CSIAM Trans. Appl. Math.
doi: 10.4208/csiam-am.2020-0025

Vol. **1**, No. 3, pp. 530-560
September 2020

A CIP-FEM for High-Frequency Scattering Problem with the Truncated DtN Boundary Condition

Yonglin Li[1], Weiying Zheng[1,2,*] and Xiaopeng Zhu[1,2]

[1] *LSEC, NCMIS, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and System Sciences, Chinese Academy of Sciences, Beijing, 100190, China.*
[2] *School of Mathematical Science, University of Chinese Academy of Sciences, Beijing 100049, China.*

Received 2 June 2020; Accepted 12 September 2020

Abstract. A continuous interior penalty finite element method (CIP-FEM) is proposed to solve high-frequency Helmholtz scattering problem by an impenetrable obstacle in two dimensions. To formulate the problem on a bounded domain, a Dirichlet-to-Neumann (DtN) boundary condition is proposed on the outer boundary by truncating the Fourier series of the original DtN mapping into finite terms. Assuming the truncation order $N \geq kR$, where k is the wave number and R is the radius of the outer boundary, then the H^j-stabilities, $j = 0,1,2$, are established for both original and dual problems, with explicit and sharp estimates of the upper bounds with respect to k. Moreover, we prove that, when $N \geq \lambda kR$ for some $\lambda > 1$, the solution to the DtN-truncation problem converges exponentially to the original scattering problem as N increases. Under the condition that k^3h^2 is sufficiently small, we prove that the preasymptotic error estimates for the linear CIP-FEM as well as the linear FEM are $C_1kh + C_2k^3h^2$. Numerical experiments are presented to validate the theoretical results.

AMS subject classifications: 65N12, 65N30, 78A40, 78A45
Key words: Helmholtz equation, high-frequency, DtN operator, CIP-FEM, wave-number-explicit estimates.

1 Introduction

In this paper, we study the exterior scattering problem of high-frequency acoustic waves by an impenetrable bounded obstacle in two dimensions. To solve the problem numerically, one should truncate the unbounded domain into a bounded one and impose accurate artificial boundary condition on the truncation boundary. For this purpose, some

*Corresponding author. *Email addresses:* liyonglin@lsec.cc.ac.cn (Y. Li), zwy@lsec.cc.ac.cn (W. Zheng), zxp@lsec.cc.ac.cn (X. Zhu)

non-reflecting boundary conditions are proposed, such as the perfectly matched layer (PML) boundary condition [4], the Dirichlet-to-Neumann (DtN) boundary condition or transparent boundary condition (TBC) [12], etc. For scattering problems of moderate frequency, there are extensive studies on the PML in the literature (see e.g. [5, 8, 21]). For TBCs of scattering problems, the authors refer to the recent papers [2, 19, 20, 22].

The DtN boundary condition, or equivalently the DtN operator, is expressed by an infinite trigonometric or spherical harmonic series. In [7], Chandler and Monk studied the high-frequency Helmholtz scattering problem with exact DtN boundary condition. They proved the inf-sup condition for the sesquilinear form $b(\cdot,\cdot)$ of weak formulation with inf-sup constant being $\mathcal{O}(k^{-1})$, where k is the wave number. In practice, the DtN operator must be truncated into the sum of a finite N terms. This leads to an approximate problem with the truncated DtN boundary condition. Obviously, using fewer terms makes the numerical methods more efficient. A major task is to prove the well-posedness of the approximate problem and that the approximate solution u^N converges exponentially to the original solution u as $N\to+\infty$.

Based on numerical experiences and heuristic analysis for Helmholtz eigenvalues on the computational domain, Harari and Hughes suggested to take $N\geq kR$ in the truncated DtN operator to enhance the solvability and accuracy of the approximate problem [14], where R is the radius of the computational domain. In [13], the authors circumvented the difficulty of uniqueness by modifying the truncated DtN operator to eliminate real eigenvalues of the Helmholtz problem. In [15], the authors proved that the approximate solution u^N converges algebraically to the exact solution u provided that $N\geq N_0$ for some N_0 large enough. Recently, Xu and Yin proved that u^N converges to u at an exponential rate q^N for some $0<q<1$ under the same condition that $N\geq N_0$ [29]. Since the papers are focused on moderate-frequency problems, both the factor q and the dependence of N_0 on the wave number k are not given explicitly. In [20], Jiang, Li and Zheng first proposed an adaptive finite element method with truncated DtN operator for solving scattering problems. The a posteriori error estimates consist of the residuals from finite element (FE) discretization and the error estimate from the truncation of the DtN operator. In [19], the authors proved that the error estimate for the DtN truncation decays exponentially as $N\to+\infty$. However, there still is a question left: how do the well-posedness of the truncated problem and the convergence rate of u^N to u depend explicitly on the wave number k? The first objective of the paper is to give an attempt for the answer.

The second objective of the paper is to study the continuous interior penalty finite element method (CIP-FEM) for solving the truncated problem. Preasymptotic FE error estimates will be given by emphasizing their explicit dependence on k. The Helmholtz equation with large wave number is highly indefinite. It is well-known that traditional finite element method (FEM) suffers from the effect of pollution errors [1, 16, 17]. For the Helmholtz equation with exact DtN boundary condition, Melenk and Sauter [26] proved that the linear FE error estimate is $\|u-u_h^{\mathrm{FEM}}\|_{H^1}\leq C_1kh+C_2k^3h^2$ provided that k^2h is sufficiently small. For impedance boundary condition and PML boundary condition, Wu et al. proved the same error estimate by assuming that k^3h^2 is small enough [11, 23, 28, 30].

The first term $\mathcal{O}(kh)$ in the error estimate is of the same order as FE interpolation error and the second term comes from the pollution error. We remark that "asymptotic error estimate" refers to the error estimate without pollution error, while "preasymptotic error estimate" refers to the one with non-negligible pollution effects. In [15, 29], the authors presented a priori error estimates for FEMs using truncated DtN boundary conditions, but did not show their explicit dependence on k. However, the wave-number-explicit estimates are crucial in our work, since we are considering the high-frequency problems. For instance, the frequency of the audio which can be detected by the human ear, denoted by F, falls between $20Hz$ and $20{,}000Hz$. Taking $F=20{,}000Hz$ as an example, the wave number is $k=2\pi F/c\approx 366$, where $c\approx 343m/s$ denotes the speed of sound in the air. To control the FE interpolation error $\mathcal{O}(kh)$, the size of the mesh is expected to be $h=\mathcal{O}(1/k)\approx 2.7\times 10^{-3}$, which means the degree of freedom is $DOF=\mathcal{O}(h^{-2})\approx 1.3\times 10^5$ when the computational domain is a unit square in two dimension. By contrast, to control the pollution error $\mathcal{O}(k^3h^2)$, it's expected that $h=\mathcal{O}(1/k^{1.5})\approx 1.4\times 10^{-4}$, which implies that $DOF\approx 4.9\times 10^7$. Obviously, it's an expensive cost for the ordinary computer and hard to compute when F is high, in other words, k is large. Although there is no strict definition for "high-frequency" in acoustics, we generally regard it as a "high-frequency" problem when the frequency F is higher than $2000Hz$, which means $k>36.6$.

Recently, the CIP-FEM, which was first proposed by Douglas and Dupont for second-order elliptic and parabolic equations [10], has shown great advantages in dealing with high-frequency Helmholtz problems [11, 23, 28, 30]. Such as, the CIP-FEM uses the same approximation space as the traditional continuous FEM but modifies the bilinear form by adding a penalty term at mesh interfaces. It is absolutely stable if the penalty parameters have negative imaginary parts (see [28, 30]). In addition, the error bound of CIP-FEM is not larger than that of the traditional FEM under the same mesh condition. Moreover, the penalty parameters of CIP-FEM can be tuned to reduce the pollution error greatly in practice. For other numerical methods, such as high-order FEMs or discontinuous Galerkin method, we refer to [11, 18, 24–26, 30].

The contributions of the paper are listed as follows.

1. We prove that, under the condition $N\geq kR$, the approximate problem with the truncated DtN boundary condition is well-posed and the associated sesquilinear form satisfies an inf-sup condition with the inf-sup constant being $\mathcal{O}(k^{-1})$.

2. By assuming that the obstacle has C^1-smooth boundary and $N\geq kR$, we prove that the approximate solution satisfies the stabilities $\|u^N\|_{H^j}\leq Ck^{j-1}$, $j=0,1,2$, where the constant C is independent of k.

3. Under the condition $N\geq kR$, we prove the well-posedness of the dual problem with the truncated DtN operator and give the H^2-stability of the solution. We remark that the dual problem plays the key role in estimating both $u-u^N$ and FE errors. However, the well-posedness of the dual problem and the H^2-stability are usually treated as assumptions in the literatures.

4. Under the condition $N \geq 1.7kR$, we prove that u^N converges to u exponentially as $N \to +\infty$. The explicit form of the convergence rate is also presented.

5. We prove the preasymptotic error estimates for the CIP-FEM under the assumption that k^3h^2 is small enough.

The outline of the paper is organized as follows. In Section 2, we introduce the model problem and its weak formulation by using the DtN operator. In Section 3, we give some wave-number-explicit analyses for the truncated DtN problem and prove the exponential convergence of u^N to u. Section 4 is devoted to the preasymptotic error estimates for the linear CIP-FEM as well as FEM. In Section 5, we use some numerical examples to verify the theoretical results.

Throughout this paper, we adopt conventional notations and definitions for Sobolev spaces, norms, and inner products (see e.g. [6]). For any open domain $D \subset \mathbb{R}^2$, $L^2(D)$ denotes the usual Hilbert space of square-integrable complex functions. The inner product and norm on $L^2(D)$ are defined by

$$(u,v)_D := \int_D u(\boldsymbol{x})\overline{v}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}, \quad \|u\|_{L^2(D)} := (u,u)_D^{1/2}.$$

For a Lipschitz-continuous curve Γ, $\langle\cdot,\cdot\rangle_\Gamma$ denotes the L^2-inner product on $L^2(\Gamma)$ or the duality paring between two dual spaces defined on Γ. We shall use C to denote a generic positive constant which is independent of the sensitive quantities, such as the mesh size h, the wave number k, the truncation number N, and the penalty parameters γ_e. We also use the shorthand notations $A \lesssim B$, $B \gtrsim A$ for the inequality $A \leq CB$ and $B \geq CA$, respectively. Moreover, $A \eqsim B$ means that $A \lesssim B$ and $B \gtrsim A$ hold simultaneously. We shall use boldface notations for vector-valued quantities, such as $\boldsymbol{L}^2(\Omega) = (L^2(\Omega))^2$.

2 Variational formulation using the DtN operator

In this section, we first introduce the model problem of acoustic scattering. Next we reformulate the problem on a truncated domain by using the DtN operator. For any $r > 0$, let B_r denote the ball with center at the origin and radius being r and let $\Gamma_r = \partial B_r$ be its boundary.

2.1 The model problem

We consider acoustic scattering by a bounded, sound soft obstacle occupying a compact set $\Omega \in \mathbb{R}^2$ with C^1-smooth boundary Γ, as seen in Fig. 1. To favor the FE error analysis in Section 4, we also assume that Γ is piecewise C^2-smooth. Suppose Ω is starlike with respect to the origin. The problem is to seek for u which satisfies the Helmholtz equation

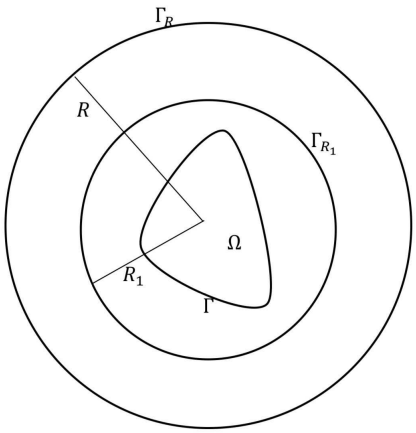

Figure 1: Geometry of the obstacle scattering problem.

with Sommerfeld radiation condition

$$\Delta u + k^2 u = f \quad \text{in } \Omega_e := \mathbb{R}^2 \setminus \Omega, \tag{2.1a}$$

$$u = 0 \quad \text{on } \Gamma, \tag{2.1b}$$

$$\frac{\partial u}{\partial r} - \mathbf{i}ku = o(r^{-\frac{1}{2}}) \quad \text{as } r := |\boldsymbol{x}| \to \infty, \tag{2.1c}$$

where $\mathbf{i} = \sqrt{-1}$ denotes the imaginary unit and k is the wave number. We assume that $f \in L^2(\Omega_e)$ and has compact support.

2.2 The DtN operator

We choose $R > R_1 > R_0 := \sup_{\boldsymbol{x}\in\Omega} |\boldsymbol{x}|$ such that $\text{supp}(f) \subset \{\boldsymbol{x} : |\boldsymbol{x}| < R_1\}$ and define

$$D_R := B_R \cap \Omega_e, \quad D_{R_1} := B_{R_1} \cap \Omega_e.$$

For convenience, we define $V_R := \{w \in H^1(D_R) : w = 0 \text{ on } \Gamma\}$ and let it be equipped with the equivalent H^1-norm

$$\|w\|_{V_R} = \left(\|\nabla w\|^2_{L^2(D_R)} + k^2 \|w\|^2_{L^2(D_R)} \right)^{1/2}. \tag{2.2}$$

Since $\{e^{\mathbf{i}m\theta} : m \in \mathbb{Z}\}$ provides a complete and orthogonal basis of $L^2(\Gamma_R)$, any function $w \in L^2(\Gamma_R)$ admits the Fourier expansion

$$w(\boldsymbol{x}) = \sum_{m\in\mathbb{Z}} w_m e^{\mathbf{i}m\theta}, \quad w_m = \frac{1}{2\pi} \int_0^{2\pi} w(R\cos\theta, R\sin\theta) e^{-\mathbf{i}m\theta} \mathrm{d}\theta.$$

The well-known Parseval equality yields

$$\|w\|^2_{L^2(\Gamma_R)} = 2\pi R \sum_{m\in\mathbb{Z}} |w_m|^2.$$

The trace space $H^s(\Gamma_R)$ is defined by

$$H^s(\Gamma_R)=\{w\in L^2(\Gamma_R):\|w\|_{H^s(\Gamma_R)}<\infty\},$$

where the $\|\cdot\|_{H^s(\Gamma_R)}$ is given by

$$\|w\|_{H^s(\Gamma_R)}=\left(2\pi R\sum_{m\in\mathbb{Z}}(1+m^2)^s|w_m|^2\right)^{1/2}.$$

The DtN operator $\mathcal{T}$: $H^{1/2}(\Gamma_R)\to H^{-1/2}(\Gamma_R)$ is defined by

$$\forall w=\sum_{m\in\mathbb{Z}}w_m e^{\mathbf{i}m\theta}\in H^{1/2}(\Gamma_R),\quad \mathcal{T}w:=\sum_{m\in\mathbb{Z}}\frac{kH_m^{(1)'}(kR)}{H_m^{(1)}(kR)}w_m e^{\mathbf{i}m\theta},\tag{2.3}$$

where $H_m^{(1)}(\cdot)$ is the Hankel function of the first kind of order m.

Let $\langle\cdot,\cdot\rangle_{\Gamma_R}$ denote the duality pairing between $H^{1/2}(\Gamma_R)$ and $H^{-1/2}(\Gamma_R)$. For any $v=\sum_{m\in\mathbb{Z}}v_m e^{\mathbf{i}m\theta}\in H^{1/2}(\Gamma_R)$, we have

$$\langle\mathcal{T}w,v\rangle_{\Gamma_R}=2\pi\sum_{m\in\mathbb{Z}}h_m(kR)w_m\overline{v}_m,\quad\text{where } h_m(z)=z\frac{H_m^{(1)'}(z)}{H_m^{(1)}(z)}.$$

From (3.4a) of [26], $\mathcal{T}$ is bounded and satisfies

$$|\langle\mathcal{T}w,v\rangle_{\Gamma_R}|\le C\|w\|_{V_R}\|v\|_{V_R}\quad\forall w,v\in V_R,\tag{2.4}$$

where $C>0$ is independent of k and R.

2.3 The weak formulation

In the exterior domain $\Omega_e\backslash\overline{B_R}$, the solution of the Helmholtz equation (2.1) can be written as a Fourier series in polar coordinates:

$$u(r,\theta)=\sum_{m\in\mathbb{Z}}\frac{H_m^{(1)}(kr)}{H_m^{(1)}(kR)}u_m(R)e^{\mathbf{i}m\theta},\quad r>R,\tag{2.5}$$

where

$$u_m(R)=\frac{1}{2\pi}\int_0^{2\pi}u(R\cos\theta,R\sin\theta)e^{-\mathbf{i}m\theta}\mathrm{d}\theta.$$

From (2.3) and (2.5), it is easy to see that

$$\frac{\partial u}{\partial r}=\mathcal{T}u\quad\text{on }\Gamma_R.$$

Therefore, we can reformulate (2.1) on the truncated domain D_R as follows

$$\Delta u + k^2 u = f \quad \text{in } D_R, \tag{2.6a}$$
$$u = 0 \quad \text{on } \Gamma, \tag{2.6b}$$
$$\frac{\partial u}{\partial \boldsymbol{n}} = \mathcal{T} u \quad \text{on } \Gamma_R, \tag{2.6c}$$

where $\boldsymbol{n}$ is the unit outer normal on Γ_R.

The weak formulation of (2.6) is proposed as follows: find $u \in V_R$ such that

$$b(u,v) = -(f,v)_{D_R} \quad \forall v \in V_R, \tag{2.7}$$

where b is the bounded sesquilinear form on V_R and is defined by

$$b(u,v) := \int_{D_R} (\nabla u \cdot \nabla \overline{v} - k^2 u \overline{v}) - \langle \mathcal{T} u, v \rangle_{\Gamma_R}.$$

From [7], b satisfies the inf-sup condition

$$\inf_{0 \neq u \in V_R} \sup_{0 \neq v \in V_R} \frac{|b(u,v)|}{\|u\|_{V_R} \|v\|_{V_R}} \gtrsim \frac{1}{k}. \tag{2.8}$$

This yields the well-posedness of (2.7).

3 The approximate problem with truncated DtN operator

Note from (2.3) that the DtN operator $\mathcal{T}$ is defined as an infinite series. Practically, it is necessary to truncate the nonlocal operator by taking finitely many terms of the expansion so as to attain a feasible algorithm.

3.1 The truncated DtN operator

Given a sufficiently large integer N, we define the truncated DtN operator

$$\mathcal{T}_N : H^{1/2}(\Gamma_R) \to H^{-1/2}(\Gamma_R)$$

as follows

$$\forall w = \sum_{m \in \mathbb{Z}} w_m e^{\mathbf{i} m\theta} \in H^{1/2}(\Gamma_R), \quad \mathcal{T}_N w := R^{-1} \sum_{|m| \leq N} h_m(kR) w_m e^{\mathbf{i} m\theta}. \tag{3.1}$$

Following the proof of [26, eq. (3.4a)], but only taking finite terms of the expansion of $\mathcal{T}$, we can easily prove that the truncated operator $\mathcal{T}_N$ is also bounded and satisfies

$$|\langle \mathcal{T}_N w, v \rangle| \leq C \|w\|_{V_R} \|v\|_{V_R} \quad \forall w, v \in V_R, \tag{3.2}$$

where the constant C is independent of k and R.

3.2 The approximate problem

Based on $\mathcal{T}_N$, an approximate problem of (2.7) is to seek for $u^N \in V_R$ such that

$$b_N(u^N,v) = -(f,v)_{D_R} \quad \forall v \in V_R, \tag{3.3}$$

where the sesquilinear form $b_N : V_R \times V_R \to \mathbb{C}$ is defined by

$$b_N(u^N,v) := \int_{D_R} (\nabla u^N \nabla \overline{v} - k^2 u^N \overline{v}) dx - \langle \mathcal{T}_N u^N, v \rangle_{\Gamma_R}. \tag{3.4}$$

It is clearly that u^N satisfies

$$\Delta u^N + k^2 u^N = f \quad \text{in } D_R, \tag{3.5a}$$

$$u^N = 0 \quad \text{on } \Gamma, \tag{3.5b}$$

$$\frac{\partial u^N}{\partial \boldsymbol{n}} = \mathcal{T}_N u^N \quad \text{on } \Gamma_R. \tag{3.5c}$$

The continuity of b_N is apparent:

$$|b_N(w,v)| \le \int_{D_R} \left(|\nabla w \cdot \nabla \overline{v}| + k^2 |w\overline{v}| \right) + |\langle \mathcal{T}_N w, v \rangle_{\Gamma_R}| \le C \|w\|_{V_R} \|v\|_{V_R}.$$

Moreover, from (3.4) we also have

$$\begin{aligned} &\mathrm{Re}\, b_N(w,w) = \|w\|^2_{V_R} - 2k^2 \|w\|^2_{L^2(D_R)} - \mathrm{Re}\langle \mathcal{T}_N w, w \rangle_{\Gamma_R}, \\ &\mathrm{Im}\, b_N(w,w) = -\mathrm{Im}\langle \mathcal{T}_N w, w \rangle_{\Gamma_R}. \end{aligned}$$

To prove the well-posedness of (3.3), we still need the inf-sup condition for b_N which will be studied in the last part of this section.

3.3 Some useful estimates for $\mathcal{T}_N$

First we prove some preliminary but useful results for the Hankel functions. The results will be used in the estimation of $\mathcal{T}_N$.

Lemma 3.1. *Assume that $kR \ge 1$. For any $m \in \mathbb{Z}$, there holds*

$$\left| \frac{H_m^{(1)'}(kR)}{H_m^{(1)}(kR)} \right| \le \frac{|m|}{kR} + 2. \tag{3.6}$$

Proof. First we consider the case of $|m| \ge 1$. The modified Hankel functions satisfy

$$\frac{2}{\pi} K_m(t) = \begin{cases} \mathbf{i}^{m+1} H_m^{(1)}(\mathbf{i}t), & -\pi \le \arg t \le \pi/2, \\ (-\mathbf{i})^{m+1} H_m^{(2)}(-\mathbf{i}t), & \pi/2 \le \arg t \le \pi. \end{cases} \tag{3.7}$$

From [27, eq. (10.29.2)], we have the recurrence relation $tK'_m(t) = -mK_m(t)+tK_{m-1}(t)$. This yields

$$\left|\frac{H_m^{(1)'}(kR)}{H_m^{(1)}(kR)}\right| = \left|\frac{H_{|m|}^{(1)'}(kR)}{H_{|m|}^{(1)}(kR)}\right| \le \left|\frac{K'_{|m|}(-\mathbf{i}kR)}{K_{|m|}(-\mathbf{i}kR)}\right| \le \frac{|m|}{kR}+\left|\frac{K_{|m|-1}(-\mathbf{i}kR)}{K_{|m|}(-\mathbf{i}kR)}\right| \le \frac{|m|}{kR}+1,$$

where the last inequality follows from the monotonicity of $K_m(t)$ in m (cf. [27, eq. (10.37.1)]).

If $m=0$, from [27, eqs. (10.17.13)–(10.17.15)] and the condition $kR\ge 1$, we have

$$\left|\frac{H_0^{(1)'}(kR)}{H_0^{(1)}(kR)}\right| = \left|\frac{H_1^{(1)}(kR)}{H_0^{(1)}(kR)}\right| \le \frac{4kR+3e^{3/(4kR)}}{4kR+e^{1/(4kR)}} \le 2.$$

This finishes the proof. □

Since we are considering the high-frequency problems, the condition $kR\ge 1$ is naturally. In the rest of this paper, we always assume that $kR\ge 1$ and will not emphasize it again.

Lemma 3.2. *Let $\lambda>1$ be a given constant. For any $t>0$ and $m\ge\max\left(\lambda t, \frac{5\lambda^2(4\lambda^2+1)}{4(\lambda^2-1)^3}+1\right)$,*

$$\frac{1}{2}\Theta(t/m) \le \left|H_m^{(1)}(t)\right| \le \frac{1.5}{(1-\lambda^{-2})^{1/4}}\Theta(t/m),$$

where

$$\Theta(s) := \sqrt{\frac{2}{\pi m}}\left(\frac{se^{\sqrt{1-s^2}}}{1+\sqrt{1-s^2}}\right)^{-m}.$$

Proof. Write $s := t/m$ for convenience. The assumptions imply $s\le 1/\lambda<1$. From [27, eq. (10.27.8)] we know that

$$H_m^{(1)}(t) = \frac{2}{\pi}(-\mathbf{i})^{m+1}K_m(-\mathbf{i}t) = \frac{2}{\pi}(-\mathbf{i})^{m+1}K_m(-\mathbf{i}ms).$$

From [3, Lemma 5.1], we have

$$\left|H_m^{(1)}(t)\right| = \frac{2}{\pi}|K_m(-\mathbf{i}ms)| = \frac{\Theta(s)}{(1-s^2)^{1/4}}|1+\eta_2(m,s)|,$$

where η_2 is called "an error term" and is estimated in [3, Lemma 5.2] as follows

$$|\eta_2(m,s)| \le \frac{M(s)}{m}e^{M(s)/m},$$

where

$$M(s)=\frac{1}{6}+\frac{1}{3\sqrt{5}}+\frac{\pi s^2(4+s^2)}{8(1-s^2)^3}\leq\frac{1}{3}+\frac{\pi\lambda^2(4\lambda^2+1)}{8(\lambda^2-1)^3}<\frac{m}{3}.$$

We infer that $|\eta_2(m,s)|<\frac{1}{3}e^{1/3}<1/2$. The proof is completed by using $1-s^2\geq 1-\lambda^{-2}$. □

Lemma 3.3. *Suppose $N\geq kR$. For any $w\in H^{1/2}(\Gamma_R)$, there hold*

$$\mathrm{Re}\langle\mathcal{T}_N w,w\rangle_{\Gamma_R}\leq 0,\ \mathrm{Im}\langle\mathcal{T}_N w,w\rangle_{\Gamma_R}\geq 0, \tag{3.8}$$

$$\mathrm{Re}\langle\mathcal{T}_N w,w\rangle_{\Gamma_R}+R\int_{\Gamma_R}\left(k^2|w|^2-|\nabla_T w|^2+|\mathcal{T}_N w|^2\right)\leq 2kR\,\mathrm{Im}\langle\mathcal{T}_N w,w\rangle_{\Gamma_R}, \tag{3.9}$$

where $\nabla_T w$ stands for the tangential gradient of w on Γ_R.

Proof. The proof is similar to that of [7, Lemma 2.1] except that w is an arbitrary function in $H^{1/2}(\Gamma_R)$ here. First we expand w into Fourier series

$$w(R,\theta)=\sum_{m\in\mathbb{Z}} w_m e^{\mathbf{i}m\theta}.$$

For convenience, we write

$$\rho=kR,\quad M_m(\rho)=\left|H_m^{(1)}(\rho)\right|,\quad N_m(\rho)=\left|H_m^{(1)'}(\rho)\right|,\quad C_m(\rho)=\frac{|w_m|^2}{M_m^2(\rho)}.$$

By (3.1), we find that

$$\begin{aligned}
\langle\mathcal{T}_N w,w\rangle_{\Gamma_R}&=2\pi\rho\sum_{|m|\leq N}\frac{H_m^{(1)'}(\rho)}{H_m^{(1)}(\rho)}|w_m|^2\\
&=2\pi\rho\sum_{|m|\leq N}\frac{H_m^{(1)'}(\rho)\overline{H_m^{(1)}(\rho)}}{\left|H_m^{(1)}(\rho)\right|^2}|w_m|^2\\
&=2\pi\rho\sum_{|m|\leq N}\left[\mathrm{Re}\left(H_m^{(1)'}(\rho)\overline{H_m^{(1)}(\rho)}\right)+\mathbf{i}\left(J_m(\rho)Y_m'(\rho)-J_m'(\rho)Y_m(\rho)\right)\right]C_m\\
&=\sum_{|m|\leq N}\left[\pi\rho\frac{\mathrm{d}M_m^2(\rho)}{\mathrm{d}\rho}+4\mathbf{i}\right]C_m,
\end{aligned}$$

where in the last identity we have used the Wronskian formula (see [27, eq. (10.5.2)]):

$$\pi\rho\left(J_m(\rho)Y_m'(\rho)-J_m'(\rho)Y_m(\rho)\right)=2. \tag{3.10}$$

Since $M_m(\rho)$ is decreasing on the positive real axis (see [7]), we obtain (3.8).

Next we prove (3.9). Since $|\nabla_T w| = \frac{1}{R}\left|\frac{\partial w}{\partial \theta}\right|$, we deduce that

$$\begin{aligned}
&R\int_{\Gamma_R}\left(k^2|w|^2-|\nabla_T w|^2+|\mathcal{T}_N w|^2\right)\\
=&2\pi\sum_{m\in\mathbb{Z}}(\rho^2-m^2)|w_m|^2+2\pi\rho^2\sum_{|m|\le N}N_m^2(\rho)C_m\\
=&2\pi\sum_{m\in\mathbb{Z}}M_m^2(\rho)(\rho^2-m^2)C_m+2\pi\rho^2\sum_{|m|\le N}N_m^2(\rho)C_m.
\end{aligned}$$

It suffices to prove

$$\sum_{|m|\le N}\left(\frac{\rho}{2}\frac{\mathrm{d}M_m^2}{\mathrm{d}\rho}+(\rho^2-m^2)M_m^2+\rho^2N_m^2-\frac{4\rho}{\pi}\right)C_m+\sum_{|m|>N}(\rho^2-m^2)M_m^2C_m\le 0.$$

The second term on the left-hand side is clearly negative since $\rho = kR \le N$. From [7, eq. (2.13)], we know that

$$\frac{\rho}{2}\frac{\mathrm{d}M_m^2}{\mathrm{d}\rho}+(\rho^2-m^2)M_m^2+\rho^2N_m^2-\frac{4\rho}{\pi}\le 0\quad\forall\rho>0.$$

The first term on the left-hand side is also negative. This proves (3.9). □

3.4 The inf-sup condition for b_N

The purpose of this subsection is to prove the inf-sup condition for b_N. It yields the well-posedness of problem (3.3) naturally. For any $\xi\in L^2(D_R)$, we consider the problem

$$\Delta w+k^2w=\xi\quad\text{in } D_R, \tag{3.11a}$$

$$w=0\quad\text{on }\Gamma, \tag{3.11b}$$

$$\frac{\partial w}{\partial \boldsymbol{n}}=\mathcal{T}_N w\quad\text{on }\Gamma_R. \tag{3.11c}$$

Its weak formulation is to seek for $w\in V_R$ such that

$$b_N(w,v)=-(\xi,v)_{D_R}\quad\forall v\in V_R. \tag{3.12}$$

First we prove the stability of the solution to (3.12).

Lemma 3.4. *If $N\ge kR$, there holds*

$$\|w\|_{V_R}\lesssim\|\xi\|_{L^2(D_R)}. \tag{3.13}$$

Proof. Integrating in D_R the two identities

$$\begin{aligned}
&\nabla\cdot(\boldsymbol{x}|w|^2)=2|w|^2+(\boldsymbol{x}\cdot\nabla w)\overline{w}+w(\boldsymbol{x}\cdot\nabla\overline{w}),\\
&\nabla\cdot(\boldsymbol{x}|\nabla w|^2)=\nabla(\boldsymbol{x}\cdot\nabla w)\cdot\nabla\overline{w}+\nabla w\cdot\nabla(\boldsymbol{x}\cdot\nabla\overline{w}),
\end{aligned}$$

we get

$$2\|w\|^2_{L^2(D_R)}+2\mathrm{Re}(w,\boldsymbol{x}\cdot\nabla w)_{D_R}=R\|w\|^2_{L^2(\Gamma_R)}, \tag{3.14}$$

$$2\mathrm{Re}(\nabla w,\nabla(\boldsymbol{x}\cdot\nabla w))_{D_R}=R\|\nabla w\|^2_{\boldsymbol{L}^2(\Gamma_R)}-\int_\Gamma \boldsymbol{x}\cdot\boldsymbol{n}|\nabla w|^2. \tag{3.15}$$

Multiplying both sides of (3.11a) by $v=\boldsymbol{x}\cdot\nabla w$ and integrating by parts in D_R, we obtain

$$b_N(w,\boldsymbol{x}\cdot\nabla w)+\int_\Gamma \frac{\partial w}{\partial \boldsymbol{n}}(\boldsymbol{x}\cdot\nabla\overline{w})=-(\xi,\boldsymbol{x}\cdot\nabla w)_{D_R}, \tag{3.16}$$

where the unit normal $\boldsymbol{n}$ on Γ points to Ω_e. Taking the real part of (3.16) and using (3.14)-(3.15), we have

$$\begin{aligned}&2k^2\|w\|^2_{L^2(D_R)}+R\|\nabla w\|^2_{\boldsymbol{L}^2(\Gamma_R)}+2\mathrm{Re}\int_\Gamma \frac{\partial w}{\partial n}(\boldsymbol{x}\cdot\nabla\overline{w})\\ =&\int_\Gamma \boldsymbol{x}\cdot\boldsymbol{n}|\nabla w|^2+2\mathrm{Re}\langle\mathcal{T}_N w,\boldsymbol{x}\cdot\nabla w\rangle_{\Gamma_R}+k^2R\|w\|^2_{L^2(\Gamma_R)}-2\mathrm{Re}(\xi,\boldsymbol{x}\cdot\nabla w)_{D_R}.\end{aligned} \tag{3.17}$$

Since $\nabla w=\frac{\partial w}{\partial \boldsymbol{n}}\boldsymbol{n}+\nabla_T w$ on $\Gamma\cup\Gamma_R$ and $w=0$ on Γ, we have $\nabla_T w=0$ on Γ. Then

$$\mathrm{Re}\int_\Gamma \frac{\partial w}{\partial \boldsymbol{n}}(\boldsymbol{x}\cdot\nabla\overline{w})=\int_\Gamma \boldsymbol{x}\cdot\boldsymbol{n}\left|\frac{\partial w}{\partial \boldsymbol{n}}\right|^2=\int_\Gamma \boldsymbol{x}\cdot\boldsymbol{n}|\nabla w|^2.$$

Similarly, since $\boldsymbol{x}\cdot\nabla w=R\frac{\partial w}{\partial \boldsymbol{n}}=R\mathcal{T}_N w$ on Γ_R, we have

$$\mathrm{Re}\langle\mathcal{T}_N w,\boldsymbol{x}\cdot\nabla w\rangle_{\Gamma_R}=R\left\|\frac{\partial w}{\partial \boldsymbol{n}}\right\|^2_{L^2(\Gamma_R)}=R\|\mathcal{T}_N w\|^2_{L^2(\Gamma_R)}.$$

Inserting the above equalities into (3.17) yields

$$\begin{aligned}2k^2\|w\|^2_{L^2(D_R)}=&-\int_\Gamma \boldsymbol{x}\cdot\boldsymbol{n}|\nabla w|^2+R\int_{\Gamma_R}\left(k^2|w|^2-|\nabla_T w|^2+|\mathcal{T}_N w|^2\right)\\ &-2\mathrm{Re}(\xi,\boldsymbol{x}\cdot\nabla w)_{D_R}.\end{aligned} \tag{3.18}$$

Taking $v=w$ in (3.12) shows

$$\|\nabla w\|^2_{L^2(D_R)}-k^2\|w\|^2_{L^2(D_R)}-\langle\mathcal{T}_N w,w\rangle_{\Gamma_R}=-(\xi,w)_{D_R}. \tag{3.19}$$

Summing (3.18) and the real part of (3.19), we get

$$\begin{aligned}\|w\|^2_{V_R}=&-2\mathrm{Re}(\xi,\boldsymbol{x}\cdot\nabla w)_{D_R}-\mathrm{Re}(\xi,w)_{D_R}+R\int_{\Gamma_R}\left(k^2|w|^2-|\nabla_T w|^2+|\mathcal{T}_N w|^2\right)\\ &-\int_\Gamma \boldsymbol{x}\cdot\boldsymbol{n}\left|\frac{\partial w}{\partial n}\right|^2+\mathrm{Re}\langle\mathcal{T}_N w,w\rangle_{\Gamma_R}.\end{aligned}$$

Since Ω is starlike, $\boldsymbol{x}\cdot\boldsymbol{n}>0$ on Γ. From Lemma 3.3 and taking the imaginary part of (3.19), we find that

$$\begin{aligned}\|w\|_{V_R}^2 &\leq -2\mathrm{Re}(\xi,\boldsymbol{x}\cdot\nabla w)_{D_R}-\mathrm{Re}(\xi,w)_{D_R}+2kR\mathrm{Im}\langle\mathcal{T}_N w,w\rangle_{\Gamma_R}\\ &\leq -2\mathrm{Re}(\xi,\boldsymbol{x}\cdot\nabla w)_{D_R}-\mathrm{Re}(\xi,w)_{D_R}+2kR\mathrm{Im}(\xi,w)_{D_R}\\ &\leq C\|\xi\|_{L^2(D_R)}^2+\frac{1}{2}\|w\|_{V_R}^2. \end{aligned}\tag{3.20}$$

This completes the proof. □

Lemma 3.5. *For any $\xi\in L^2(D_R)$, if $N\geq kR$, problem* (3.11) *has a unique solution.*

Proof. From (3.1) and Lemma 3.1, it is easy to see that

$$\begin{aligned}\|\mathcal{T}_N g\|_{H^{-1/2}(\Gamma_R)} &= R^{-1}\left(2\pi R\sum_{|m|\leq N}(1+m^2)^{-1/2}|h_m(kR)|^2|g_m|^2\right)^{1/2}\\ &\leq R^{-1}(N+2kR)\left(2\pi R\sum_{|m|\leq N}|g_m|^2\right)^{1/2}\\ &= R^{-1}(N+2kR)\|g\|_{L^2(\Gamma_R)}\end{aligned}$$

for any $g=\sum_{m\in\mathbb{Z}}g_m e^{\mathbf{i}m\theta}\in H^{1/2}(\Gamma_R)$. Since $H^{1/2}(\Gamma_R)$ is embedded compactly into $L^2(\Gamma_R)$, $\mathcal{T}_N$: $H^{1/2}(\Gamma_R)\to H^{-1/2}(\Gamma_R)$ is a compact operator. Therefore, (3.11) is a Fredholm-type problem. Lemma 3.4 actually proves the uniqueness of the solution w to (3.11). The existence of w follows from its uniqueness. □

Theorem 3.1. *If $N\geq kR$, the sesquilinear form b_N admits the inf-sup condition*

$$\|w\|_{V_R}\lesssim k\sup_{0\neq v\in V_R}\frac{|b_N(w,v)|}{\|v\|_{V_R}}\quad\forall w\in V_R.\tag{3.21}$$

Proof. For any $\xi,\eta\in H^{1/2}(\Gamma_R)$, suppose their Fourier expansions are given by

$$\xi=\sum_{m\in\mathbb{Z}}\xi_m e^{\mathbf{i}m\theta},\quad \eta=\sum_{m\in\mathbb{Z}}\eta_m e^{\mathbf{i}m\theta}.$$

From (3.1), it is easy to see

$$\langle\mathcal{T}_N\xi,\eta\rangle_{\Gamma_R}=2\pi\sum_{|m|\leq N}h_m(kR)\xi_m\overline{\eta}_m=\langle\mathcal{T}_N\overline{\eta},\overline{\xi}\rangle_{\Gamma_R}.\tag{3.22}$$

For any $w\in V_R$, we consider the dual problem of (3.11): find $\phi\in V_R$ such that

$$(\nabla v,\nabla\phi)_{D_R}-k^2(v,\phi)_{D_R}-\langle\mathcal{T}_N v,\phi\rangle_{\Gamma_R}=2k^2(v,w)_{D_R}\quad\forall v\in V_R.\tag{3.23}$$

Using (3.22), the complex conjugate of ϕ satisfies

$$(\nabla\overline{\phi},\nabla\overline{v})_{D_R}-k^2(\overline{\phi},\overline{v})_{D_R}-\langle\mathcal{T}_N\overline{\phi},\overline{v}\rangle_{\Gamma_R}=2k^2(\overline{w},\overline{v})_{D_R}\quad\forall v\in V_R. \tag{3.24}$$

By the arbitrariness of v, $\overline{\phi}\in V_R$ is the solution to the problem

$$b_N(\overline{\phi},v)=2k^2(\overline{w},v)_{D_R}\quad\forall v\in V_R.$$

From Lemmas 3.4 and 3.5, $\overline{\phi}$ exists uniquely and satisfies

$$\|\phi\|_{V_R}\lesssim 2k^2\|w\|_{L^2(D_R)}\lesssim k\|w\|_{V_R}. \tag{3.25}$$

Now using $\mathrm{Re}\langle\mathcal{T}_N w,w\rangle\leq 0$ and Eq. (3.23) with $v=w$, we have

$$\|w\|_{V_R}^2\leq 2k^2\|w\|_{L^2(D_R)}^2+\mathrm{Re}b_N(w,w)=\mathrm{Re}b_N(w,w+\phi).$$

Note from (3.25) that $\|w+\phi\|_{V_R}\lesssim k\|w\|_{V_R}$. Then the inf-sup condition follows directly

$$\|w\|_{V_R}\leq\frac{\mathrm{Re}b_N(w,w+\phi)}{\|w\|_{V_R}}\lesssim k\frac{|b_N(w,w+\phi)|}{\|w+\phi\|_{V_R}}\lesssim k\sup_{0\neq v\in V_R}\frac{|b_N(w,v)|}{\|v\|_{V_R}}.$$

The proof is finished. □

3.5 Convergence rate of u^N

Now we are going to prove that the approximate solution u^N converges to the scattering solution u exponentially as N increases. The Fourier series of u is given as follows

$$u(r,\theta)=\sum_{m\in\mathbb{Z}}u_m(r)e^{\mathbf{i}m\theta},\quad u_m(r)=\frac{1}{2\pi}\int_0^{2\pi}u(r\cos\theta,r\sin\theta)e^{-\mathbf{i}m\theta}d\theta\quad\forall r>R_0. \tag{3.26}$$

Lemma 3.6. *Let $\lambda>1$ be a constant and $N\geq\max\left(\lambda kR,\frac{5\lambda^2(4\lambda^2+1)}{4(\lambda^2-1)^3}+1\right)$. Then*

$$|u_m(R)|\leq\frac{3e^{\frac{kR}{2\sqrt{\lambda^2-1}}}}{(1-\lambda^{-2})^{1/4}}\left(\frac{R_1}{R}\right)^{|m|}|u_m(R_1)|\quad\forall m\in\mathbb{Z}\ \textit{satisfying}\ |m|\geq N. \tag{3.27}$$

Proof. Comparing (3.26) with (2.5), we obtain the relationship between $u_m(R)$ and $u_m(R_1)$

$$u_m(R)=\frac{H_m^{(1)}(kR)}{H_m^{(1)}(kR_1)}u_m(R_1). \tag{3.28}$$

It suffices to estimate $B_m:=\left|H_m^{(1)}(kR)/H_m^{(1)}(kR_1)\right|$. Since $B_{-m}=B_m$, we only consider the case of $m>0$ without loss of generality.

From Lemma 3.2, for $m>N\geq\max\left(\lambda kR,\frac{5\lambda^2(4\lambda^2+1)}{4(\lambda^2-1)^3}+1\right)$, we have

$$\frac{1}{2}\Theta(kR/m)\leq\left|H_m^{(1)}(kR)\right|\leq\frac{1.5}{(1-\lambda^{-2})^{1/4}}\Theta(kR/m), \tag{3.29}$$

where

$$\Theta(s):=\sqrt{\frac{2}{\pi m}}\left(\frac{se^{\sqrt{1-s^2}}}{1+\sqrt{1-s^2}}\right)^{-m}.$$

Direct calculations show that

$$\frac{\Theta(kR/m)}{\Theta(kR_1/m)}\leq\left(\frac{R_1}{R}\right)^m e^{\frac{kR[1-(R_1/R)^2]}{\sqrt{(m/(kR))^2-(R_1/R)^2}+\sqrt{(m/kR)^2-1}}}\leq\left(\frac{R_1}{R}\right)^m e^{\frac{kR}{2\sqrt{\lambda^2-1}}}. \tag{3.30}$$

Using (3.29) and (3.30), an upper bound of B_m is estimated as follows

$$B_m\leq\frac{3}{(1-\lambda^{-2})^{1/4}}\frac{\Theta(kR/m)}{\Theta(kR_1/m)}\leq\frac{3e^{\frac{kR}{2\sqrt{\lambda^2-1}}}}{(1-\lambda^{-2})^{1/4}}\left(\frac{R_1}{R}\right)^m.$$

The proof is finished. □

Theorem 3.2. *Suppose $N\geq1.7kR$ and $R\geq2R_1$. Let u and u^N be the solutions to* (2.7) *and* (3.3)*, respectively. Then*

$$\|u-u^N\|_{V_R}\lesssim k\left(\frac{R_1}{R}\right)^{2N/3}\|f\|_{L^2(D_{R_1})}. \tag{3.31}$$

Proof. By the inf-sup condition (3.21), we have

$$\|u-u^N\|_{V_R}\lesssim k\sup_{0\neq v\in V_R}\frac{|b_N(u-u^N,v)|}{\|v\|_{V_R}}=k\sup_{0\neq v\in V_R}\frac{|\langle(\mathcal{T}-\mathcal{T}_N)u,v\rangle_{\Gamma_R}|}{\|v\|_{V_R}}. \tag{3.32}$$

Using Lemma 3.6 and $(1-\lambda^{-2})\gtrsim1$, and the estimate in Lemma 3.1, we have

$$\begin{aligned}
|\langle(\mathcal{T}-\mathcal{T}_N)u,v\rangle_{\Gamma_R}|&=\left|2\pi\sum_{|m|>N}h_m(kR)u_m(R)\overline{v}_m(R)\right|\\
&\lesssim e^{\frac{kR}{2\sqrt{\lambda^2-1}}}\sum_{|m|>N}\left(\frac{R_1}{R}\right)^m(|m|+2kR)\,|u_m(R_1)|\,|\overline{v}_m(R)|\\
&\lesssim e^{\frac{kR}{2\sqrt{\lambda^2-1}}}\left(\frac{R_1}{R}\right)^N\sum_{|m|>N}\sqrt{1+m^2}\,|u_m(R_1)|\,|\overline{v}_m(R)|\\
&\lesssim e^{\frac{kR}{2\sqrt{\lambda^2-1}}}\left(\frac{R_1}{R}\right)^N\|u\|_{H^{1/2}(\Gamma_{R_1})}\|v\|_{H^{1/2}(\Gamma_R)}\\
&\lesssim e^{\frac{kR}{2\sqrt{\lambda^2-1}}}\left(\frac{R_1}{R}\right)^N\|u\|_{V_R}\|v\|_{V_R}.
\end{aligned} \tag{3.33}$$

We choose $\lambda=1.7$ and let $R\geq 2R_1$. Since $N\geq \lambda kR$, some simple calculations yield

$$e^{\frac{kR}{2\sqrt{\lambda^2-1}}}\left(\frac{R_1}{R}\right)^{N/3}\leq e^{\frac{N}{2\lambda\sqrt{\lambda^2-1}}}2^{-N/3}\leq (e^{0.215}2^{-1/3})^N<1.$$

Substituting the estimate into the last inequality of (3.33) yields

$$|\langle(\mathcal{T}-\mathcal{T}_N)u,v\rangle_{\Gamma_R}|\lesssim\left(\frac{R_1}{R}\right)^{2N/3}\|u\|_{V_R}\|v\|_{V_R},$$

which together with (3.32) leads to

$$\begin{aligned}\|u-u^N\|_{V_R}&\lesssim k\left(\frac{R_1}{R}\right)^{2N/3}\|u\|_{V_R}\\&\lesssim k\left(\frac{R_1}{R}\right)^{2N/3}\|u-u^N\|_{V_R}+k\left(\frac{R_1}{R}\right)^{2N/3}\|u^N\|_{V_R}.\end{aligned}$$

Noting $N\geq\lambda kR$ and then $k\big(\frac{R_1}{R}\big)^{2N/3}\ll\frac{1}{2}$ for large k, and using the stability estimate $\|u^N\|_{V_R}\lesssim\|f\|_{L^2(D_{R_1})}$ (cf. Lemma 3.4), we conclude the proof of this theorem. □

3.6 The H^2-stability

Next we prove the H^2-stability of the solutions to problems (3.5) and (3.11). It will be used in proving error estimates for the CIP-FEM in the next section.

First we consider an elliptic problem with DtN boundary condition:

$$\Delta z=\eta \quad \text{in } D_R, \tag{3.34a}$$

$$z=0 \quad \text{on } \Gamma, \tag{3.34b}$$

$$\frac{\partial z}{\partial \boldsymbol{n}}=\mathcal{T}_N z \quad \text{on } \Gamma_R, \tag{3.34c}$$

where $\eta\in L^2(D_R)$. Its weak formulation is to seek for $z\in V_R$ such that

$$(\nabla z,\nabla v)_{D_R}-\langle\mathcal{T}_N z,v\rangle_{\Gamma_R}=(\eta,v)_{D_R}\quad\forall v\in V_R. \tag{3.35}$$

From (3.8), $\mathrm{Re}\langle\mathcal{T}_N v,v\rangle_{\Gamma_R}\leq 0$. The left-hand side of (3.35) provides a coercive sesquilinear form on V_R. The Lax-Milgram lemma implies that problem (3.34) has a unique solution. The following lemma states the H^2-stability for z.

Lemma 3.7. *The solution to* (3.34) *satisfies*

$$|z|_{H^2(D_R)}\leq\|\eta\|_{L^2(D_R)}. \tag{3.36}$$

Proof. For any $v\in C_0^\infty(\Omega_e)$, the formula of integral by parts shows

$$\int_{\Gamma_R} v_x(\overline{v}_{yy}n_x-\overline{v}_{yx}n_y)=(v_{xx},v_{yy})_{D_R}-\|v_{xy}\|^2_{L^2(D_R)}=\int_{\Gamma_R}\overline{v}_y(v_{xx}n_y-v_{xy}n_x),$$

where $\boldsymbol{n}=(n_x,n_y)$ and $\boldsymbol{t}=(-n_y,n_x)$ denote the unit outer normal and unit tangential vectors on Γ_R, respectively. Note that

$$v_x=\frac{\partial v}{\partial \boldsymbol{n}}n_x-\frac{\partial v}{\partial \boldsymbol{t}}n_y,\quad v_y=\frac{\partial v}{\partial \boldsymbol{n}}n_y+\frac{\partial v}{\partial \boldsymbol{t}}n_x\quad\text{on }\Gamma_R.$$

Simple calculations show

$$v_x(\overline{v}_{yy}n_x-\overline{v}_{yx}n_y)+v_y(\overline{v}_{xx}n_y-\overline{v}_{xy}n_x)=\frac{\partial v}{\partial \boldsymbol{n}}\frac{\partial^2\overline{v}}{\partial \boldsymbol{t}^2}-\frac{\partial v}{\partial \boldsymbol{t}}\frac{\partial^2\overline{v}}{\partial \boldsymbol{n}\partial \boldsymbol{t}}+\frac{1}{R}|\nabla v|^2\quad\text{on }\Gamma_R.$$

This leads to

$$2\mathrm{Re}(v_{xx},v_{yy})_{D_R}=2\|v_{xy}\|^2_{L^2(D_R)}+\int_{\Gamma_R}\left(\frac{\partial v}{\partial \boldsymbol{n}}\frac{\partial^2\overline{v}}{\partial \boldsymbol{t}^2}-\frac{\partial v}{\partial \boldsymbol{t}}\frac{\partial^2\overline{v}}{\partial \boldsymbol{n}\partial \boldsymbol{t}}\right)+\frac{1}{R}\|\nabla v\|^2_{\boldsymbol{L}^2(\Gamma_R)}.$$

Moreover, we have

$$\begin{aligned}\|\Delta v\|^2_{L^2(D_R)}&=\|v_{xx}\|^2_{L^2(D_R)}+\|v_{yy}\|^2_{L^2(D_R)}+2\mathrm{Re}(v_{xx},v_{yy})_{D_R}\\&=|v|^2_{H^2(D_R)}+\frac{1}{R}\|\nabla v\|^2_{\boldsymbol{L}^2(\Gamma_R)}+\left\langle\frac{\partial v}{\partial \boldsymbol{n}},\frac{\partial^2 v}{\partial \boldsymbol{t}^2}\right\rangle_{\Gamma_R}-\left\langle\frac{\partial v}{\partial \boldsymbol{t}},\frac{\partial^2 v}{\partial \boldsymbol{n}\partial \boldsymbol{t}}\right\rangle_{\Gamma_R}.\end{aligned}\tag{3.37}$$

Now we prove the H^2-stability of z. Note that $\mathcal{T}_N z\in C^\infty(\Gamma_R)$ and Γ, Γ_R are C^1-smooth curves. The regularity results for elliptic problems show $z\in H^2(D_R)$. Since the space $\{v|_{D_R}: v\in C_0^\infty(\Omega_e)\}$ is dense in $H^2(D_R)\cap V_R$, we can choose a sequence of functions $v_n\in\{v|_{D_R}: v\in C_0^\infty(\Omega_e)\}$ such that

$$\lim_{n\to+\infty}\|v_n-z\|_{H^2(D_R)}=0.$$

Since each v_n satisfies Eq. (3.37), we can take limits on both sides of the equation by passing $n\to+\infty$ and get

$$\|\Delta z\|^2_{L^2(D_R)}=|z|^2_{H^2(D_R)}+\frac{1}{R}\|\nabla z\|^2_{\boldsymbol{L}^2(\Gamma_R)}+\left\langle\frac{\partial z}{\partial \boldsymbol{n}},\frac{\partial^2 z}{\partial \boldsymbol{t}^2}\right\rangle_{\Gamma_R}-\left\langle\frac{\partial z}{\partial \boldsymbol{t}},\frac{\partial^2 z}{\partial \boldsymbol{n}\partial \boldsymbol{t}}\right\rangle_{\Gamma_R},\tag{3.38}$$

where the third and fourth terms on the right-hand of (3.38) are understood as the duality pairing between $H^{1/2}(\Gamma_R)$ and $H^{-1/2}(\Gamma_R)$. Using the DtN boundary condition (3.34c), the above equality is equivalent to

$$\|\Delta z\|^2_{L^2(D_R)}=|z|^2_{H^2(D_R)}+\frac{1}{R}\|\nabla z\|^2_{\boldsymbol{L}^2(\Gamma_R)}+\left\langle\mathcal{T}_N z,\frac{\partial^2 z}{\partial \boldsymbol{t}^2}\right\rangle_{\Gamma_R}-\left\langle\frac{\partial z}{\partial \boldsymbol{t}},\frac{\partial}{\partial \boldsymbol{t}}(\mathcal{T}_N z)\right\rangle_{\Gamma_R}.$$

From (3.1) and the fact that $\frac{\partial}{\partial t}=\frac{1}{r}\frac{\partial}{\partial\theta}$, we find that

$$\left\langle \mathcal{T}_N z,\frac{\partial^2 z}{\partial t^2}\right\rangle_{\Gamma_R}-\left\langle \frac{\partial z}{\partial t},\frac{\partial}{\partial t}(\mathcal{T}_N z)\right\rangle_{\Gamma_R}=-\frac{4\pi}{R^2}\sum_{|m|\le N}m^2|z_m|^2\mathrm{Re}h_m(kR).$$

From [26, eq. (3.24a)], we know that $\mathrm{Re}h_m(kR)\le 0$. This means that the left-hand side of the above equation is positive. Therefore, we have $|z|_{H^2(D_R)}\le\|\Delta z\|_{L^2(D_R)}$. The proof is finished by using (3.34a). □

Theorem 3.3. *Let $w\in V_R$ be the solution to* (3.11)*, then $w\in H^2(D_R)$. Moreover, if $N\ge kR$, the following estimate holds*

$$|w|_{H^2(D_R)}\lesssim k\|\xi\|_{L^2(D_R)}. \tag{3.39}$$

Proof. From Lemma 3.7, we have

$$|w|_{H^2(D_R)}\lesssim\|\xi\|_{L^2(D_R)}+k^2\|w\|_{L^2(D_R)}\lesssim\|\xi\|_{L^2(D_R)}+k\|w\|_{V_R}.$$

Then inequality (3.39) follows directly from Lemma 3.4. □

Corollary 3.1. *Let $u^N\in V_R$ be the solution to* (3.5)*, then $u^N\in H^2(D_R)$. Moreover, if $N\ge kR$, the following estimate holds*

$$\|u^N\|_{V_R}+k^{-1}|u^N|_{H^2(D_R)}\lesssim\|f\|_{L^2(D_R)}.$$

Proof. The corollary is a direct consequence of Lemma 3.4 and Theorem 3.3. □

4 CIP-FEM and its preasymptotic error analysis

In this section, we propose a CIP-FEM for the truncated DtN problem (3.3) and give a preasymptotic error analysis for the discrete solution.

4.1 Finite element meshes

Note that both Γ and Γ_R are curved boundaries. We can not subdivide D_R into the union of triangles which only have straight edges. To solve the issue, we follow Melenk and Sauter [26] to partition D_R into meshes of curved triangles.

Let $\mathcal{M}_H$ be a coarse triangulation of D_R such that $\overline{D}_R=\cup_{K\in\mathcal{M}_H}K$. For each element $K\in\mathcal{M}_H$, let h_K denote the diameter of the smallest ball which contains K. The mesh size of $\mathcal{M}_H$ is denoted by $H:=\max_{K\in\mathcal{M}_H}h_K$. Now we make some assumptions on the macro partition $\mathcal{M}_H$.

(A1) Each $K\in\mathcal{M}_H$ has at most one curved edge on ∂D_R, or equivalently, at most two vertices on ∂D_R.

(A2) If K has at most one vertex on ∂D_R, then all edges of K are straight.

(A3) For any $K\in\mathcal{M}_H$, there are two closed balls $B_K^{(1)}$, $B_K^{(2)}$ such that

$$B_K^{(1)}\subset K\subset B_K^{(2)},\quad \text{diameter}(B_K^{(1)})\geq\sigma_1 H,\quad \text{diameter}(B_K^{(2)})\leq\sigma_2 H,$$

where σ_1,σ_2 are two positive constants independent of K and H.

In fact, assumption (A3) implies that elements in $\mathcal{M}_H$ are quasi-uniform and shape-regular.

For each $K\in\mathcal{M}_H$, let T_K denote the closed triangle whose three edges are straight and whose vertices are the same as the vertices of K. Clearly $T_K=K$ if K has at most one vertex on ∂D_R. By assumption (A3), $\widehat{\mathcal{M}}_H:=\{T_K:K\in\mathcal{M}_H\}$ provides a quasi-uniform, shape-regular, and body-fitted mesh of D_R. We define a polygonal domain by

$$\hat{D}_R:=\text{interior}\left(\cup_{T_K\in\widehat{\mathcal{M}}_H}T_K\right).$$

Since Γ is piecewise C^2-smooth, it is reasonable to assume that

(A4) There is a bi-Lipschitz and one-to-one mapping $\boldsymbol{\chi}$: $\overline{D_R}\to\overline{\hat{D}_R}$ which is C^2-smooth on each $K\in\mathcal{M}_H$ and satisfies $T_K=\boldsymbol{\chi}(K)$.

Let $\{\widehat{\mathcal{M}}_h:h<H\}$ be a family of triangular meshes of $\hat{D}_R$, which are obtained successively by uniform refinements of $\widehat{\mathcal{M}}_H$. In other words, each triangle is divided into four triangles of the same shape. The fine meshes of D_R are defined by

$$\mathcal{M}_h:=\left\{K=\boldsymbol{\chi}^{-1}(\hat{K}):\forall\hat{K}\in\widehat{\mathcal{M}}_h\right\},\quad h:=\max_{K\in\mathcal{M}_h}h_K.$$

For any $K\in\mathcal{M}_h$, $\boldsymbol{\chi}_K:=\boldsymbol{\chi}|_K$ may depend on the macro mesh-size H, but is independent of h. Since $\boldsymbol{\chi}_K$ is one-to-one and C^2-smooth, it is reasonable to assume that

(A5) There is a constant C which depends on H but is independent of h such that

$$\|\boldsymbol{\chi}_K\|_{W^{2,\infty}(K)}\leq C,\quad \left\|\boldsymbol{\chi}_K^{-1}\right\|_{W^{2,\infty}(\hat{K})}\leq C\quad\forall K\in\mathcal{M}_h,\ \hat{K}=\boldsymbol{\chi}(K).$$

Since $\widehat{\mathcal{M}}_h$ is quasi-uniform and shape-regular, it is easy to show that $\mathcal{M}_h$ is also quasi-uniform and shape-regular. The details are omitted.

4.2 Finite element spaces

Now we define the FE space associated with $\mathcal{M}_h$. First let the FE space on the polygonal domain $\hat{D}_R$ be defined by

$$\hat{V}_h=\left\{v_h\in H^1(\hat{D}_R):\ v_h|_{\hat{\Gamma}}=0,\ v_h|_{\hat{K}}\in P_1(\hat{K}),\ \forall\hat{K}\in\widehat{\mathcal{M}}_h\right\},$$

where $\hat{\Gamma}:=\boldsymbol{\chi}(\Gamma)$ and $P_1(\hat{K})$ denotes the space of all linear polynomials on $\hat{K}$. The FE space on D_R is defined by

$$V_h=\left\{v_h\circ\boldsymbol{\chi}\colon\ \forall v_h\in\hat{V}_h\right\}. \tag{4.1}$$

Since $\boldsymbol{\chi}$: $D_R\to\hat{D}_R$ is continuous and one-to-one, it is easy to see that $V_h\subset V_R$.

4.3 Finite element approximation of (3.3)

Now we formulate the CIP-FE approximation to problem (3.3). For convenience, we endow each element $K\in\mathcal{M}_h$ a unique index $i_K\in\mathbb{N}$. Let $\mathcal{E}_h^I$ be the set of all interior edges of $\mathcal{M}_h$. The jump of a function v across an edge $e=\partial K\cap\partial K'\in\mathcal{E}_h^I$ is defined as follows

$$[v]_e:=\begin{cases}v|_K-v|_{K'}, & \text{if } i_K>i_{K'},\\ v|_{K'}-v|_K, & \text{if } i_{K'}>i_K.\end{cases} \tag{4.2}$$

We define the spaces of piecewise regular functions by

$$H^2(\mathcal{M}_h):=\{v\in L^2(D_R)\colon\ v|_K\in H^2(K),\ \forall K\in\mathcal{M}_h\},\quad W:=V_R\cap H^2(\mathcal{M}_h),$$

which are equipped with the semi-norm

$$|v|_{H^2(\mathcal{M}_h)}:=\left(\sum_{K\in\mathcal{M}_h}|v|^2_{H^2(K)}\right)^{1/2}.$$

Remember that $\boldsymbol{\chi}$ is C^2-smooth on each element $K\in\mathcal{M}_h$. Therefore, we have

$$V_h\subset W.$$

For any $w,v\in W$, we define the sesquilinear form of interior Neumann penalties by

$$J(w,v):=\sum_{e\in\mathcal{E}_h^I}\gamma_e h_e\langle[\nabla w\cdot\boldsymbol{n}],[\nabla v\cdot\boldsymbol{n}]\rangle_e, \tag{4.3}$$

where h_e is the length of edge e and γ_e is a parameter with nonpositive imaginary part. Let b_h: $W\times W\to\mathbb{C}$ be a sesquilinear form defined by

$$b_h(w,v):=b_N(w,v)+J(w,v)\quad\forall w,v\in W.$$

The CIP-FEM for problem (3.3) is defined as follows: find $u_h^N\in V_h$ such that

$$b_h(u_h^N,v_h)=-(f,v_h)_{D_R}\quad\forall v_h\in V_h. \tag{4.4}$$

Note that $J(w,v)=0$ for all $w\in H^2(D_R)$ and $v\in W$, the solution of (3.3) also satisfies

$$b_h(u^N,v)=-(f,v)_{D_R}\quad\forall v\in W. \tag{4.5}$$

This yields the Galerkin orthogonality

$$b_h(u^N-u_h^N,v_h)=0\quad\forall v_h\in V_h. \tag{4.6}$$

4.4 Elliptic projection operator

The elliptic projection operator will play an important role in the preasymptotic error analysis [11, 23, 28, 30]. For simplicity, we assume $\gamma_e \equiv \gamma$ and define

$$a_h(w,v) = (\nabla w, \nabla v)_{D_R} + k^2 (w,v)_{D_R} - \langle \mathcal{T}_N w, v\rangle_{\Gamma_R} + J(w,v).$$

The elliptic projection $\mathcal{P}_h$: $V_R \cap H^2(D_R) \to V_h$ is defined by: for any $w \in V_R \cap H^2(D_R)$, find $\mathcal{P}_h w \in V_h$ such that

$$a_h(v_h, \mathcal{P}_h w) = a_h(v_h, w) \quad \forall v_h \in V_h. \tag{4.7}$$

Since $\mathrm{Re}\langle \mathcal{T}_N v, v\rangle_{\Gamma_R} \le 0$ and $|\langle \mathcal{T}_N w, v\rangle_{\Gamma_R}| \lesssim \|w\|_{V_R} \|v\|_{V_R}$, by arguments similar to the proof of [23, Lemma 4.2], we get the continuity and coercivity of a_h. The details are omitted.

Lemma 4.1. *There exists a constant $\gamma_0 > 0$ such that, if $\mathrm{Re}\gamma \ge -\gamma_0$ and $|\gamma| \lesssim 1$, then for any $w, v \in W$ and $v_h \in V_h$,*

$$|a_h(w,v)| \lesssim (\|w\|_{V_R} + h|w|_{H^2(\mathcal{M}_h)})(\|v\|_{V_R} + h|v|_{H^2(\mathcal{M}_h)}), \tag{4.8}$$

$$\mathrm{Re} a_h(v_h, v_h) \gtrsim \|v_h\|^2_{V_R}. \tag{4.9}$$

In fact, Lemma 4.1 shows that problem (4.7) has a unique solution and $\mathcal{P}_h$ is well-defined. Now we are ready to prove the error estimates for the elliptic projection.

Lemma 4.2. *Let the condition of Lemma 4.1 be satisfied and suppose $kh \lesssim 1$. Then*

$$\|w - \mathcal{P}_h w\|_{L^2(D_R)} + h\|w - \mathcal{P}_h w\|_{V_R} \lesssim h^2 \|w\|_{H^2(D_R)} \quad \forall w \in V_R \cap H^2(D_R). \tag{4.10}$$

Proof. Let $\hat{\pi}_h$: $C_B(\hat{D}_R) \to \hat{V}_h$ be the Lagrangian FE interpolation operator (cf. [6, 9]). It is well-known that

$$\|\hat{w} - \hat{\pi}_h \hat{w}\|_{L^2(\hat{K})} + h|\hat{w} - \hat{\pi}_h \hat{w}|_{H^1(\hat{K})} \lesssim h^2 |\hat{w}|_{H^2(\hat{K})} \quad \forall \hat{w} \in H^2(\hat{K}),\ \hat{K} \in \widehat{\mathcal{M}}_h.$$

For any $w \in H^2(D_R) \hookrightarrow C_B(D_R)$, since $\boldsymbol{\chi}^{-1}$ is continuous, we have $\hat{w} := w \circ \boldsymbol{\chi}^{-1} \in C_B(\hat{D}_R)$. The FE interpolation $\pi_h w \in V_h$ of w is defined by

$$\pi_h w := (\hat{\pi}_h \hat{w}) \circ \boldsymbol{\chi} \quad \text{on } D_R.$$

Using assumption (A5), we deduce that

$$\|w - \pi_h w\|_{L^2(K)} \lesssim \|\hat{w} - \hat{\pi}_h \hat{w}\|_{L^2(\hat{K})} \lesssim h^2 |\hat{w}|_{H^2(\hat{K})} \lesssim h^2 \|w\|_{H^2(K)},$$
$$|w - \pi_h w|_{H^1(K)} \lesssim \|\hat{w} - \hat{\pi}_h \hat{w}\|_{H^1(\hat{K})} \lesssim h|\hat{w}|_{H^2(\hat{K})} \lesssim h\|w\|_{H^2(K)}.$$

Together with the assumption $kh \lesssim 1$, the above estimates yield

$$\|w - \pi_h w\|_{L^2(D_R)} + h\|w - \pi_h w\|_{V_R} \lesssim h^2 \|w\|_{H^2(D_R)} \quad \forall w \in H^2(D_R). \tag{4.11}$$

Now we estimate $\|\pi_h w-\mathcal{P}_h w\|_{V_R}$. For any $v_h\in V_h$, $\hat{v}_h:=v_h\circ\boldsymbol{\chi}^{-1}\in\hat{V}_h$ is a piecewise linear function. Using assumption (A5), we find that

$$\begin{aligned}&|v_h|_{H^2(K)}\lesssim|\hat{v}_h|_{H^1(\hat{K})}+|\hat{v}_h|_{H^2(\hat{K})}=|\hat{v}_h|_{H^1(\hat{K})}\lesssim|v_h|_{H^1(K)},\\&|v_h-w|_{H^2(K)}\lesssim|\hat{v}_h-\hat{w}|_{H^1(\hat{K})}+|\hat{v}_h-\hat{w}|_{H^2(\hat{K})}\lesssim|v_h-w|_{H^1(K)}+\|w\|_{H^2(K)},\end{aligned}$$

for any $K\in\mathcal{M}_h$ and $\hat{K}=\boldsymbol{\chi}(K)$. The coercivity and continuity of a_h in Lemma 4.1 show

$$\begin{aligned}&\|\pi_h w-\mathcal{P}_h w\|_{V_R}^2\\\lesssim&\mathrm{Re}a_h(\pi_h w-\mathcal{P}_h w,\pi_h w-\mathcal{P}_h w)\\=&\mathrm{Re}a_h(\pi_h w-\mathcal{P}_h w,\pi_h w-w)\\\lesssim&\left(\|\pi_h w-\mathcal{P}_h w\|_{V_R}+h|\pi_h w-\mathcal{P}_h w|_{H^2(\mathcal{M}_h)}\right)\left(\|\pi_h w-w\|_{V_R}+h|\pi_h w-w|_{H^2(\mathcal{M}_h)}\right)\\\lesssim&\|\pi_h w-\mathcal{P}_h w\|_{V_R}\left(\|\pi_h w-w\|_{V_R}+h\|w\|_{H^2(D_R)}\right).\end{aligned}$$

Together with (4.11), this shows

$$\|w-\mathcal{P}_h w\|_{V_R}\lesssim h\|w\|_{H^2(D_R)}. \tag{4.12}$$

To estimate $\|w-\mathcal{P}_h w\|_{L^2(D_R)}$, we consider the elliptic dual problem

$$-\Delta z+k^2 z=w-\mathcal{P}_h w \quad \text{in } D_R, \tag{4.13a}$$

$$z=0 \quad \text{on } \Gamma, \tag{4.13b}$$

$$\frac{\partial z}{\partial n}=\mathcal{T}_N z \quad \text{on } \Gamma_R. \tag{4.13c}$$

From Lemma 3.7, we have

$$|z|_{H^2(D_R)}\lesssim\|w-\mathcal{P}_h w\|_{L^2(D_R)}+k^2\|z\|_{L^2(D_R)}. \tag{4.14}$$

Multiplying both sides of (4.13a) by $\overline{z}$ and integrating by parts in D_R, we obtain

$$\|z\|_{V_R}^2-\langle\mathcal{T}_N z,z\rangle_{\Gamma_R}=(w-\mathcal{P}_h w,z)_{D_R}.$$

Taking real part of the equality and using Young's inequality, we have

$$\|z\|_{V_R}\lesssim\frac{1}{k}\|w-\mathcal{P}_h w\|_{L^2(D_R)}.$$

Combining the estimate with (4.14) yields

$$\|z\|_{H^2(D_R)}\lesssim\|w-\mathcal{P}_h w\|_{L^2(D_R)}. \tag{4.15}$$

Next multiplying both sides of (4.13a) by the conjugate of $w-\mathcal{P}_h w$ and integrating by parts in D_R, we get

$$\begin{aligned}\|w-\mathcal{P}_h w\|^2_{L^2(D_R)} &= a_h(z,w-\mathcal{P}_h w) = a_h(z-\pi_h z,w-\mathcal{P}_h w)\\ &\lesssim (\|z-\pi_h z\|_{V_R}+h|z-\pi_h z|_{H^2(\mathcal{M}_h)})(\|w-\mathcal{P}_h w\|_{V_R}+h|w-\mathcal{P}_h w|_{H^2(\mathcal{M}_h)})\\ &\lesssim h^2\|z\|_{H^2(D_R)}\|w\|_{H^2(D_R)}.\end{aligned}$$

Using (4.15), we get $\|w-\mathcal{P}_h w\|_{L^2(D_R)}\lesssim h^2\|w\|_{H^2(D_R)}$. The proof is finished. □

4.5 Preasymptotic error analysis

The purpose of this section is to estimate the error between the scattering solution u and the FE solution u_h^N. First we prove the preasymptotic error estimate between the truncated continuous solution u^N and the discrete solution u_h^N.

Lemma 4.3. *Let the assumptions of Lemma* 4.1 *be satisfied and assume $N\geq kR$. Let u^N, u_h^N be the solutions to the truncated problem* (3.3) *and the discrete problem* (4.4)*, respectively. There exists a positive C_0 independent of k, h, and γ such that, if $k^3h^2\leq C_0$, then*

$$\|u^N-u_h^N\|_{V_R}\lesssim (kh+k^3h^2)\|f\|_{L^2(D_R)}, \tag{4.16}$$

$$\|u^N-u_h^N\|_{L^2(D_R)}\lesssim k^2h^2\|f\|_{L^2(D_R)}. \tag{4.17}$$

Proof. Write $e_h=u^N-u_h^N\in V_R$, $\eta_h=\overline{u_h^N}-\mathcal{P}_h\overline{u^N}\in V_h$ and $\zeta_h=\overline{u^N}-\mathcal{P}_h\overline{u^N}\in V_R$ for convenience. First we consider the dual problem: find $w\in V_R$ such that

$$b_N(v,w)=(v,e_h)_{D_R}\quad \forall v\in V_R.$$

Since b_N satisfies the inf-sup condition in Theorem 3.1, the problem has a unique solution. By arguments similar to (3.23)-(3.24), the conjugate $\overline{w}$ is the solution to the problem

$$b_N(\overline{w},v)=(\overline{e_h},v)_{D_R}\quad \forall v\in V_R.$$

Using Lemma 3.4 and Theorem 3.3, we get the stability estimates

$$\|w\|_{V_R}+k^{-1}|w|_{H^2(D_R)}\lesssim\|e_h\|_{L^2(D_R)}. \tag{4.18}$$

Similarly, from Corollary 3.1, we get the stability of u^N

$$\|u^N\|_{V_R}+k^{-1}|u^N|_{H^2(D_R)}\lesssim\|f\|_{L^2(D_R)}. \tag{4.19}$$

Using the Galerkin orthogonality (4.6), error estimates (4.10)-(4.11), and inequalities (4.18)-(4.19), we find that

$$\begin{aligned}\|e_h\|^2_{L^2(D_R)} &= b_N(e_h,w) = b_h(e_h,w) = b_h(e_h,w-\mathcal{P}_h w)\\ &= a_h(e^N,w-\mathcal{P}_h w) - 2k^2(e_h,w-\mathcal{P}_h w)\\ &= a_h(u^N-\pi_h u^N,w-\mathcal{P}_h w) - 2k^2(e_h,w-\mathcal{P}_h w)\\ &\lesssim h^2\|u^N\|_{H^2(D_R)}\|w\|_{H^2(D_R)} + k^2h^2\|e_h\|_{L^2(D_R)}\|w\|_{H^2(D_R)}\\ &\lesssim k^2h^2\|f\|_{L^2(D_R)}\|e_h\|_{L^2(D_R)} + k^3h^2\|e_h\|^2_{L^2(D_R)}.\end{aligned}$$

Therefore, there exists a positive constant C_0 such that, when $k^3h^2 \le C_0$,

$$\|e_h\|_{L^2(D_R)} \lesssim k^2h^2\|f\|_{L^2(D_R)}. \tag{4.20}$$

Next, we infer from (4.10), (4.19), and (4.20) that

$$\begin{aligned}&\|\zeta_h\|_{L^2(D_R)} \lesssim h^2\|u^N\|_{H^2(D_R)} \lesssim kh^2\|f\|_{L^2(D_R)},\\ &\|\eta_h\|_{L^2(D_R)} \lesssim \|\overline{e_h}\|_{L^2(D_R)} + \|\zeta_h\|_{L^2(D_R)} \lesssim k^2h^2\|f\|_{L^2(D_R)}.\end{aligned}$$

It is easy to see that

$$a_h\left(\overline{\mathcal{P}_h\overline{u^N}},v_h\right) = a_h\left(\overline{v}_h,\mathcal{P}_h\overline{u^N}\right) = a_h\left(\overline{v}_h,\overline{u^N}\right) = a_h(u^N,v_h) \quad \forall v_h \in V_h. \tag{4.21}$$

From (4.9), (4.21) and (4.6), we find that

$$\begin{aligned}\|\overline{\eta_h}\|^2_{V_R} &\lesssim \mathrm{Re}\, a_h(\overline{\eta_h},\overline{\eta_h}) = \mathrm{Re}\, a_h(u_h^N-u^N,\overline{\eta_h})\\ &= 2k^2\,\mathrm{Re}\left(u_h^N-u^N,\overline{\eta_h}\right)_{D_R} \lesssim k^6h^4\|f\|^2_{L^2(D_R)}.\end{aligned}$$

Therefore, using the triangle inequality, we have

$$\|e_h\|_{V_R} = \|\overline{e_h}\|_{V_R} \le \left\|\overline{u^N}-\mathcal{P}_h\overline{u^N}\right\|_{V_R} + \|\eta_h\|_{V_R} \lesssim (kh+k^3h^2)\|f\|_{L^2(D_R)}.$$

The proof is finished. □

Finally, we present the main result of the paper, which is a direct consequence of Theorem 3.2 and Lemma 4.3.

Theorem 4.1. *Let u be the solution to the scattering problem* (2.1) *and let u_h^N be the solution to the finite problem* (4.4)*. Assume that*

- $N \ge 1.7kR$, $R \ge 2R_1$ *and*
- *the assumptions of Lemma* 4.1 *are satisfied.*

There exists a positive C_0 independent of k, h, and the penalty parameter γ such that, when $k^3h^2 \le C_0$,

$$\|u-u_h^N\|_{V_R} \lesssim \left[k(R_1/R)^{2N/3}+kh+k^3h^2\right]\|f\|_{L^2(D_R)}.$$

Remark 4.1. Theorem 4.1 also holds for the linear FEM, that is, for $\gamma=0$. Moreover, the theories in this paper can be easily extended to higher-order (CIP-)FEMs. Here we do not elaborate on the details and refer to [11, 30] for relevant studies. Some specific remarks on the theories are given below.

- The FE preasymptotic error estimates contain two terms. The first term in (4.16) is the order of $\mathcal{O}(kh)$ and comes from FE interpolation error. The second term is the order of $\mathcal{O}(k^3h^2)$ and comes from high-frequency pollution error (see [1]).
- Although the theoretical analysis in our work only indicate that the error bound of CIP-FEM is not larger than that of the traditional FEM, the CIP-FEM is more efficient than FEM in solving high-frequency Helmholtz equation in practice. In fact, from the numerical experiments in the next section, we can see that the penalty parameter γ (namely, γ_e in (4.3)) may be tuned to reduce the pollution error greatly.

5 Numerical experiments

In this section, we solve the scattering problem (2.1) numerically by the CIP-FEM and the truncated DtN boundary condition defined on a circle of radius $R=1$. The obstacle is defined by

$$\Omega := \{\boldsymbol{x}=(x,y)\in\mathbb{R}^2 :\ x^2+y^2\le R_0^2\},\quad R_0=0.5.$$

The penalty parameters for the CIP-FEM are obtained by a dispersion analysis for Helmholtz equation on equilateral triangulations and have the expression

$$\gamma_e = -\frac{\sqrt{3}}{24}-\frac{\sqrt{3}}{1728}(kh_e)^2 \quad \forall e\in\mathcal{E}_h^I.$$

Fig. 2 shows a quasi-uniform and body-fitted mesh of D_R. All triangles are approximately equilateral and their union forms a polygonal domain $\hat{D}_R$. For simplicity, we use $\hat{D}_R$ and the space $\hat{V}_h$ of piecewise linear functions (see Subsection 4.2) to replace D_R and V_h in practical computations.

From Theorem 4.1, the FE error estimate is given by

$$\|u-u_h^N\|_{V_R} \le C_1kh+C_2k^3h^2+C_3k(R_1/R)^{2N/3},$$

where C_j, $j=1,2,3$ are dependent of f. We remark that the third term is negligible for sufficiently large N. Moreover, from Lemma 4.3, the FE error estimate also requires $N\ge$

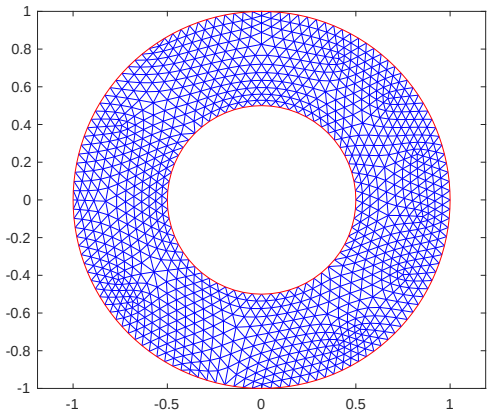

Figure 2: A sample mesh for the truncated DtN method.

kR. Therefore, we shall take $N=kR$ for the truncation of the DtN operator, except for specific declaration. Recall that the truncated DtN operator is defined by

$$\mathcal{T}_N w = R^{-1} \sum_{|m|\le N} h_m(kR) w_m e^{im\theta}, \quad h_m(r) = r\frac{H_m^{(1)'}(r)}{H_m^{(1)}(r)}.$$

It's well-known that computing $H_m^{(1)}(r)$ is difficult for large m. Fortunately, we have the recurrence relations (see [27, §10.6])

$$\frac{H_m^{(1)'}(r)}{H_m^{(1)}(r)} = \frac{H_{m-1}^{(1)}(r)}{H_m^{(1)}(r)} - \frac{m}{r}, \quad \frac{H_{m-1}^{(1)'}(r)}{H_{m-1}^{(1)}(r)} = -\frac{H_m^{(1)}(r)}{H_{m-1}^{(1)}(r)} + \frac{m-1}{r},$$

which provide an efficient way for computing $h_m(r)$

$$h_m(r) = r\frac{H_m^{(1)'}(r)}{H_m^{(1)}(r)} = \frac{r^2}{m-1-h_{m-1}(r)} - m \quad \forall m \in \mathbb{Z}.$$

Example 5.1. This example is to investigate the effect of pollution errors due to high-frequency. The exact solution is chosen as

$$u(\boldsymbol{x}) = H_0^{(1)}(k|\boldsymbol{x}|) \quad \text{in } \Omega_e.$$

Fig. 3 plots the relative H^1-errors of the (CIP-)FE solutions and FE interpolations for k=8, 32 and 128, respectively. It is shown that, for k=8, the error curves of FE and CIP-FE solutions fit the curves of the FE interpolations very well. This means that the pollution errors have small effects on the discrete solutions. While for large k, the FE errors decay

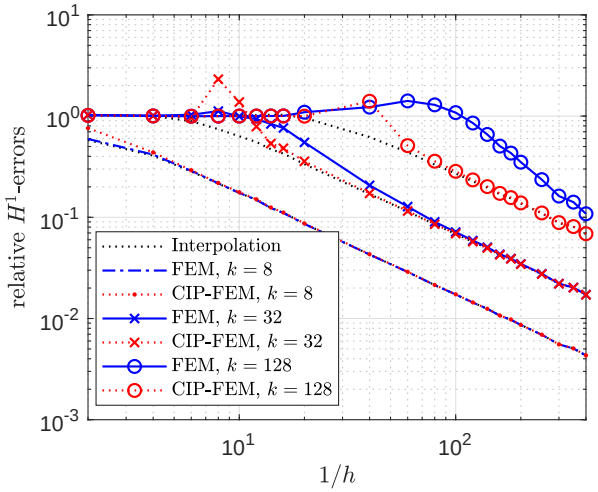

Figure 3: Relative H^1-errors of the (CIP-)FE solutions and FE interpolations for $k=8$, 32, 128, respectively.

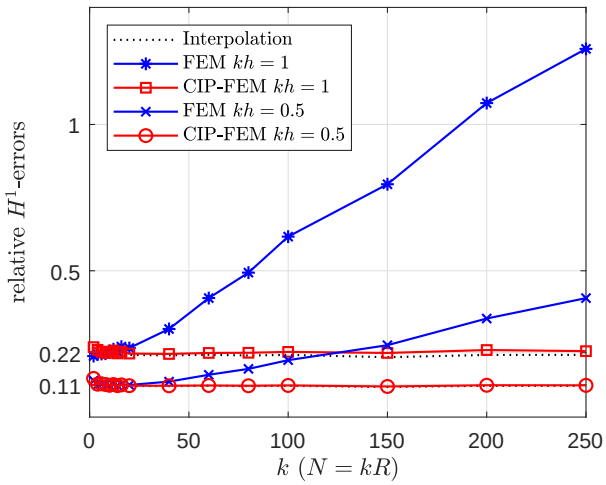

Figure 4: Relative H^1-errors of the (CIP-)FE solutions and FE interpolations for $kh=1$ and $kh=0.5$.

very slowly in a range of mesh sizes far from the decaying point of the corresponding FE interpolations. This shows clearly the effect of pollution errors of traditional FEM. The errors of CIP-FE solutions behave similarly, but begin to decay much earlier than traditional FEM, which implies that the CIP-FEM reduces the pollution effect greatly.

Next we fix $kh=1$ and $kh=0.5$, respectively, and investigate the pollution errors. Fig. 4 shows that the pollution error of FEM grows fast, while the pollution error of the CIP-FEM is almost invariant for k up to 250. This means that the CIP-FEM is much more stable and efficient than the FEM for high-frequency problem. Moreover, we find that the

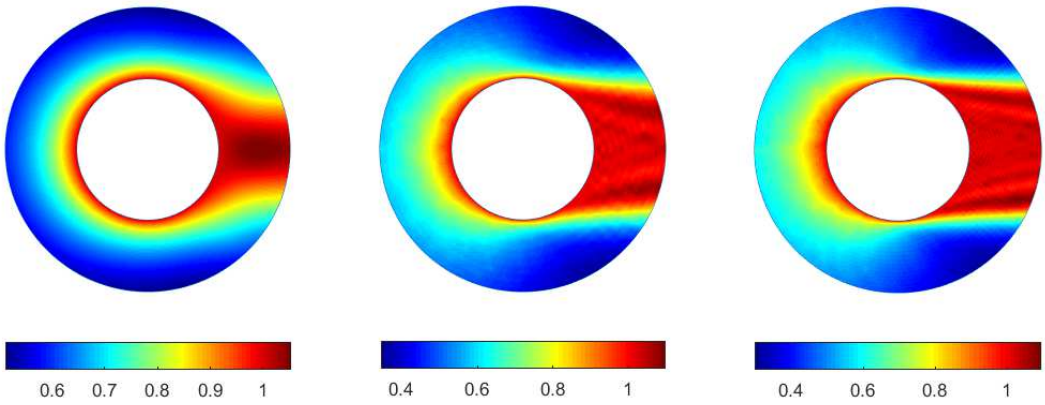

Figure 5: The amplitudes of the CIP-FE solutions for $k=8$ (left), 64 (middle), and 128 (right).

errors of the FE interpolation (or CIP-FEM) for $kh=1$ are almost twice for $kh=0.5$. This indicates the asymptotic behavior

$$\text{the error of FE interpolation (or CIP-FEM when } k\leq 250) \sim kh.$$

However, the errors of the FEM for $kh=1$ are almost four times that of $kh=0.5$, which corresponds to the asymptotic behavior

$$\text{the pollution error of FEM} \sim k(kh)^2.$$

Example 5.2. This example is to investigate the effect of the truncation order N on the approximation error. The exact solution is given by

$$u=-\sum_{m\in\mathbb{Z}} \boldsymbol{i}^m \frac{J_m(kR_0)}{H_m^{(1)}(kR_0)} H_m^{(1)}(kr)e^{\boldsymbol{i}m\theta}, \tag{5.1}$$

which satisfies Eq. (2.1a) with $f=0$ and the boundary condition $u=-e^{ikx}$ on $\Gamma=\Gamma_{R_0}$. In practice, the infinite series in (5.1) is truncated so that the relative change due to an additional term is less than 10^{-6}. The amplitudes of the CIP-FE solutions for $k=8$, 64 and 128 are shown in Fig. 5, respectively.

In the left graph of Fig. 6, we fix $k=32$ and choose $h=1/16$, $1/32$, $1/48$ and $1/64$, respectively. The log-log plots of errors show that the truncation errors of the DtN operator decay extremely fast when N reaches a specific value. For example, the H^1-error for $h=1/64$ decays slowly before $N=18$, but encounters a sudden drop at $N=20$. It implies that the truncation error is much larger than the discrete error for small N, but decays exponentially when $N\geq 20$. So there exists a threshold $N\geq \alpha k$ for the exponential convergence of the truncated DtN operator with a constant $\alpha>0$.

Now we investigate the theoretical threshold $N\geq kR$ for convergence by numerical experiments for different values of k. In the right graph of Fig. 6, we choose a small

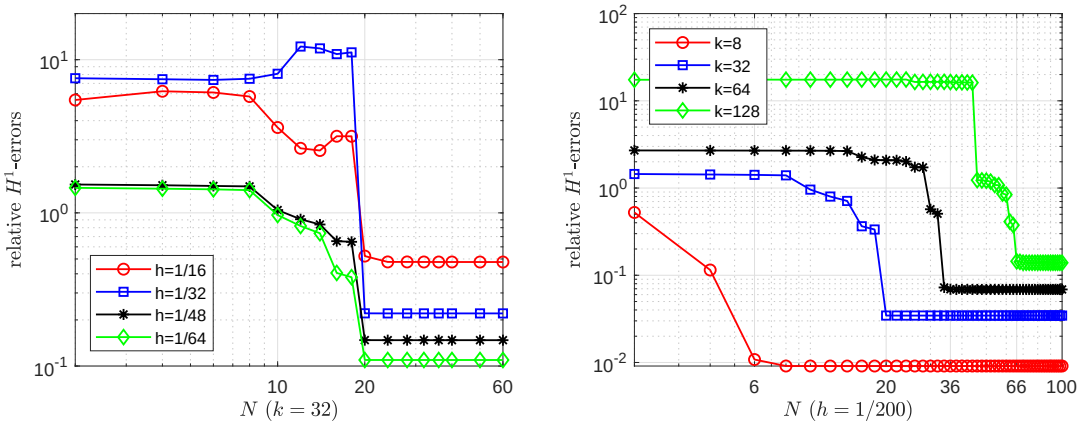

Figure 6: The relative H^1 errors of the CIP-FEM with different h (left) and k (right).

mesh-size $h=1/200$ so that the CIP-FE discrete errors are negligible. The graph shows that the optimal truncation orders of the DtN operator are $N_{\text{opt}}=6$, 20, 36 and 66, while the theoretical predictions are $kR=8$, 32, 64 and 128, respectively. Based on the above observations, we conclude that the truncation condition $N\geq kR$ is safe and necessary for the truncated DtN method.

Acknowledgments

The authors would like to thank professor Peijun Li from Purdue University, USA, and professor Haijun Wu from Nanjing University, China, for their valuable suggestions on this work.

This work was supported by the National Science Fund for Distinguished Young Scholars 11725106 and by China NSF grant 11831016.

References

[1] I. M. 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(6):2392–2423, 1997.

[2] G. Bao, P. Li, and X. Yuan. An adaptive finite element DtN method for the open cavity scattering problems. *CSIAM Trans. Appl. Math.*, to appear, 2020.

[3] G. Bao and H. Wu. Convergence analysis of the perfectly matched layer problems for time-harmonic Maxwell's equations. *SIAM J. Numer. Anal.*, 43(5):2121–2143, 2005.

[4] J. P. Berenger. A perfectly matched layer for the absorption of electromagnetic waves. *J. Comp. Phys.*, 114(2):185–200, 1994.

[5] J. Bramble and J. Pasciak. Analysis of a finite PML approximation for the three dimensional time-harmonic Maxwell and acoustic scattering problems. *Math. Comp.*, 76(258):597–614, 2007.

[6] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*, volume 15. Springer Science & Business Media, 2007.
[7] S. N. Chandler-Wilde and P. Monk. Wave-number-explicit bounds in time-harmonic scattering. *SIAM J. Math. Anal.*, 39(5):1428–1455, 2008.
[8] Z. Chen and X. Liu. An adaptive perfectly matched layer technique for time-harmonic scattering problems. *SIAM J. Numer. Anal.*, 43(2):645–671, 2005.
[9] P. G. Ciarlet. *The finite element method for elliptic problems*, volume 40. Siam, 2002.
[10] J. Douglas and T. Dupont. Interior penalty procedures for elliptic and parabolic Galerkin methods. In *Computing methods in applied sciences*, pages 207–216. Springer, 1976.
[11] Y. Du and H. Wu. Preasymptotic error analysis of higher order FEM and CIP-FEM for helmholtz equation with high wave number. *SIAM J. Numer. Anal.*, 53(2):782–804, 2015.
[12] D. Givoli and J. Keller. A finite element method for large domains. *Comput. Methods Appl. Mech. Engrg.*, 76(1):41–66, 1989.
[13] M. Grote and J. Keller. On nonreflecting boundary conditions. *J. Comput. Phys*, 122(2):231–243, 1995.
[14] I. Harari and T. Hughes. Analysis of continuous formulations underlying the computation of time-harmonic acoustics in exterior domains. *Comput. Methods Appl. Mech. Engrg.*, 97(1):103–124, 1992.
[15] G. C. Hsiao, N. Nigam, J.E. Pasciak, and L. Xu. Error analysis of the DtN-FEM for the scattering problem in acoustics via Fourier analysis. *J. Comput. Appl. Math.*, 235(17):4949–4965, 2011.
[16] F. Ihlenburg. *Finite element analysis of acoustic scattering*, volume 132. Springer Science & Business Media, 2006.
[17] F. Ihlenburg and I. Babuška. Finite element solution of the Helmholtz equation with high wave number part I. the *h*-version of the FEM. *Comput. Math. Appl.*, 30(9):9–37, 1995.
[18] F. Ihlenburg and I. Babuška. 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(1):315–358, 1997.
[19] X. Jiang, P. Li, J. Lv, and W. Zheng. An adaptive finite element method for the wave scattering with transparent boundary condition. *Journal of Scientific Computing*, 72(3):936–956, 2017.
[20] X. Jiang, P. Li, and W. Zheng. Numerical solution of acoustic scattering by an adaptive DtN finite element method. *Commun Comput. Phys.*, 13(5):1277–1244, 2013.
[21] M. Lassas and E. Somersalo. On the existence and convergence of the solution of PML equations. *Computing*, 60(3):229–241, 1998.
[22] P. Li and X. Yuan. Convergence of an adaptive finite element DtN method for the elastic wave scattering by periodic structures. *Comput. Methods Appl. Mech. Engrg.*, 360:1127221, 2020.
[23] Y. Li and H. Wu. FEM and CIP-FEM for Helmholtz equation with high wave number and perfectly matched layer truncation. *SIAM J. Numer. Anal.*, 57(1):96–126, 2019.
[24] J. Melenk. *On generalized finite-element methods*. PhD thesis, research directed by Dept. of Mathematics. University of Maryland at College Park, 1995.
[25] J. Melenk, A. Parsania, and S. Sauter. General DG-methods for highly indefinite Helmholtz problems. *Journal of Scientific Computing*, 57(3):536–581, 2013.
[26] J. Melenk and S. Sauter. Convergence analysis for finite element discretizations of the Helmholtz equation with Dirichlet-to-Neumann boundary conditions. *Math. Comp.*, 79(272):1871–1914, 2010.
[27] F. Olver, D. Lozier, R. Biosvert, and C. Clark. *NIST Handbook of Mathematical Functions*. Cambridge University Press, New York, 2010.

[28] H. Wu. Pre-asymptotic error analysis of CIP-FEM and FEM for the Helmholtz equation with high wave number. part I: Linear version. *IMA J. Numer. Anal.*, 34(3):1266–1288, 2014.

[29] L. Xu and T. Yin. Analysis of the Fourier series Dirichlet-to-Neumann boundary condition of the Helmholtz equation and its application to finite element methods. *arXiv preprint arXiv:1609.00583*, 2016.

[30] L. Zhu and H. Wu. Preasymptotic error analysis of CIP-FEM and FEM for Helmholtz equation with high wave number. part II: hp version. *SIAM J. Numer. Anal.*, 51(3):1828–1852, 2013.