocal boundary value problem

We introduce an artificial circular boundary Γ_R of radius R (see Figure 1 (right)) which is large enough to enclose the entire region Ω. The artificial boundary decomposes the exterior domain Ω^c into two subdomains denoted by Ω_R and ${\Omega_R}^c$ respectively, where Ω_R is the annular region between Γ and Γ_R, and ${\Omega_R}^c = \mathbb{R}^2 \setminus \overline{\Omega \cup \Omega_R}$ the unbounded exterior region. The boundary value problem (1)–(3) can be equivalently replaced by the following nonlocal boundary value problem: *Given $\partial u^i / \partial n$, find $u(x) \in C^2(\Omega_R) \cap C^1(\overline{\Omega_R})$ such that*

$$\Delta u + k^2 u \;=\; 0 \quad \text{in} \quad \Omega_R, \tag{5}$$

$$\frac{\partial u}{\partial n} \;=\; -\frac{\partial u^i}{\partial n} \quad \text{on} \quad \Gamma, \tag{6}$$

$$\frac{\partial u}{\partial n} \;=\; Tu \quad \text{on} \quad \Gamma_R. \tag{7}$$

Here, the DtN mapping $T : H^s(\Gamma_R) \mapsto H^{s-1}(\Gamma_R)$, for $\forall\, \varphi \in H^s(\Gamma_R)$, $1/2 \leq s \in \mathbb{R}$, is defined as

$$T\varphi := \sum_{n=0}^{\infty}{}' \frac{k{H_n^{(1)}}'(kR)}{\pi H_n^{(1)}(kR)} \int_0^{2\pi} \varphi(R, \phi) \cos(n(\theta - \phi)) d\phi \;. \tag{8}$$

Here and throughout the presentation, the prime $'$ behind the summation means that the first term in the summation is multiplied by $1/2$. Condition (7) on Γ_R in terms of the DtN mapping T also defines a nonlocal boundary condition for u on Γ_R since the Dirichlet data u over the entire boundary Γ_R are required to compute the Neumann data $\partial u / \partial n$ at a single point $x \in \Gamma_R$. Prior to the discussion of mapping properties for T, we point out some properties for the Hankel function $H_n^{(1)}(\cdot)$ in the next two Lemmas ([25]). To simplify the presentation throughout the dissertation, we shall denote by $c > 0$ a generic constant whose precise value is not required and may change line by line.

Lemma 3.1. *There exists a positive constant c such that*

$$\left| \frac{H_{n-1}^{(1)}(z)}{H_n^{(1)}(z)} \right| \leq c, \quad \forall\, n \in \mathbb{Z}, \tag{9}$$

where the constant c is dependent on the argument z but independent of n.

Lemma 3.2. *There exists a positive constant c such that*

$$\frac{1}{1+|n|}\left|\frac{H_n^{(1)'}(z)}{H_n^{(1)}(z)}\right| \le \frac{1}{(1+n^2)^{\frac{1}{2}}}\left|\frac{H_n^{(1)'}(z)}{H_n^{(1)}(z)}\right| \le c, \quad \forall\, n \in \mathbb{Z}, \qquad (10)$$

where the constant c is dependent on the argument z but independent of n.

Theorem 3.1. *The DtN mapping T in (8) is a bounded linear operator from $H^s(\Gamma_R)$ to $H^{s-1}(\Gamma_R)$ for any constants $s \ge \frac{1}{2}$.*

Proof: For the convenience of proof, we expand the function φ into the Fourier series

$$\varphi(R,\theta) = \sum_{n\in\mathbb{Z}} a_n H_n^{(1)}(kR)e^{in\theta} = \sum_{n\in\mathbb{Z}} \varphi_n e^{in\theta}$$

for $\forall\, \varphi \in H^s(\Gamma_R)$, $s \ge 1/2$. Here, φ_n is defined as

$$\varphi_n = a_n H_n^{(1)}(kR) = \frac{1}{2\pi}\int_0^{2\pi} \varphi(R,\phi)e^{-in\phi}d\phi\ .$$

As a consequence, we have an equivalent form of (8)

$$T\varphi := \sum_{n\in\mathbb{Z}} ka_n H_n^{(1)'}(kR)e^{in\theta} = \sum_{n\in\mathbb{Z}} k\varphi_n \frac{H_n^{(1)'}(kR)}{H_n^{(1)}(kR)}e^{in\theta}\ .$$

Now, we use an alternative of (4)

$$\|v\|_{H^s(\Gamma_R)}^2 = \sum_{n\in\mathbb{Z}}(1+n^2)^s|v_n|^2\ , \quad \forall\, s \in \mathbb{R}$$

for $\forall\, v(R,\theta) \in H^s(\Gamma_R)$. Therefore, we have, by Lemma 3.2,

$$\begin{aligned}
\|T\varphi\|_{H^{s-1}(\Gamma_R)}^2 &= \sum_{n\in\mathbb{Z}}(1+n^2)^s|\varphi_n|^2 \frac{|k|^2}{1+n^2}\left|\frac{H_n^{(1)'}(kR)}{H_n^{(1)}(kR)}\right|^2 \\
&\le c\sum_{n\in\mathbb{Z}}(1+n^2)^s|\varphi_n|^2 = c\|\varphi\|_{H^s(\Gamma_R)}^2
\end{aligned}$$

for $\forall\, s \ge 1/2$, and this leads to

$$\|T\varphi\|_{H^{s-1}(\Gamma_R)} \le c\|\varphi\|_{H^s(\Gamma_R)}.$$

7

Here $c > 0$ is a constant dependent on kR but independent of φ. This completes the proof.

The following uniqueness for the nonlocal boundary value problem (5)–(7) can be easily established.

Theorem 3.2. *The nonlocal boundary value problem (5)–(7) has at most one solution.*

Proof: It is sufficient to prove that the corresponding homogeneous boundary value problem of (5)–(7) has only the trivial solution. Suppose u_0 is a solution of the corresponding homogeneous boundary value problem of (5)–(7). Now let u_1 be the solution of the exterior Dirichlet problem for the Helmholtz equation:

$$\Delta u_1 + k^2 u_1 = 0, \quad \text{in} \quad \Omega_R{}^c, \tag{11}$$

$$u_1 = u_0, \quad \text{on} \quad \Gamma_R, \tag{12}$$

$$\lim_{r \to \infty} r^{\frac{1}{2}} \left(\frac{\partial u_1}{\partial r} - iku_1 \right) = 0. \tag{13}$$

Then u_1 has the representation in the form

$$u_1(r, \theta) = \sum_{n \in \mathbb{Z}} a_n H_n^{(1)}(kr) e^{in\theta}, \quad \forall r \geq R. \tag{14}$$

Computing the normal derivative for (14) and taking the limit as $r \to R$, we obtain, on Γ_R,

$$
\begin{aligned}
\frac{\partial u_1}{\partial n} &= \sum_{n \in \mathbb{Z}} \frac{k H_n^{(1)'}(kR)}{2\pi H_n^{(1)}(kR)} \int_0^{2\pi} u_1(R, \phi) e^{in(\theta - \phi)} d\phi \\
&= T u_1 \\
&= T u_0
\end{aligned} \tag{15}
$$

because of the boundary condition (12). In the mean time, the nonlocal boundary condition (7) gives

$$\frac{\partial u_0}{\partial n} = T u_0 \quad \text{on} \quad \Gamma_R. \tag{16}$$

Therefore, we have

$$\frac{\partial u_1}{\partial n} - \frac{\partial u_0}{\partial n} = T u_0 - T u_0 = 0, \quad \text{on} \quad \Gamma_R. \tag{17}$$

8

If we define the function $u \in C^2(\Omega_R \cup \Omega_R{}^c) \cap C^1(\overline{\Omega_R})$ as

$$
u = \begin{cases} u_0, & x \in \Omega_R, \\[2mm] u_1, & x \in \Omega_R{}^c, \end{cases}
$$

then by (12) and (17), we can see that both u and $\partial u / \partial n$ are continuous across the interface Γ_R. Therefore, u is the solution of the homogeneous transmission problem which is equivalent to the corresponding homogeneous boundary value problem of (1)–(3). The latter has been proved to be uniquely solvable. That leads to

$$
u \equiv 0 \quad \text{in} \quad \Omega^c, \quad \Rightarrow \quad u_0 \equiv 0 \quad \text{in} \quad \Omega_R.
$$

This completes the proof.

3.1. Modified nonlocal boundary value problem

One needs to truncate the infinite series of the exact DtN mapping at a finite order in practical computations to obtain an approximate DtN mapping written as

$$
T^N \varphi = \sum_{n=0}^{N} {}' \frac{k H_n^{(1)}{}'(kR)}{\pi H_n^{(1)}(kR)} \int_0^{2\pi} \varphi(R, \phi) \cos(n(\theta - \phi)) d\phi \tag{18}
$$

for $\forall \varphi \in H^s(\Gamma_R)$, $s \geq 1/2$. Here, the non-negative integer N is called the truncation order of the DtN mapping. Consequently, we arrive at a modified nonlocal boundary value problem consisting of (5), (6) and

$$
\frac{\partial u}{\partial n} = T^N u \quad \text{on} \quad \Gamma_R . \tag{19}
$$

To end this section, we include the point estimate for the difference of T and T^N in the next theorem. This estimate will be needed later.

Theorem 3.3. *Suppose DtN mappings T and T^N are defined as in (8) and (18) respectively. Then, for given $\varphi \in H^s(\Gamma_R)$, $s \in \mathbb{R}$, there holds, for $\forall t \geq 0$,*

$$
\|(T - T^N)\varphi\|_{H^{s-1}(\Gamma_R)} \leq c \frac{\epsilon(N, \varphi)}{N^t} \|\varphi\|_{H^{s+t}(\Gamma_R)} , \tag{20}
$$

where $c > 0$ is a constant dependent on kR but independent of φ and N, and $\epsilon(N, \varphi) \leq 1$ is a function of the truncation order N and the function φ satisfying $\epsilon(N, \varphi) \to 0$ as $N \to \infty$.

Proof: Suppose $\varphi(R, \theta)$ assumes the form

$$\varphi(R, \theta) = \sum_{n=0}^{\infty}{}' (a_n \cos(n\theta) + b_n \sin(n\theta)) \ ,$$

and hence

$$\|\varphi\|_{H^s(\Gamma_R)} = \{\frac{|a_0|^2}{2} + \sum_{n=1}^{\infty}(1+n^2)^s(|a_n|^2 + |b_n|^2)\}^{1/2} < +\infty \ . \qquad (21)$$

Then $T\varphi$ and $T^N\varphi$ on Γ_R read

$$T\varphi = \sum_{n=0}^{\infty}{}' \frac{kH_n^{(1)'}(kR)}{H_n^{(1)}(kR)}(a_n \cos(n\theta) + b_n \sin(n\theta)) \qquad (22)$$

and

$$T^N\varphi = \sum_{n=0}^{N}{}' \frac{kH_n^{(1)'}(kR)}{H_n^{(1)}(kR)}(a_n \cos(n\theta) + b_n \sin(n\theta)) \ , \qquad (23)$$

respectively. Subtracting (23) from (22), we arrive at

$$(T - T^N)\varphi = \sum_{n=N+1}^{\infty} \frac{kH_n^{(1)'}(kR)}{H_n^{(1)}(kR)}(a_n \cos(n\theta) + b_n \sin(n\theta)) \ . \qquad (24)$$

By the definition of the norm on the Sobolev space $H^s(\Gamma_R)$ and Lemma 3.2, we have

$$
\begin{aligned}
\|(T - T^N)\varphi\|_{H^{s-1}(\Gamma_R)} &= \{\sum_{n=N+1}^{\infty}(1+n^2)^{s-1}|k|^2\left|\frac{H_n^{(1)'}(kR)}{H_n^{(1)}(kR)}\right|^2(|a_n|^2 + |b_n|^2)\}^{1/2} \\
&= \{\sum_{n=N+1}^{\infty}(1+n^2)^s|k|^2\frac{1}{1+n^2}\left|\frac{H_n^{(1)'}(kR)}{H_n^{(1)}(kR)}\right|^2(|a_n|^2 + |b_n|^2)\}^{1/2} \\
&\leq c\{\sum_{n=N+1}^{\infty}(1+n^2)^s(|a_n|^2 + |b_n|^2)\}^{1/2} \\
&\leq \frac{c}{N^t}\{\sum_{n=N+1}^{\infty}(1+n^2)^{s+t}(|a_n|^2 + |b_n|^2)\}^{1/2} \\
&= c\frac{\epsilon(N, \varphi)}{N^t}\{\frac{|a_0|^2}{2} + \sum_{n=1}^{\infty}(1+n^2)^{s+t}(|a_n|^2 + |b_n|^2)\}^{1/2} \\
&= c\frac{\epsilon(N, \varphi)}{N^t}\|\varphi\|_{H^{s+t}(\Gamma_R)} \ . \qquad (25)
\end{aligned}
$$

10

Here, $c > 0$ is a constant dependent on kR but independent of φ and n, and $\epsilon(N, \varphi)$ is defined as

$$\epsilon(N, \varphi) = \frac{\{\sum\limits_{n=N+1}^{\infty} (1 + n^2)^{s+t}(|a_n|^2 + |b_n|^2)\}^{\frac{1}{2}}}{\{\frac{|a_0|^2}{2} + \sum\limits_{n=1}^{\infty}(1 + n^2)^{s+t}(|a_n|^2 + |b_n|^2)\}^{\frac{1}{2}}} \leq 1 , \qquad (26)$$

and a function of the truncation order N and φ for given values of s and t, generated by the addition of leading terms to the summation with positive terms for the construction of the norm on the space $H^{s+t}(\Gamma_R)$. Clearly, $\epsilon(N, \varphi) \to 0$ for $\forall \varphi \in H^s(\Gamma_R)$ as $N \to \infty$ because of (21). This completes the proof.

4. Weak formulation

We study in this section the weak formulation of (5)–(7), and its corresponding modified weak formulation of (5), (6) and (19). Essential mathematical features for both formulations will be presented.

The standard weak formulation of the nonlocal boundary value problem (5)–(7) reads: *Given $\partial u^i/\partial n$, find $u(x) \in H^1(\Omega_R)$ such that*

$$a(u, v) + b(u, v) = \ell(v) , \quad \forall v \in H^1(\Omega_R) , \qquad (27)$$

where $a(u, v) = \int_{\Omega_R} \nabla u \cdot \nabla \bar{v} dx - k^2 \int_{\Omega_R} u \bar{v} dx$ and $b(u, v) = -\int_{\Gamma_R} (Tu)\bar{v} ds$ are sesquilinear forms defined on $H^1(\Omega_R) \times H^1(\Omega_R)$, and ℓ defined by $\ell(v) = \int_{\Gamma} \frac{\partial u^i}{\partial n} \bar{v} ds \in H^{-1/2}(\Gamma)$ is a linear functional on $H^1(\Omega_R)$ dependent on $\frac{\partial u^i}{\partial n} \in H^{-1/2}(\Gamma)$. In addition, we point out that the operator T is self-adjoint, and hence the sesquilinear form $a(\cdot, \cdot) + b(\cdot, \cdot)$ defined in (27) is Hermitian.

Theorem 4.1. *The sesquilinear form $a(u, v) + b(u, v)$ in (27) satisfies*

$$|a(u, v) + b(u, v)| \leq c\|u\|_{H^1(\Omega_R)}\|v\|_{H^1(\Omega_R)}, \quad \forall u, v \in H^1(\Omega_R), \qquad (28)$$

where $c > 0$ is the continuity constant independent of u and v.

In order to obtain the existence for a weak solution of the variational equation (27), we need the next theorem.

Theorem 4.2. *The sesquilinear form $a(u, v) + b(u, v)$ in (27) satisfies a Gårding's inequality in the form*

$$Re\{a(v, v) + b(v, v)\} \geq \alpha \|v\|^2_{H^1(\Omega_R)} - \beta \|v\|^2_{H^{1-\epsilon}(\Omega_R)}, \quad \forall v \in H^1(\Omega_R), \quad (29)$$

where $\alpha > 0$, $\beta \geq 0$, and $1/2 > \epsilon > 0$ are constants independent of v.

Proof: We begin with the sesquilinear form $a(v, v)$ which reads

$$
\begin{aligned}
a(v, v) &= \int_{\Omega_R} |\nabla v|^2 dx - k^2 \int_{\Omega_R} |v|^2 dx \\
&= \|v\|^2_{H^1(\Omega_R)} - (k^2 + 1)\|v\|^2_{H^0(\Omega_R)}.
\end{aligned}
\quad (30)
$$

Therefore, the *Sobolev* embedding theorem gives

$$Re\{a(v, v)\} \geq \|v\|^2_{H^1(\Omega_R)} - c\|v\|^2_{H^{1-\epsilon}(\Omega_R)}, \quad (31)$$

where $c > 0$ and $1/2 > \epsilon > 0$ are constants. Next, we consider the sesquilinear form $b(v, v)$ which takes the form

$$
\begin{aligned}
b(v, v) &= -\int_{\Gamma_R} T v \bar{v} ds \\
&= -\frac{kR}{\pi} \sum_{n=0}^{\infty} {}' \frac{H_n^{(1)'}(kR)}{H_n^{(1)}(kR)} \int_0^{2\pi} \int_0^{2\pi} v(R, \phi)\overline{v(R, \theta)} \cos(n(\theta - \phi))d\theta d\phi .
\end{aligned}
\quad (32)
$$

Applying the recurrence relations of the Hankel function

$$\frac{H_n^{(1)'}(kR)}{H_n^{(1)}(kR)} = \frac{H_{n-1}^{(1)}(kR)}{H_n^{(1)}(kR)} - \frac{n}{kR}$$

to the right-hand side of (32) , we arrive at

$$
\begin{aligned}
b(v, v) &= \frac{1}{\pi} \sum_{n=1}^{\infty} n \int_0^{2\pi} \int_0^{2\pi} v(R, \phi)\overline{v(R, \theta)} \cos(n(\theta - \phi))d\theta d\phi \\
&\quad - \frac{kR}{\pi} \sum_{n=0}^{\infty} {}' \frac{H_{n-1}^{(1)}(kR)}{H_n^{(1)}(kR)} \int_0^{2\pi} \int_0^{2\pi} v(R, \phi)\overline{v(R, \theta)} \cos(n(\theta - \phi))d\theta d\phi \\
&= b_1(v, v) - b_2(v, v),
\end{aligned}
$$

12

where

$$b_1(v, v) = \frac{1}{\pi} \sum_{n=1}^{\infty} n \int_0^{2\pi} \int_0^{2\pi} v(R, \phi)\overline{v(R, \theta)} \cos(n(\theta - \phi))d\theta d\phi,$$

and

$$b_2(v, v) = \frac{kR}{\pi} \sum_{n=0}^{\infty} {}' \frac{H_{n-1}^{(1)}(kR)}{H_n^{(1)}(kR)} \int_0^{2\pi} \int_0^{2\pi} v(R, \phi)\overline{v(R, \theta)} \cos(n(\theta - \phi))d\theta d\phi.$$

Therefore, we have

$$b_1(v, v) \geq \pi \sum_{n=1}^{\infty} n(|a_n|^2 + |b_n|^2) \geq 0 , \tag{33}$$

where a_n and b_n are coefficients of the Fourier series of $v \in H^{1/2}(\Gamma_R)$. In addition, from Lemma 3.1, there holds

$$
\begin{aligned}
b_2(v, v) &\leq |b_2(v, v)| \\
&= \left| \frac{kR}{\pi} \sum_{n=0}^{\infty} {}' \frac{H_{n-1}^{(1)}(kR)}{H_n^{(1)}(kR)} \int_0^{2\pi} \int_0^{2\pi} v(R, \phi)\overline{v(R, \theta)} \cos(n(\theta - \phi))d\theta d\phi \right| \\
&\leq c\|v\|_{H^0(\Gamma_R)}^2,
\end{aligned} \tag{34}
$$

where $c > 0$ is a constant. By (33) and (34), we arrive at

$$Re\{b(v, v)\} = Re\{b_1(v, v) - b_2(v, v)\} \geq -c\|v\|_{H^0(\Gamma_R)}^2. \tag{35}$$

Furthermore, the *Sobolev* embedding theorem and the trace theorem yield

$$\|v\|_{H^0(\Gamma_R)}^2 \leq c\|v\|_{H^{\frac{1}{2}-\epsilon}(\Gamma_R)}^2 \leq c\|v\|_{H^{1-\epsilon}(\Omega_R)}^2, \tag{36}$$

where $c > 0$ and $1/2 > \epsilon > 0$ are constants. Consequently, (35) and (36) lead us to

$$Re\{b(v, v)\} \geq -c\|v\|_{H^{1-\epsilon}(\Omega_R)}^2, \tag{37}$$

where $c > 0$ is a constant. Finally, the combination of (31) and (37) yields (29). This completes the proof.

Now, the existence result follows immediately from the *Fredholm Alternative* theorem: Uniqueness implies the existence. As a consequence of Theorems 3.2 and 4.2, we have the following theorem.

13

Theorem 4.3. *The variational equation* (27) *admits a unique solution* $u \in H^1(\Omega_R)$.

4.1. Modified weak formulation

We now consider the modified variational equation of (27) for $u_N \in H^1(\Omega_R)$,

$$a(u_N, v) + b^N(u_N, v) = \ell(v), \quad \forall v \in H^1(\Omega_R) , \qquad (38)$$

where $b^N(u_N, v) = -\int_{\Gamma_R} (T^N u_N) \bar{v} ds$.

Theorem 4.4. *The sesquilinear form* $a(u, v) + b^N(u, v)$ *satisfies:*

1. $|a(u, v) + b^N(u, v)| \leq c \|u\|_{H^1(\Omega_R)} \|v\|_{H^1(\Omega_R)} \quad \forall u, v \in H^1(\Omega_R);$

2. $Re\{a(v, v) + b^N(v, v)\} \geq \alpha \|v\|^2_{H^1(\Omega_R)} - \beta \|v\|_{H^{1-\epsilon}(\Omega_R)}, \quad \forall v \in H^1(\Omega_R),$

where $c > 0$, $\alpha > 0$, $\beta \geq 0$ *and* $1/2 > \epsilon > 0$ *are all constants independent of* u *and* v.

Proof: We only show the proof of the second part. Following the same argument in Theorem 4.2, we obtain

$$Re\{a(v, v)\} \geq \|v\|^2_{H^1(\Omega_R)} - c\|v\|^2_{H^{1-\epsilon}(\Omega_R)}, \quad \forall v \in H^1(\Omega_R), \qquad (39)$$

where $c > 0$ and $1/2 > \epsilon > 0$ are constants. In addition, $b^N(v, v)$ can be written in the form

$$
\begin{aligned}
b^N(v, v) &= \frac{1}{\pi} \sum_{n=1}^{N} n \int_0^{2\pi} \int_0^{2\pi} v(R, \phi) \overline{v(R, \theta)} \cos(n(\theta - \phi)) d\theta d\phi \\
&\quad - \frac{kR}{\pi} \sum_{n=0}^{N} {}' \frac{H_{n-1}^{(1)}(kR)}{H_n^{(1)}(kR)} \int_0^{2\pi} \int_0^{2\pi} v(R, \phi) \overline{v(R, \theta)} \cos(n(\theta - \phi)) d\theta d\phi \\
&= b_1^N(v, v) - b_2^N(v, v), \qquad (40)
\end{aligned}
$$

where

$$b_1^N(v, v) = \frac{1}{\pi} \sum_{n=1}^{N} n \int_0^{2\pi} \int_0^{2\pi} v(R, \phi) \overline{v(R, \theta)} \cos(n(\theta - \phi)) d\theta d\phi \qquad (41)$$

and

$$b_2^N(v, v) = \frac{kR}{\pi} \sum_{n=0}^{N} {}' \frac{H_{n-1}^{(1)}(kR)}{H_n^{(1)}(kR)} \int_0^{2\pi} \int_0^{2\pi} v(R, \phi) \overline{v(R, \theta)} \cos(n(\theta-\phi)) d\theta d\phi. \qquad (42)$$

14

We are able to show that

$$
\begin{aligned}
b_1^N(v,v) &= \frac{1}{\pi} \sum_{n=1}^{N} n \int_0^{2\pi} \int_0^{2\pi} v(R,\phi)\overline{v(R,\theta)} \cos(n(\theta-\phi))d\theta d\phi \\
&= \pi \sum_{n=1}^{N} n(|a_n|^2 + |b_n|^2) \\
&\geq 0,
\end{aligned}
\tag{43}
$$

and

$$
\begin{aligned}
b_2^N(v,v) &\leq |b_2^N(v,v)| \\
&= \left| \frac{kR}{\pi} \sum_{n=0}^{N}{}' \frac{H_{n-1}^{(1)}(kR)}{H_n^{(1)}(kR)} \int_0^{2\pi} \int_0^{2\pi} v(R,\phi)\overline{v(R,\theta)} \cos(n(\theta-\phi))d\theta d\phi \right| \\
&\leq c\|v\|_{H^0(\Gamma_R)}^2 .
\end{aligned}
\tag{44}
$$

Here, a_n and b_n are coefficients of the Fourier series of $v \in H^{1/2}(\Gamma_R)$, and $c > 0$ is a constant. Inequalities (43) and (44) yield

$$
Re\{b^N(v,v)\} \geq -c\|v\|_{H^0(\Gamma_R)}^2,
\tag{45}
$$

and this implies further

$$
Re\{b^N(v,v)\} \geq -c\|v\|_{H^{1-\epsilon}(\Omega_R)}^2,
\tag{46}
$$

due to the *Sobolev* embedding theorem and the trace theorem. Here, $c > 0$ and $1/2 > \epsilon > 0$ are constants. Consequently, by the combination of (39) and (46), we complete the proof of the second part immediately.

Theorem 4.5. *There exists a constant $N_0 \geq 0$ such that the modified variational equation (38) has at most one solution $u_N \in H^1(\Omega)$ for $N \geq N_0$.*

Proof: We argue by contradiction. If the theorem does not hold, then for each N_0, there is an $N = N(N_0) \geq N_0$ and $u_N = u_{N(N_0)} \in H^1(\Omega_R)$ such that

$$
a(u_N,\varphi) + b^N(u_N,\varphi) = 0 , \quad \forall \varphi \in H^1(\Omega_R) ,
\tag{47}
$$

and $\|u_N\|_{H^1(\Omega_R)} = 1$. There is a subsequence denoted by $\{u_{N(i)}\}$ which converges weakly to some $u \in H^1(\Omega_R)$. We may assume that this subsequence also

converges strongly to u in $H^{1-s}(\Omega_R)$ for $0 < s < 1$ since $H^1(\Omega_R)$ is compactly embedded in $H^{1-s}(\Omega_R)$.

Now for $N = N(i)$, we write

$$
\begin{aligned}
0 &= a(u_N, \varphi) + b^N(u_N, \varphi) \\
&= a(u_N - u, \varphi) + b^N(u_N - u, \varphi) + b^N(u, \varphi) - b(u, \varphi) \\
&\quad + a(u, \varphi) + b(u, \varphi) \ .
\end{aligned}
\tag{48}
$$

For smooth test function $\varphi \in C^\infty(\bar{\Omega}_R)$, we see that for $s < 1/2$, by the generalized Cauchy-Schwartz inequality (see p.50 of [26] or p.166 of [24]),

$$
|a(u_N - u, \varphi)| \leq c\|u_N - u\|_{H^{1-s}(\Omega_R)}\|\varphi\|_{H^{1+s}(\Omega_R)} \to 0 \quad \text{as} \quad i \to \infty \ ,
$$

and similarly,

$$
|b^N(u_N - u, \varphi)| \leq c\|u_N - u\|_{H^{1-s}(\Omega_R)}\|\varphi\|_{H^{1+s}(\Omega_R)} \to 0 \quad \text{as} \quad i \to \infty \ ,
$$

in view of the trace theorem. Also,

$$
\begin{aligned}
|b^N(u, \varphi) - b(u, \varphi)| &= |\langle (T - T^N)u, \varphi \rangle_{\Gamma_R}| \\
&\leq \|(T - T^N)u\|_{H^{-1/2-s}(\Gamma_R)}\|\varphi\|_{H^{1/2+s}(\Gamma_R)} \to 0 \quad \text{as} \quad i \to \infty
\end{aligned}
$$

as a consequence of Theorem 3.3. Finally, taking the limit of (48) as $i \to \infty$ gives

$$
a(u, \varphi) + b(u, \varphi) = 0 \ .
\tag{49}
$$

By density of smooth functions in $H^1(\Omega_R)$, (49) holds for all $\varphi \in H^1(\Omega_R)$ which further implies that $u = 0$ because of Theorem 4.3. Therefore, $u_{N(i)} \to 0$ in $L^2(\Omega_R)$, so

$$
\|u_{N(i)}\|_{H^0(\Omega_R)} \to 0 \quad \text{as} \quad i \to \infty \ .
\tag{50}
$$

For $N = N(i)$, on the other hand, we have

$$
\begin{aligned}
\|u_N\|_{H^1(\Omega_R)}^2 &= (k^2 + 1)\|u_N\|_{H^0(\Omega_R)}^2 + a(u_N, u_N) \\
&\leq (k^2 + 1)\|u_N\|_{H^0(\Omega_R)}^2 + a(u_N, u_N) + b_1^N(u_N, u_N) \\
&= (k^2 + 1)\|u_N\|_{H^0(\Omega_R)}^2 + a(u_N, u_N) + b^N(u_N, u_N) + b_2^N(u_N, u_N) \\
&= (k^2 + 1)\|u_N\|_{H^0(\Omega_R)}^2 + b_2^N(u_N, u_N)
\end{aligned}
\tag{51}
$$

because of (30), (43), (40) and (47) , respectively. Moreover, we see from (34)

$$b_2(u_N, u_N) \le c\|u_N\|^2_{H^{1/2+\epsilon}(\Omega_R)} \le c\|u_N\|^{1+2\epsilon}_{H^1(\Omega_R)}\|u_N\|^{1-2\epsilon}_{H^0(\Omega_R)} \qquad (52)$$

due to the interpolation inequality for $0 < \theta = 1/2 + \epsilon < 1$ ([27]). Finally, (51) and (52), together with the fact that $\|u_N\|_{H^1(\Omega_R)} = 1$, give

$$1 \le (k^2 + 1)\|u_N\|^2_{H^0(\Omega_R)} + c\|u_N\|^{1-2\epsilon}_{H^0(\Omega_R)} \qquad (53)$$

which contradicts (50). This completes the proof.

Theorem 4.4 means that that one can apply the *Fredholm Alternative*: Uniqueness implies existence. By Theorem 4.5, we have the following theorem.

Theorem 4.6. *There exists a constant $N_0 \ge 0$ such that the modified variational equation* (38) *admits a unique solution $u_N \in H^1(\Omega_R)$ for $N \ge N_0$.*

5. Finite element analysis

Our main goal in this section is to establish a priori error estimates for the finite element solution of (38) in terms of the finite element meshsize h and the truncation order N in the appropriate *Sobolev* spaces.

5.1. Galerkin formulation

Let S_h be the standard finite element space. Now we consider the Galerkin formulation of (38): *Given $\partial u^i/\partial n$, find $u_h \in S_h \subset H^1(\Omega_R)$ satisfying*

$$a(u_h, v_h) + b^N(u_h, v_h) = \ell(v_h), \quad \forall\, v_h \in S_h. \qquad (54)$$

We can show ([20]) that the discrete sesquilinear form $a(u_h, v_h) + b^N(u_h, v_h)$ satisfies the BBL-condition as implication of the following:

Gårding's inequality + Uniqueness + Approximation property of S_h $\Rightarrow$ BBL-condition.

Theorem 5.1. *If the sesquilinear form $a(v, w) + b^N(v, w)$ in* (38) *satisfies the following conditions:*

1. $Re\{a(v,v) + b^N(v,v) + (Cv,v)_{H^1(\Omega_R)}\} \geq \alpha\|v\|^2_{H^1(\Omega_R)}, \quad \forall\, v \in H^1(\Omega_R);$

2. $\{v \in H^1(\Omega_R)|a(v,w) + b^N(v,w) = 0, \quad \forall\, w \in H^1(\Omega_R)\} = \{0\};$

3. *Finite element space $S_h \subset H^1(\Omega_R)$ satisfies the standard approximation property.*

Then, there exists a constnat $h_0 > 0$ such that $a(v,w) + b^N(v,w)$ for $0 < h \leq h_0$ satisfies the BBL condition in the form

$$\sup_{0 \neq w_h \in S_h} \frac{|a(v_h, w_h) + b^N(v_h, w_h)|}{\|w_h\|_{H^1(\Omega_R)}} \geq \gamma\|v_h\|_{H^1(\Omega_R)}, \quad \forall\, v_h \in S_h \ . \quad (55)$$

Here, C is a compact operator from $H^1(\Omega_R)$ to $H^1(\Omega_R)$, $(\cdot,\cdot)_{H^1(\Omega_R)}$ stands for the inner product on $H^1(\Omega_R)$, $\alpha > 0$ is a constant, and $\gamma > 0$ is the inf-sup constant independent of h.

Remark: In [28], it has been shown that the inf-sup constant γ in (55) has the order of $1/k$ for the one-dimensional Helmholtz equation ; more precisely, there exist positive constants c_1, c_2 independent of the wave number k such that $\frac{c_1}{k} \leq \gamma \leq \frac{c_2}{k}$. In brief, the larger for the magnitude of wave number k, the more oscillate for the solution of the Helmholtz equation, i.e. the finer mesh required in finite element discretization. In this paper, we have no desire to deal with high frequency acoustic waves. Interested readers are refered to [28], [29], [30], and to name a few.

Once the BBL condition (55) is established, we are in the position to derive a priori error estimates for the finite element solution $u_h \in S_h$.

5.2. Asymptotic error estimates

A priori error estimates including error effects of both the numerical discretization and the truncation of infinite series seem to be more reasonable. To this end, we first derive an upper bound of numerical errors analogous to the well-known *Céa*'s lemma in the positive definite case.

Theorem 5.2. *There exist constants $h_0 > 0$ and $N_0 \geq 0$ such that for any $0 < h \leq h_0$ and $N_0 \leq N$*

$$\|u - u_h\|_{H^1(\Omega_R)} \leq c\{\inf_{w_h \in S_h} \|u - w_h\|_{H^1(\Omega_R)} + \sup_{0 \neq w_h \in S_h} \frac{|b(u, w_h) - b^N(u, w_h)|}{\|w_h\|_{H^1(\Omega_R)}}\},$$

(56)

where $c > 0$ is a constant independent of h and N.

Proof: We begin with the BBL condition (55)

$$\gamma \|v_h\|_{H^1(\Omega_R)} \leq \sup_{0 \neq w_h \in S_h} \frac{|a(v_h, w_h) + b^N(v_h, w_h)|}{\|w_h\|_{H^1(\Omega_R)}}, \quad \forall v_h \in S_h,$$

where $\gamma > 0$ is a constant. Replacing v_h with $u_h - v_h \in S_h$ in the above inequality, we arrive at

$$\gamma \|u_h - v_h\|_{H^1(\Omega_R)} \leq \sup_{0 \neq w_h \in S_h} \frac{|a(u_h - v_h, w_h) + b^N(u_h - v_h, w_h)|}{\|w_h\|_{H^1(\Omega_R)}}, \quad \forall v_h \in S_h.$$

(57)

According to (27) and (54), we have, for $\forall w_h \in S_h$,

$$a(u, w_h) + b^N(u, w_h) = \ell(w_h) + b^N(u, w_h) - b(u, w_h)$$

(58)

and

$$a(u_h, w_h) + b^N(u_h, w_h) = \ell(w_h),$$

(59)

respectively. Therefore, subtracting (58) from (59) leads to

$$a(u_h - u, w_h) + b^N(u_h - u, w_h) = b(u, w_h) - b^N(u, w_h), \quad \forall w_h \in S_h.$$

(60)

In the mean time, a simple manipulation gives

$$
\begin{aligned}
a(u_h - v_h, w_h) + b^N(u_h - v_h, w_h) &= a(u_h - u + u - v_h, w_h) + b^N(u_h - u + u - v_h, w_h) \\
&= a(u_h - u, w_h) + b^N(u_h - u, w_h) + \\
&\quad a(u - v_h, w_h) + b^N(u - v_h, w_h).
\end{aligned}
$$

(61)

Therefore, by (60) and (61), the inequality (57) implies that

$$
\begin{aligned}
\gamma \|u_h - v_h\|_{H^1(\Omega_R)} &\leq \sup_{0 \neq w_h \in S_h} \frac{|a(u_h - v_h, w_h) + b^N(u_h - v_h, w_h)|}{\|w_h\|_{H^1(\Omega_R)}} \\
&= \sup_{0 \neq w_h \in S_h} \frac{|a(u - v_h, w_h) + b^N(u - v_h, w_h) + a(u_h - u, w_h) + b^N(u_h - u, w_h)|}{\|w_h\|_{H^1(\Omega_R)}} \\
&= \sup_{0 \neq w_h \in S_h} \frac{|a(u - v_h, w_h) + b^N(u - v_h, w_h) + b(u, w_h) - b^N(u, w_h)|}{\|w_h\|_{H^1(\Omega_R)}} \\
&\leq \sup_{0 \neq w_h \in S_h} \frac{|a(u - v_h, w_h) + b^N(u - v_h, w_h)|}{\|w_h\|_{H^1(\Omega_R)}} + \sup_{0 \neq w_h \in S_h} \frac{|b(u, w_h) - b^N(u, w_h)|}{\|w_h\|_{H^1(\Omega_R)}} \\
&\leq c \|u - v_h\|_{H^1(\Omega_R)} + \sup_{0 \neq w_h \in S_h} \frac{|b(u, w_h) - b^N(u, w_h)|}{\|w_h\|_{H^1(\Omega_R)}}, \quad (62)
\end{aligned}
$$

where c is a positive constant. Consequently, the triangular inequality and the formulation (62) yield, $\forall\, v_h \in S_h$,

$$
\begin{aligned}
\|u - u_h\|_{H^1(\Omega_R)} &\leq \|u - v_h\|_{H^1(\Omega_R)} + \|u_h - v_h\|_{H^1(\Omega_R)} \\
&\leq (1 + \frac{c}{\gamma}) \|u - v_h\|_{H^1(\Omega_R)} + \frac{1}{\gamma} \sup_{0 \neq w_h \in S_h} \frac{|b(u, w_h) - b^N(u, w_h)|}{\|w_h\|_{H^1(\Omega_R)}},
\end{aligned}
$$

and this further leads to

$$
\|u - u_h\|_{H^1(\Omega_R)} \leq c \{ \inf_{v_h \in S_h} \|u - v_h\|_{H^1(\Omega_R)} + \sup_{0 \neq w_h \in S_h} \frac{|b(u, w_h) - b^N(u, w_h)|}{\|w_h\|_{H^1(\Omega_R)}} \},
$$

where $c > 0$ is a constant independent of h and N. This completes the proof.

According to the estimate (56), we are able to observe that the numerical errors are dominated by two single terms. We study the first term correlated with the meshsize h by using the approximation theory, and the second term dependent on the truncation order N by employing the Fourier analysis. In the following, starting with the estimate (56), we first derive a priori error estimates in the energy space $H^1(\Omega_R)$, and then a priori error estimates measured in L^2-norm to conclude this section.

Theorem 5.3. *Suppose that $u \in H^t(\Omega_R)$, for $\forall\, 2 \leq t \in \mathbb{R}$. Then, there exist constants $h_0 > 0$ and $N_0 \geq 0$ such that for any $0 < h \leq h_0$ and $N_0 \leq N$*

$$
\|u - u_h\|_{H^1(\Omega_R)} \leq c(h^{t-1} + \frac{\epsilon(N, u)}{N^{t-1}}) \|u\|_{H^t(\Omega_R)}, \quad (63)
$$

where $c > 0$ is a constant independent of h and N, and $\epsilon(N, u) \le 1$ is a function

of the truncation order N and the function u satisfying $\epsilon(N, u) \to 0$ as $N \to \infty$.

Proof: We have known from Theorem 5.2 that

$$\|u - u_h\|_{H^1(\Omega_R)} \le c\{ \inf_{v_h \in S_h} \|u - v_h\|_{H^1(\Omega_R)} + \sup_{0 \neq w_h \in S_h} \frac{|b(u, w_h) - b^N(u, w_h)|}{\|w_h\|_{H^1(\Omega_R)}} \},$$

$$(64)$$

where c is a positive constant. Regarding the first term in (64), the approximation property of the finite element space S_h gives

$$\inf_{v_h \in S_h} \|u - v_h\|_{H^1(\Omega_R)} \le c h^{t-1} \|u\|_{H^t(\Omega_R)},$$

$$(65)$$

where c is a positive constant independent of h. We now consider the second term in (64). According to the trace theorem, there exists a bounded linear operator $\gamma : H^1(\Omega_R) \to H^{1/2}(\Gamma_R)$ such that

$$\frac{|b(u, w_h) - b^N(u, w_h)|}{\|w_h\|_{H^1(\Omega_R)}} = \frac{|\langle (T - T^N)\gamma u, \gamma w_h \rangle_{\Gamma_R}|}{\|w_h\|_{H^1(\Omega_R)}} = \frac{|\langle \gamma^*(T - T^N)\gamma u, w_h \rangle_{\Omega_R}|}{\|w_h\|_{H^1(\Omega_R)}},$$

$$(66)$$

where $\langle \cdot, \cdot \rangle_{\Gamma_R}$ is the standard L^2 duality pairing between $H^{-1/2}(\Gamma_R)$ and $H^{1/2}(\Gamma_R)$, and $\gamma^* : H^{-1/2}(\Gamma_R) \to (H^1(\Omega_R))'$ is the adjoint operator of γ. As a consequence, we have

$$
\begin{aligned}
\sup_{0 \neq w_h \in S_h} \frac{|b(u, w_h) - b^N(u, w_h)|}{\|w_h\|_{H^1(\Omega_R)}} &= \sup_{0 \neq w_h \in S_h} \frac{|\langle \gamma^*(T - T^N)\gamma u, w_h \rangle_{\Omega_R}|}{\|w_h\|_{H^1(\Omega_R)}} \\
&= \|\gamma^*(T - T^N)\gamma u\|_{(H^1(\Omega_R))'} \\
&\le c\|(T - T^N)\gamma u\|_{H^{-1/2}(\Gamma_R)} \\
&\le c\frac{\epsilon(N, u)}{N^{t-1}}\|\gamma u\|_{H^{t-1/2}(\Gamma_R)} \\
&\le c\frac{\epsilon(N, u)}{N^{t-1}}\|u\|_{H^t(\Omega_R)}
\end{aligned}
$$

$$(67)$$

because of Theroem 3.3 and the boundedness of operators γ and γ^*. The proof is hence established by following a combination of (65) and (67).

Remark: The estimate (63) can be easily simplified by replacing $\epsilon(N, u)$ by 1, since $\epsilon(N, u) \le 1$. However, we decide to keep the form (63). We note that

$\epsilon(N, u)$ depends on both N and u although its precise dependence can not be determined explicitly. Numerical tests in Section 6 show that the convergence of the function $\epsilon(N, u)$ is extremely fast, and its rate decays as the number kR increases.

We now extend the error estimate in the energy space to the one measured in the $L^2(\Omega_R)$ space.

Theorem 5.4. *Suppose that $u \in H^t(\Omega_R)$, for $\forall 2 \le t \in \mathbb{R}$. Then there exist constants $h_0 > 0$ and $N_0 \ge 0$ such that for any $0 < h \le h_0$ and $N_0 \le N$*

$$\|u - u_h\|_{L^2(\Omega_R)} \le c(h^t + \frac{\epsilon(N, u)}{N^t})\|u\|_{H^t(\Omega_R)}, \tag{68}$$

where $c > 0$ is a constant independent of h and N, and $\epsilon(N, u)$ is a function of the truncation order N and the function u satisfying $\epsilon(N, u) \to 0$ as $N \to \infty$.

Proof: Suppose that u is the solution of the variational equation (27), and u_h is the finite element solution of the variational equation (54). By (60) in Theorem 5.2, we have

$$a(e, v_h) + b^N(e, v_h) + b(u, v_h) - b^N(u, v_h) = 0, \quad \forall v_h \in S_h, \tag{69}$$

where $e = u - u_h$ is the finite element error. Now, we consider the following boundary value problem: *Find $w \in C^2(\Omega_R) \cap C^1(\overline{\Omega_R})$ satisfying*

$$\Delta w + k^2 w = e \quad \text{in} \quad \Omega_R, \tag{70}$$

$$\frac{\partial w}{\partial n} = 0 \quad \text{on} \quad \Gamma, \tag{71}$$

$$\frac{\partial w}{\partial n} = Tw \quad \text{on} \quad \Gamma_R. \tag{72}$$

Let w be a weak solution of nonlocal boundary value problem (70)–(72), and hence w satisfies

$$a(v, w) + b^N(v, w) + b(v, w) - b^N(v, w) = (v, e)_{L^2(\Omega_R)}, \quad \forall v \in H^1(\Omega_R), \tag{73}$$

where $(\cdot, \cdot)_{L^2(\Omega_R)}$ stands for the inner product on $L^2(\Omega_R)$. In particular, we choose v to be e in (73), and then obtain

$$a(e, w) + b^N(e, w) + b(e, w) - b^N(e, w) = (e, e)_{L^2(\Omega_R)} = \|e\|_{L^2(\Omega_R)}^2. \tag{74}$$

22

Subtracting (69) from (74) leads to, $\forall \, v_h \in S_h$,

$$\|e\|_{L^2(\Omega_R)}^2 = a(e, w - v_h) + b^N(e, w - v_h) + b(e, w) - b^N(e, w) + b^N(u, v_h) - b(u, v_h). \tag{75}$$

Theorem 4.4, the approximation property of S_h and the regularity theory imply that

$$
\begin{aligned}
|a(e, w - v_h) + b^N(e, w - v_h)| &\leq c\|e\|_{H^1(\Omega_R)}\|w - v_h\|_{H^1(\Omega_R)} \\
&\leq ch\|e\|_{H^1(\Omega_R)}\|w\|_{H^2(\Omega_R)} \\
&\leq ch\|e\|_{H^1(\Omega_R)}\|e\|_{L^2(\Omega_R)}, \tag{76}
\end{aligned}
$$

where c is a positive constant. Following the same argument in Theorem 5.3 and choosing $t = 2$, we arrive at, by the regularity theory,

$$|b(e, w) - b^N(e, w)| \leq c_1 \, \frac{\epsilon_1(N, u)}{N}\|e\|_{H^1(\Omega_R)}\|e\|_{L^2(\Omega_R)}. \tag{77}$$

Similarly, we also have

$$
\begin{aligned}
|b(u, v_h) - b^N(u, v_h)| &\leq |b(u, w - v_h) - b^N(u, w - v_h)| + |b(u, w) - b^N(u, w)| \\
&\leq c_2 \, \frac{\epsilon_2(N, u)h}{N^{t-1}}\|u\|_{H^t(\Omega_R)}\|e\|_{L^2(\Omega_R)} + c_3 \, \frac{\epsilon_3(N, u)}{N^t}\|u\|_{H^t(\Omega_R)}\|e\|_{L^2(\Omega_R)}. \tag{78}
\end{aligned}
$$

In above formulations, $\{\epsilon_j(N, u)\}|_{j=1}^{j=3} \leq 1$ are similar to the function $\epsilon(N, u)$ in Theorem 5.3, and $\{c_j\}|_{j=1}^{j=3}$ are positive constants. Therefore, by the combination of the inequalities (76)–(78) and (63), the equation (75) yields

$$\|u - u_h\|_{L^2(\Omega_R)} \leq c(h^t + \frac{\epsilon(N, u)}{N^t})\|u\|_{H^t(\Omega_R)}, \quad \forall \, 2 \leq t \in \mathbb{R}, \tag{79}$$

where $c > 0$ is a constant independent of h and N. This completes the proof.

The error estimates (63) and (68) demonstrate that the finite element approximation u_h converges to u, the weak solution of variational equation (27), as $h \rightarrow 0$ and $N \rightarrow \infty$.

23

6. Numerical experiments

In this section, we present the results of several numerical tests to validate our theoretical results.

6.1. A model problem

We first introduce a model problem whose analytical solutions can be obtained easily so that we are able to evaluate the accuracy of the numerical solutions. We compute the scattering, by an infinite circular cylinder of radius R_0, of a plane wave $u^i = e^{ikx \cdot d}$ propagating along the positive x_1 axis with the sound-hard boundary condition on the surface of scatterer. $x = (x_1, x_2) \in \mathbb{R}^2$ and $d = (1, 0)$ which is the unit vector describing the direction of traveling of the incident wave. The mathematical model can be formulated as the exterior boundary value problem (1)–(3) with the boundary Γ to be a circle of radius R_0. In this case the exact solution u assumes the form

$$u(r, \theta) = -\sum_{n \in \mathbb{Z}} i^n \frac{J_n{}'(kR_0)}{H_n^{(1)'}(kR_0)} H_n^{(1)}(kr) e^{in\theta}, \quad \forall r \geq R_0 . \tag{80}$$

In the following simulations, the infinite Fourier series (80), representing the exact solution, are truncated when the relative change due to an additional mode in the fields is below 10^{-6}.

We choose the artificial boundary Γ_R to be a circle of radius R ($R > R_0$) with the same center as Γ. Therefore, the computational region Ω_R is the annulus between Γ and Γ_R (see Figure 3). We map the computational annulus region $\{x | R_0 \leq |x| \leq R\}$ into a rectangle $\{(r, \theta) | r \in [R_0, R], \theta \in [0, 2\pi)\}$ ([2]) discretized by uniform rectangle elements, and employ the piecewise linear basis functions $\{\varphi_i\}|_{i=1}^{i=NP}$ in terms of r and θ to construct the finite element space S_h. Here $NP = (N_r + 1) \times N_\theta$ is total number of elements, and N_r and N_θ denote the number of elements in the radial and angular direction respectively. During the following numerical tests, if there is no specification, we use the correlation rule $N_\theta \sim 4kR_0N_r$ to guide our discretization of the computational domain Ω_R. A direct solver is employed for the solutions of the resulting linear system.
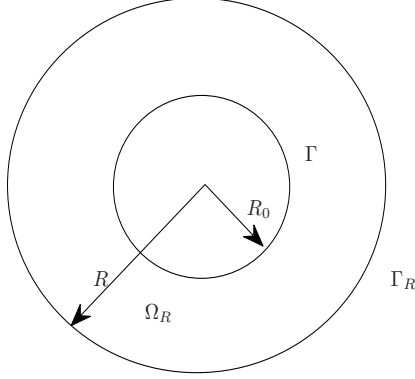
Figure 2: The computational domain of the model problem

To find the finite element solution of (54), we must be able to numerically impose the nonlocal boundary condition $\frac{\partial u}{\partial n} = T^N u$ into the evaluation of the sesquilinear form $b^N(u,v) = -\int_{\Gamma_R} T^N u \overline{v} ds$. In the discrete formulation, this amounts to computing the integrals

$$\int_{\Gamma_R} T^N \varphi_j \overline{\varphi_i} ds. \tag{81}$$

As for our computation, the finite element space S_h consists of piecewise linear functions φ_j, $j = 1, 2, \ldots, NP$, and most of them will vanish on the boundary Γ_R correspondingly eliminating the complexity of the above procedure. More precisely, the computation of integrals (81) amounts to evaluating the following series ([25])

$$
\begin{aligned}
\int_{\Gamma_R} T^N \varphi_j \overline{\varphi_i} ds &= \sum_{n=0}^{N} {}' \frac{kR H_n^{(1)'}(kR)}{\pi H_n^{(1)}(kR)} \int_0^{2\pi} \int_0^{2\pi} \varphi_j(R,\phi) \overline{\varphi_i(R,\theta)} \cos(n(\theta - \phi)) d\theta d\phi \\
&= \frac{4kR}{\pi \Delta\theta^2} \sum_{n=0}^{N} {}' \frac{H_n^{(1)'}(kR)}{H_n^{(1)}(kR)} \frac{(1 - \cos(n\Delta\theta))^2}{n^4} \cos(n(\theta_j - \theta_i)). \tag{82}
\end{aligned}
$$

Here the prime $'$ behind the summation implies that the first term in the summation is multiplied by $1/2$. The term in the summation as $n = 0$ can be obtained by taking the limit $n \to 0$.

25

6.2. Numerical tests

In the first test, we compute the model problem to report the effects of numerical discretization errors. According to Theorem 5.3 and Theorem 5.4, as the truncation order N of the DtN mapping is appropriate, we should be able to observe that $\|u - u_h\|_{H^1(\Omega_R)} = O(h)$ and $\|u - u_h\|_{L^2(\Omega_R)} = O(h^2)$ for the finite element space S_h. We choose $R_0 = 1$ and $R = 2$. Three different cases for the wave numbers $k = 1, 2$ and 4 are considered. Figure 3 shows the log-log plot of errors measured in L^2 and H^1-norms with respect to $1/h = N_r/(R - R_0)$ (here and in the sequel, we refer to this equation for the size of h). Slopes of -2 on the left and -1 on the right verify the convergence order of $O(h^2)$ and $O(h)$.
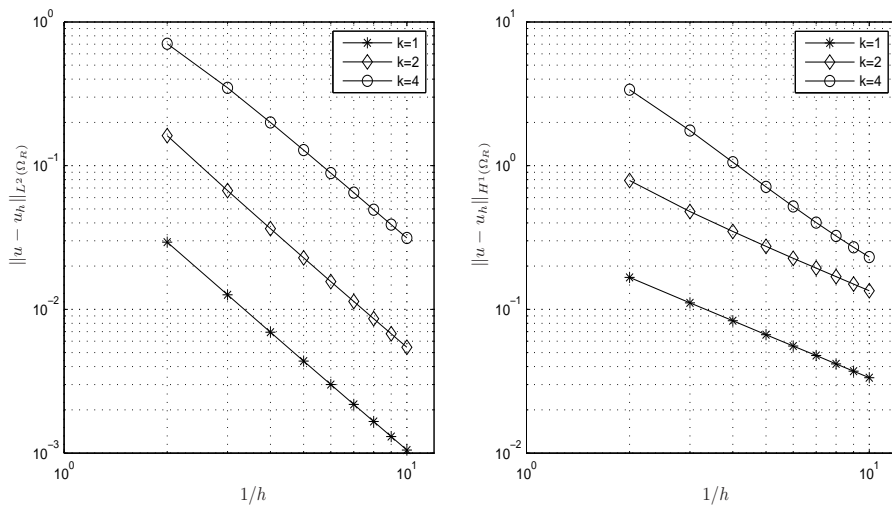

Figure 3: Log-log plot vs $1/h = N_r/(R - R_0)$ for errors in L^2-norm (left), and H^1-norm (right).

Remark: The quality of discrete numerical solutions to the Helmholtz equation depends significantly on the physical wave number k. It is known that the meshsize h in the finite element computations should be proportional to the wave number k ([31]). Therefore, Figure 3 also indicates that the accuracy de-

cays correspondingly as the wave number k increases under the same resolution.

The second numerical test is concerned with the effects of truncation order N on the total numerical errors, and its correlation with that of finite element discretization. Here, we only consider the errors measured in L^2-norm, and should observe the convergence order $O(\epsilon(N, u)(\frac{1}{N})^2)$ according to (68), provided sufficiently small meshsize h. We set $R_0 = 0.5$, $R = 1$, and $kR = 4$, and compute for four different values of $h = 1/4, 1/12, 1/20, 1/30$, respectively. The log-log plots of errors are presented in Figure 4 (left) showing that the errors due to the truncation of the DtN mapping decay extremely fast. For instance, we can see the super-exponential convergence order for all different values of h. It is actually expected since we are aware that $\epsilon(N, u) \to 0$ exponentially for sufficiently smooth functions u as $N \to \infty$. Meanwhile, the term $O((\frac{1}{N})^2)$ contributes more to the convergence rate. Secondly, the accuracy arrives at the optimal as $N = N_o = 3$ for $h = 1/4$ and $N_o = 4$ for all other values of h. Here, N_o, the optimal truncation order of the DtN mapping, is defined as the minimum number of N required to attain the optimal order of accuracy with respect to a given set of meshsize h and the number kR (we will show N_o is also dependent on kR in the next test). It implies that the rate at which $1/N_o$ decreases is much lower than the rate h decays for optimal order of accuracy. More precisely, $1/N_o = 1/3$ is required as $h = 1/4$, while $1/N_o = 1/4$ is needed as $h = 1/30$. In addition, we observe that there are no numerical improvements as the truncation order $N > N_o$ is employed for each value of h, and this point can be easily understood because of the inherent restriction of accuracy related with the value of h. We experience the similar restriction of accuracy related with the value of truncation order N as well, i.e. there are no improvments of accuracy for the employment of h less than some value corresponding to a given value of N. For instance, looking at the vertical line for $N = 3$ in the Figure 4 (left), we can see that the error decays from 10^{-1} to 10^{-2} as the mesh size h decreases from $1/4$ down to $1/12$, and remains stable for $h = 1/20$ and $1/30$.

Finally, we revisit the numerical rule $N \geq kR$ proposed by Harari *et al.*
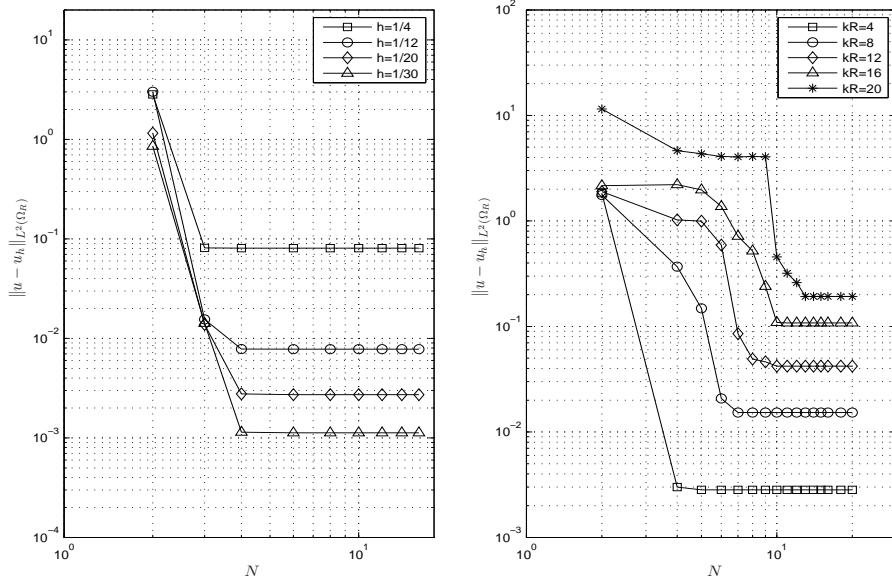
Figure 4: Log-log plot of errors in L^2-norm vs the truncation order N for $R_0 = 0.5$, $R = 1$ and $kR = 4$ (left); The correlation between the truncation order N and the parameter kR as $h = 1/15$ (right).

([32, 33]) in the third numerical test. We choose the inner radius $R_0 = 1$ and the outer radius $R = 2$, and the invariant meshsize $h = 1/15$. The log-log plots of numerical errors measured in L^2-norm are presented in Figure 4 (right) showing that the optimal truncation order N_o increases linearly with kR in order to arrive at the optimal order of accuracy. The optimal truncation order N_o equals to 4 as $kR = 4$, while N_o goes up to 13 as $kR = 20$. In addition, with respect to each value of kR, as long as the optimal order of accuracy is attained, no improvement of accuracy can be observed as $N > N_o$. Our numerical results are in good agreement with the numerical rule $N \geq kR$. Since the meshsize h is invariant, we also can see that the optimal accuracy decays as the wave number k increases. Finally, we want to indicate that, although the value of R stays invariant in the presented results, the optimal truncation order N_o increases with kR as well if the wave number k remains unchanged.

28

As a summary of above numerical tests, we give the following Chart (see Figure 5) guiding numerical computations as applying the coupling of the finite element method and the analytical method for the solution of exterior acoustic scattering problems. Here, we use $\epsilon(N) = \epsilon(N, u)$, and the L^2 error, and the interaction Chart with error estimates in the energy space can be attained accordingly.
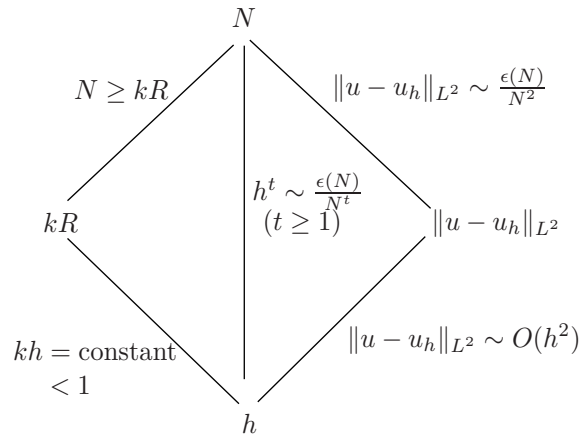

Figure 5: The correlation among numerical errors and parameters.

References

[1] C. Johnson, J. Nedelec, On the coupling of boundary integral and finite element methods, Mathematics of Computation 35 (1980) 1063–1079.

[2] R. MacCamy, S. Marin, A finite element method for exterior interface problems, Internat. J. Math. and Math. Sci. 3 (1980) 311–350.

[3] K. Feng, Finite element method and natural boundary reduction, Proceedings of the international congress of mathematicians, Warsaw (1983) 1439–1453.

[4] J. Keller, D. Givoli, Exact non-reflecting boundary conditions, J. Comput. Phys. 82 (1989) 172–192.

[5] G. Hsiao, The coupling of boundary element and finite element methods, ZAMM Z.angew.Math.Mech. 70 (1990) 493–503.

[6] D. Givoli, Numerical methods for problems in infinite domains, volume 33, Elsevier Scientific Publishing Co., 1992.

[7] G. Grote, J. Keller, On nonrefleting boundary conditions, J. of Comp. Phys. 122 (1995) 231–243.

[8] D. Givoli, Recent advances in the dtn-fe method, Arch. Comput. Methods Eng. 6 (1999) 71–116.

[9] D. Nicholls, N. Nigma, Exact non-reflecting boundary conditions on general domains, J. Comp. Phys. 194 (2004) 278–303.

[10] D. Nicholls, N. Nigma, Error analysis of an enhanced dtn-fe method for exterior scattering problems, Numer. Math. 105 (2006) 267–298.

[11] L. Chindelevitch, D. Nicholls, N. Nigam, Error analysis and preconditioning for an enhanced dtn-fe algorithm for exterior scattering problems, J. Comput. Appl. Math. 204 (2007) 493–504.

[12] T. Binford, D. Nicholls, N. Nigam, T. Warburtan, Exact non-reflecting boundary conditions on perturbed domains and hp-finite elements, J. Sci. Comput. 39 (2009) 265–292.

[13] B. Engquist, A. Majda, Absorbing boundary conditions for numerical simulation of waves, Math. Comp. 31 (1977) 629–651.

[14] A. Bayliss, E. Turkel, Radiation boundary conditions for wave-like equations, Comm. Pure and Appl. Math. 33 (1980) 707–725.

[15] K. Feng, Asymptotic radiation conditions for reduced wave equation, J. of Comp. Math. 2 (1984) 130–138.

[16] H. Han, X. Wu, Approximation of infinite boundary condition and its application to finite element methods, J. of Comp. Math. 3 (1985) 179–192.

[17] D. Koyama, Error estimates of the dtn finite element method for the exterior helmholtz problem, J. of Comp. Appl. Math. 200 (2007) 21–31.

[18] C. Goldstein, A finite element method for solving helmholtz type equations in waveguides and other unbounded domains, Mathematics of Computation 39 (1982) 309–324.

[19] A. Schatz, An observation concerning ritz-galerkin methods with indefinite bilinear forms, Math. Comp. 28 (1974) 959–962.

[20] G. Hsiao, W. Wendland, Boundary element methods: Foundation and error analysis, in: E. Stein, R. de Borst, T. J. R. Hughes (Eds.), Encyclopedia of Computational Mechanics, volume 1, John Wiley and Sons, Ltd., 2004, pp. 339–373.

[21] S. Brenner, L. Scott, The mathematical theory of finite element methods (second edition), volume 15, Springer-verlag, 2002.

[22] A. Schatz, J. Wang, Some new error estimates for ritz-galerkin methods with minimal regularity assumptions, Math. Comp. 65 (1996) 19–27.

[23] G. Hellwig, Partial differential equations, An introduction, Blaisdell Publishing Company, 1964.

[24] G. Hsiao, W. Wendland, Boundary integral equations, volume 164, Springer-verlag, 2008.

[25] O. G. Ernst, A finite-element capacitance matrix method for exterior helmholtz problems, Numer. Math. 75 (1996) 175–204.

[26] R. Adams, Sobolev spaces, Academic Press, 1975.

[27] J. Lions, E. Magenes, Non-Homogeneous boundary value problems and applications, volume I, Springer-Verlag, 1972.

[28] I. Babuska, S. Sauter, Is the pollution effect of the fem avoidable for the helmholtz equation considering high wave number, SIAM Review 42 (2000) 451–484.

[29] F. Ihlenburg, I. Babuska, Finite element solution of the helmholtz equation with high wave number, part i: The h-version of the fem, Computers Math. Applic. 30 (1995) 9–37.

[30] F. Ihlenburg, I. Babuska, Finite element solution of the helmholtz equation with high wave number, part ii: The h-p version of the fem, SIAM J. Numer. Anal. 34 (1997) 315–358.

[31] F. Ihlenburg, Finite element analysis of acoustic scattering, volume 132, Springer-verlag, 1998.

[32] I. Harari, T. Hughes, Finite element method for the helmholtz equation in exterior domain: Model problems, Comp. Meth. Appl. Mech. Eng. 87 (1991) 59–96.

[33] I. Harari, T. Hughes, Analysis of continuous formulations underlying the computation of time-harmonic acoustics in exterior domains, Comp. Meth. Appl. Mech. Eng. 97 (1992) 103–124.