# Weierstraß-Institut für Angewandte Analysis und Stochastik Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint

ISSN 2198-5855

# PML and high-accuracy boundary integral equation solver for wave scattering by a locally defected periodic surface

Xiuchen Yu<sup>1</sup>, Guanghui Hu<sup>2</sup>, Wangtao Lu<sup>3</sup>, Andreas Rathsfeld<sup>4</sup>

submitted: August 16, 2021

School of Mathematical Sciences
 Zhejiang University
 Hangzhou 310027
 China
 E-Mail: yuxiuchen@zju.edu.cn

 <sup>2</sup> School of Mathematical Sciences Nankai University Tianjin 300071 China E-Mail: ghhu@nankai.edu.cn School of Mathematical Sciences Zhejiang University Hangzhou 310027 China E-Mail: wangtaolu@zju.edu.cn

3

Weierstrass Institute Mohrenstr. 39 10117 Berlin Germany E-Mail: Andreas.Rathsfeld@wias-berlin.de

> No. 2866 Berlin 2021



2020 Mathematics Subject Classification. 78A45, 78M10, 35J50, 65R20.

4

*Key words and phrases.* Wave scattering, locally defected periodic surface, boundary element method, perfectly matched layer, Neumann-to-Dirichlet operator, boundary integral equations, Riccati equation, Neumann-marching operator, recursive doubling procedure.

The third author is partially supported by NSF of Zhejiang Province for Distinguished Young Scholars (LR21A010001).

Edited by Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS) Leibniz-Institut im Forschungsverbund Berlin e. V. Mohrenstraße 39 10117 Berlin Germany

Fax:+493020372-303E-Mail:preprint@wias-berlin.deWorld Wide Web:http://www.wias-berlin.de/

# PML and high-accuracy boundary integral equation solver for wave scattering by a locally defected periodic surface

Xiuchen Yu, Guanghui Hu, Wangtao Lu, Andreas Rathsfeld

#### Abstract

This paper studies the perfectly-matched-layer (PML) method for wave scattering in a half space of homogeneous medium bounded by a two-dimensional, perfectly conducting, and locally defected periodic surface, and develops a high-accuracy boundary-integral-equation (BIE) solver. Along the vertical direction, we place a PML to truncate the unbounded domain onto a strip and prove that the PML solution converges to the true solution in the physical subregion of the strip with an error bounded by the reciprocal PML thickness. Laterally, we divide the unbounded strip into three regions: a region containing the defect and two semi-waveguide regions, separated by two vertical line segments. In both semi-waveguides, we prove the well-posedness of an associated scattering problem so as to well define a Neumann-to-Dirichlet (NtD) operator on the associated vertical segment. The two NtD operators, serving as exact lateral boundary conditions, reformulate the unbounded strip problem as a boundary value problem over the defected region. Due to the periodicity of the semi-waveguides, both NtD operators turn out to be closely related to a Neumann-marching operator, governed by a nonlinear Riccati equation. It is proved that the Neumann-marching operators are contracting, so that the PML solution decays exponentially fast along both lateral directions. The consequences culminate in two opposite aspects. Negatively, the PML solution cannot converge exponentially to the true solution in the whole physical region of the strip. Positively, from a numerical perspective, the Riccati equations can now be efficiently solved by a recursive doubling procedure and a high-accuracy PML-based BIE method so that the boundary value problem on the defected region can be solved efficiently and accurately. Numerical experiments demonstrate that the PML solution converges exponentially fast to the true solution in any compact subdomain of the strip.

# 1 Introduction

Due to its nearly reflectionless absorption of outgoing waves, perfectly matched layer (PML), since its invention by Bérenger in 1994 [4], has become a primary truncation technique in a broad class of unbounded wave scattering problems [11, 31, 16], ranging from quantum mechanics, acoustics, electromagnetism (optics), to seismology. Mathematically, a PML can be equivalently understood as a complexified transformation of a coordinate [12]. A wave outgoing along the coordinate is then analytically continued in the complex plane and becomes exponentially decaying in the PML. However, it is such a double-edged feature that makes PML be placed only in the direction where the medium structure is invariant so as to guarantee the validity of analytic continuation. Consequently, PML loses its prominence for some complicated structures, such as periodic structures [20]. Motivated by this, this paper studies wave scattering in a half space of homogeneous medium bounded by a two-dimensional, perfectly conducting, and locally defected periodic surface, and investigates the potential of PML in designing an accurate boundary integral equation (BIE) solver for the scattering problem.

Let a cylindrical wave due to a line source, or a downgoing plane wave be specified above the defected surface. Then, a primary question is to understand clearly how the scattered wave radiates at infinity.

Intrinsically, PML is highly related to the well-known Sommerfeld radiation condition (SRC), which, arguably, is an alternative way of saying *"the wave is purely outgoing at infinity"*. However, SRC is considered to be no longer valid for characterizing the scattered wave even when the surface is flat [2]. Instead, upward propagation radiation condition (UPRC), a.k.a. angular spectrum representation condition [14] is commonly used, and can well pose the present problem or even more general rough surface scattering problems [5, 7, 8]. Milder than SRC, UPRC only requires that the scattered wave contain no downgoing waves on top of a straight line above the surface, allowing waves incoming horizontally from infinity.

If the surface has no defects, the total wave field for the plane-wave incidence is quasi-periodic so that the original scattering problem can be formulated in a single unit cell, bounded laterally but unbounded vertically. According to UPRC, the scattered wave at infinity can then be expressed in terms of upgoing Bloch waves, so that a transparent boundary condition or PML of a local/nonlocal boundary condition can be successfully used to terminate the unit cell vertically. Readers are referred to [3, 10, 26, 34] and the references therein for related numerical methods as well as results on exponential convergence due to a PML truncation. But, if the incident wave is nonquasi-periodic, e.g., the cylindrical wave, or if the surface is locally defected, much fewer numerical methods or theories have been developed as it is no longer straightforward to laterally terminate the scattering domain. Existing laterally truncating techniques include recursive doubling procedure (RDP) [33, 15], Floquet-Bloch mode expansion [17, 19, 24], and Riccati-equation based exact boundary condition [21].

In a recent work [18], we proved that the total field for the cylindrical incidence, a.k.a. the Green function, satisfies the standard SRC on top of a straight line above the surface. Based on this, we further revealed that for the plane-wave incidence, the perturbed part of the total field due to the defect satisfies the SRC as well. Consequently, this suggests to use a PML to terminate the vertical variable so as to truncate the unbounded domain to a strip, bounded vertically but unbounded laterally. In fact, such a natural setup of PML had already been adopted in the literature [33, 6, 32], without a rigorous justification of the outgoing behavior, though. It is worthwhile to mention that Chandler-Wilde and Monk in [6] rigorously proved that under a Neumann-condition PML, the PML solution converges to the true solution in the whole physical region of the strip at a rate of only algebraic order of PML thickness. They further revealed that the PML solution due to the cylindrical incidence for a flat surface decays exponentially at infinity of a rectangular strip. However, it remains unclear how the PML solution radiates at infinity of the more generally curved strip under consideration. On the other hand, no literally rigorous theory has been developed to clearly understand why this PML-truncated strip can further be laterally truncated to a bounded domain by the aforementioned techniques without introducing artificial ill-posedness. In other words, the well-posedness of scattering problems in exterior regions of the truncated domain is unjustified.

To address these questions, we first prove in this paper that, under a Dirichlet-condition PML, the PML solution due to the cylindrical incidence, i.e., the Green function of the strip, converges to the true solution in the physical subregion of the strip at an algebraic order of the PML thickness. Next, we split the strip into three regions: a bounded region containing the defect and two semi-waveguide regions of a single-directional periodic surface, separated by two vertical line segments. By use of the Green function of the strip, transparent boundary conditions can be developed to truncate the unbounded semi-waveguides. Based on this, we apply the method of variational formulation and Fredholm's alternative to prove the well-posedness of the scattering problem in either semi-waveguide so as to define a Neumann-to-Dirichlet (NtD) operator on its associated vertical segment. The two NtD operators serve exactly as lateral boundary conditions to terminate the strip and to reformulate the unbounded strip problem as a boundary value problem on the defected region. Due to the periodicity of the semi-waveguides, both NtD operators turn out to be closely related to a Neumann-marching

operator, which is the solution of a nonlinear Riccati equation. It is proved that the Neumann-marching operators are contracting, indicating that the PML solution decays exponentially fast along both lateral directions even for the curved strip. The consequences culminate in two opposite aspects. Positively, from a numerical perspective, the Riccati equations can be efficiently solved by an RDP method so that the strip can be laterally truncated with ease. Negatively, the PML solution shall never exponentially converge to the true solution in the whole physical region of the strip. Nevertheless, as conjectured in [6], exponential convergence is optimistically expected to be realizable in any compact subdomain of the strip.

To validate the above conjecture numerically, we employ a high-accuracy PML-based boundary integral equation (BIE) method [28] to execute the RDP so that the two Riccati equations can be accurately solved for the two Neumann-marching operators, respectively, and hence the two NtD operators terminating the strip can be obtained. With the two NtD operators well-prepared, the boundary value problem in the defected region can be accurately solved by the PML-based BIE method again. By carrying out several numerical experiments, we observe that the PML truncation error for the wave field over the defected part of the surface decays exponentially fast as the PML absorbing strength or the thickness increases. This indicates that there is a chance that the PML solution still converges to the true solution exponentially in any compact subdomain of the strip, the justification of which remains open.

The remaining part of this paper is organized as follows. In Section 2, we introduce the half-space scattering problem and present some known well-posedness results. In Section 3, we introduce a Dirichlet-condition PML, prove the well-posedness of the PML-truncated problem and study the prior error estimate of the PML truncation. In Section 4, we study well-posedness of the semi-waveguide problems. In Section 5, we establish lateral boundary conditions, prove the exponentially decaying property of the PML solution at infinity of the strip, and develop an RDP technique to get the lateral boundary conditions. In Section 6, we present a PML-based BIE method to numerically solve the scattering problem. In Section 7, numerical experiments are carried out to demonstrate the performance of the proposed numerical method and to validate the proposed theory. We draw our conclusion finally in Section 8 and propose some future plans.

### 2 **Problem formulation**

Let  $\Omega \times \mathbb{R} \subset \mathbb{R}^3$  be an  $x_3$ -invariant domain bounded by a perfectly-conducting surface  $\Gamma \times \mathbb{R}$ , where  $\Gamma := \partial \Omega \subset \mathbb{R}^2$ , bounding domain  $\Omega \subset \mathbb{R}^2$ , is a local perturbation of a *T*-periodic curve  $\Gamma_T \subset \mathbb{R}^2$  periodic in  $x_1$ -direction (cf. Figure 1(a)). We denote points in the Cartesian coordinate system of  $\mathbb{R}^3$  by  $(x_1, x_2, x_3)$  and let  $x := (x_1, x_2) \in \Omega$ . Throughout this paper, we shall assume that  $\Gamma$  is Lipschitz and that  $\Omega$  satisfies the following geometrical condition

$$(\text{GC1}): (x_1, x_2) \in \Omega \Rightarrow (x_1, x_2 + a) \in \Omega, \quad \forall a \ge 0.$$

For simplicity, suppose  $\Gamma$  only perturbs one periodic part of  $\Gamma_T$ , say  $\{x \in \Gamma : |x_1| < T/2\}$ .

Let the unbounded domain  $\Omega \times \mathbb{R}$  be filled by a homogeneous medium of refractive index n. For a time-harmonic transverse-electric (TE) polarized electro-magnetic wave excited by an  $x_3$ -invariant incoming field of time dependence  $e^{-i\omega t}$  with angular frequency  $\omega$ , the  $x_3$ -component of the electric field, denoted by  $u^{\text{tot}}$ , is  $x_3$ -invariant. It has the same time dependence and satisfies the following boundary value problem for the two-dimensional (2D) Helmholtz equation:

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

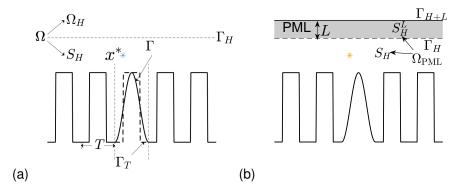


Figure 1: (a) A sketch of the half-space scattering problem. (b) A PML placed above  $\Gamma$ . The scattering surface  $\Gamma$  locally perturbs the periodic curve  $\Gamma_T$  of period T. Point  $x^*$  represents the point of an exciting source.  $\Gamma_H$  is an artificial interface, on which a DtN map is defined or above which a PML is placed.

$$u^{\text{tot}} = 0, \text{ on } \Gamma,$$
 (2)

where  $\Delta := \partial_{x_1}^2 + \partial_{x_2}^2$  is the 2D Laplacian and  $k := k_0 n$  with  $k_0 := 2\pi/\lambda$  denoting the free-space wavenumber for wavelength  $\lambda$ .

Let an incident wave  $u^{\text{inc}}$  be specified in  $\Omega$  and let  $x := (x_1, x_2) \in \Omega$ . In this paper, we shall mainly focus on the following two cases of incidence:

- (i) plane wave  $u^{\rm inc}(x) = e^{{\rm i}k(\cos\theta\,x_1 \sin\theta\,x_2)}$  for the incident angle  $\theta \in (0,\pi)$
- (ii) cylindrical wave  $u^{\text{inc}}(x;x^*) = G(x;x^*) := \frac{\mathbf{i}}{4}H_0^{(1)}(k|x-x^*|)$  excited by a source at  $x^* := (x_1^*, x_2^*) \in \Omega$ .

In the latter case, equation (1) should be replaced by

$$\Delta u^{\text{tot}} + k^2 u^{\text{tot}} = -\delta(x - x^*) \tag{3}$$

so that  $u^{\text{tot}}(x; x^*)$  in fact represents the Green function excited at the source point  $x^*$ . For simplicity, we assume that  $|x_1^*| < T/2$  so that  $x^*$  is right above the perturbed part of  $\Gamma$ .

Let  $u^{\mathrm{sc}} := u^{\mathrm{tot}} - u^{\mathrm{inc}}$  denote the scattered wave. One may enforce the following UPRC:

$$u^{\rm sc}(x) = 2 \int_{\Gamma_H} \frac{\partial G(x;y)}{\partial y_2} u^{\rm sc}(y) \mathrm{d}s(y), \tag{4}$$

where  $\Gamma_H := \{(x_1, H) : x_1 \in \mathbb{R}\}$  denotes a straight line strictly above  $\Gamma$  for some H > 0 and where  $y := (y_1, y_2)$ . According to [7], the UPRC helps to define the Dirichlet-to-Neumann map  $\mathcal{T} : H^{1/2}(\Gamma_H) \to H^{-1/2}(\Gamma_H)$  for the domain  $\Omega_H := \{x \in \Omega : x_2 > H\}$ , such that, for any  $\phi \in H^{1/2}(\Gamma_H)$ , we get  $\mathcal{T}\phi = \mathcal{F}^{-1}M_z\hat{\phi}$ , where  $\hat{\phi}(H;\xi) := [\mathcal{F}\phi](H;\xi)$  denotes the following normalized Fourier transform

$$[\mathcal{F}\phi](H;\xi) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \phi(x_1, H) e^{-\mathbf{i}\xi x_1} \mathrm{d}x_1,$$

and the operator  $M_z$  in the space of Fourier transforms is the operator of multiplication by

$$z(\xi) := \begin{cases} -\mathbf{i}\sqrt{k^2 - \xi^2}, & \text{for } |\xi| \le k, \\ \sqrt{\xi^2 - k^2}, & \text{for } |\xi| > k. \end{cases}$$
(5)

Then, we may enforce the boundary condition

$$\partial_{\nu} u^{\rm sc} = -\mathcal{T} u^{\rm sc}, \quad \text{on} \quad \Gamma_H,$$
 (6)

where  $\nu$  always denotes the outer unit normal vector on  $\Gamma_H$ . Condition UPRC guarantees the wellposedness of our scattering problem [7], but does not force  $u^{sc}$  to be purely outgoing at infinity, largely limiting its applications in designing numerical algorithms.

Nevertheless, our recent work [18] has shown a stronger Sommerfeld-type condition for the aforementioned two incidences, which still preserves the well-posedness. Note that [18] assumes further the following condition:

(GC2): some (and hence any) period of  $\Gamma_T$  contains a line segment,

which guarantees a local behavior of the Green function  $u^{\text{tot}}(x; y)$  for any x, y sufficiently close to each line segment. Let  $S_H := \Omega \cap \{x : x_2 < H\}$  be the strip between  $\Gamma_H$  and  $\Gamma$ . The radiation condition reads as follows:

(i). In the case of plane-wave incidence, the outgoing wave is  $u^{\text{og}} := u^{\text{tot}} - u^{\text{tot}}_{\text{ref}}$ , where  $u^{\text{tot}}_{\text{ref}}$  is the reference scattered field for the unperturbed scattering curve  $\Gamma_T$ , and satisfies the following half-plane Sommerfeld radiation condition (hSRC): For a sufficiently large R > 0 and any  $\rho < 0$ ,

$$\lim_{r \to \infty} \sup_{\alpha \in [0,\pi]} \sqrt{r} \left| \partial_r u^{\mathrm{og}}(x) - \mathbf{i} k u^{\mathrm{og}}(x) \right| = 0, \ \sup_{r \ge R} \sqrt{r} |u^{\mathrm{og}}(x)| < \infty, \ \mathrm{and} \ u^{\mathrm{og}} \in H^1_\rho(S^R_H),$$
(7)

where  $x = (r \cos \alpha, H + r \sin \alpha)$ ,  $S_H^R := S_H \cap \{x : |x_1| > R\}$ , and  $H_\rho^1(\cdot)$  denotes the weighted Sobolev space  $H_\rho^1(\cdot) := (1 + x_1^2)^{-\rho/2} H^1(\cdot)$ . We defer the computation of  $u_{\text{ref}}^{\text{tot}}$  to Section 6.3.

(ii). For the cylindrical incidence, the total field is the outgoing wave  $u^{\text{og}} := u^{\text{tot}}$  and satisfies the hSRC (7) in  $\Omega_H$ . Thus, the scattered field  $u^{\text{sc}}$  satisfies (7) as well since  $u^{\text{inc}}$  satisfies (7).

Certainly,  $u^{\text{og}}$  satisfies the UPRC condition (4) such that (6) holds for  $u^{\text{og}}$  in place of  $u^{\text{sc}}$  [9, Them. 2.9(ii)]. In the following, we shall consider the cylindrical incidence only and the plane-wave incidence case can be analyzed similarly.

We recall some important results from [7]. To remove the singularity of the right-hand side of (3), let

$$u_r^{\text{og}}(x;x^*) := u^{\text{og}}(x;x^*) - \chi(x;x^*)u^{\text{inc}}(x;x^*), \tag{8}$$

where the cut-off function  $\chi(x; x^*)$  is one in a neighborhood of  $x^*$  and has a sufficiently small support enclosing  $x^*$ . Let  $V_H := \{\phi|_{S_H} : \phi \in H_0^1(\Omega)\}$ . Then, it is equivalent to seek  $u_r^{\text{og}} \in V_H$  that satisfies the following boundary value problem:

$$\begin{aligned} \Delta u_r^{\text{og}} + k^2 u_r^{\text{og}} &= g, \quad \text{on} \quad S_H, \\ \partial_\nu u_r^{\text{og}} &= -\mathcal{T} u_r^{\text{og}}, \quad \text{on} \quad \Gamma_H, \end{aligned}$$

where  $g := -[\Delta \chi] u^{\text{inc}} - 2 \sum_{j=1}^{2} \partial_{x_j} \chi \partial_{x_j} u^{\text{inc}} \in L^2(S_H)$  such that  $\operatorname{supp} g$  is in the neighborhood of  $x^*$  contained in  $\overline{S}_H$ . An equivalent variational formulation reads as follows: Find  $u_r^{\text{og}} \in V_H$  such that, for any  $\phi \in V_H$ , there holds the variational equation  $b(u_r^{\text{og}}, \phi) = -\langle g, \phi \rangle_{S_H}$ , where the sesquilinear form  $b(\cdot, \cdot) \colon V_H \times V_H \to \mathbb{C}$  is given by

$$b(\phi,\psi) := \int_{S_H} \left( \nabla \phi \cdot \nabla \bar{\psi} - k^2 \phi \bar{\psi} \right) \mathrm{d}x + \int_{\Gamma_H} \mathcal{T} \phi \, \bar{\psi} \, \mathrm{d}s.$$

It has been shown in [7] that b satisfies the following inf-sup condition: For all  $v \in V_H$ ,

$$\gamma ||v||_{V_H} \le \sup_{\phi \in V_H} \frac{|b(v,\phi)|}{||\phi||_{V_H}},$$
(9)

where  $\gamma > 0$  depends on H, k and  $\Omega$ . Furthermore, b defines an invertible operator  $\mathcal{A} : V_H \to V_H^*$ such that  $\langle \mathcal{A}\phi, \psi \rangle = b(\phi, \psi)$  and  $||\mathcal{A}^{-1}|| \leq \gamma^{-1}$ . Thus, we obtain the representation  $u_r^{\text{og}} = -\mathcal{A}^{-1}g$ so that  $u^{\text{og}} = -\mathcal{A}^{-1}g + \chi u^{\text{inc}}$ .

The hSRC (7) suggests to compute the outgoing wave  $u^{\text{og}}$  numerically, as the PML technique[4, 6] could apply now to truncate the  $x_2$ -direction. In the following sections, we shall first introduce the setup of a PML to truncate  $x_2$  and then develop an accurate lateral boundary condition to truncate  $x_1$ .

### 3 PML Truncation

Mathematically, the PML truncating  $x_2$  introduces a complexified coordinate transformation

$$\tilde{x}_2 := x_2 + \mathbf{i}S \int_0^{x_2} \sigma(t) \mathrm{d}t,$$

where  $\sigma(x_2) = 0$  for  $x_2 \leq H$  and  $\sigma(x_2) \geq 0$  for  $x_2 \geq H$ . Note that such a tilde notation can also be used to define  $\tilde{y}_2$  and  $\tilde{x}_2^*$  in the following. As shown in Figure 1(b), the strip  $S_H^L := \mathbb{R} \times [H, H + L]$ with nonzero  $\sigma$  is called the PML region so that L represents its thickness. In this paper, we choose an  $m \geq 0$  and

$$\sigma(x_2) := \begin{cases} \frac{2f_2^m}{f_1^m + f_2^m}, & x_2 \in [H, H + L/2] \\ 2, & x_2 \ge H + L/2, m \ne 0 \\ 1, & x_2 \ge H + L/2, m = 0 \end{cases}$$
(10)

where we note that  $\sigma \equiv 1$  if m = 0, and

$$f_1 := \left(\frac{1}{2} - \frac{1}{m}\right)\xi^3 + \frac{\xi}{m} + \frac{1}{2}, \quad f_2 := 1 - f_1, \quad \xi := \frac{2x_2 - (2H + L/2)}{L/2}.$$

Let  $L_c := \tilde{x}_2(H+L) - H = L + iS_cL$ , where  $S_c := S/L \int_H^{H+L} \sigma(t) dt \ge S$ . Both the real and the imaginary part of  $L_c$  affect the absorbing strength of the PML (cf. [12]).

Now, let  $\tilde{x} := (x_1, \tilde{x}_2)$ . For  $x^* \in \Omega$  with  $x_2 > H$ , by analytic continuation of (4) we can define

$$u^{\mathrm{og}}(\tilde{x}; x^*) := 2 \int_{\Gamma_H} \frac{\partial G(\tilde{x}; y)}{\partial y_2} u^{\mathrm{og}}(y; x^*) \mathrm{d}s(y),$$

satisfying

$$\tilde{\Delta}u^{\mathrm{og}}(\tilde{x};x^*) + k^2 u^{\mathrm{og}}(\tilde{x};x^*) = -\delta(x-x^*),$$

where  $\tilde{\Delta} = \partial_{x_1}^2 + \partial_{\tilde{x}_2}^2$ . By chain rule, we see that  $\tilde{u}^{\text{og}}(x; x^*) := u^{\text{og}}(\tilde{x}; x^*)$  satisfies

$$\nabla \cdot (\mathbf{A}\nabla \tilde{u}^{\mathrm{og}}) + k^2 \alpha \tilde{u}^{\mathrm{og}} = -\delta(x - x^*), \quad \text{on } \Omega_{\mathrm{PML}},$$
(11)

$$\tilde{u}^{\mathrm{og}} = 0, \quad \mathrm{on} \ \Gamma.$$
 (12)

where  $\mathbf{A} := \operatorname{diag}\{\alpha(x_2), 1/\alpha(x_2))\}$  with  $\alpha(x_2) := 1 + \mathbf{i}S\sigma(x_2)$  and where the PML region  $\Omega_{\text{PML}} := \Omega \cap \{x : x_2 \leq H + L\}$  consists of the physical region  $S_H$  and the PML region  $S_H^L$ .

On the PML boundary  $\Gamma_{H+L} := \{x : x_2 = H + L\}$ , we use the homogeneous Dirichlet boundary condition

$$\tilde{u}^{\text{og}} = 0, \quad \text{on} \quad \Gamma_{H+L}.$$
 (13)

The authors in [6] adopted a Neumann condition on the PML boundary  $\Gamma_{H+L}$  and proved the wellposedness of the related PML truncation problem. Here, we choose the Dirichlet condition (13) since, as we shall see, our numerical results indicate that the Dirichlet-PML seems more stable than the Neumann-PML. Furthermore, we need the Green function of the strip  $\tilde{u}^{\text{og}}(x; x^*)$  for any  $x^* \in \Omega_{\text{PML}}$ and not only for  $x^* \in S_H$  to establish lateral boundary conditions. For completeness, we shall, following the idea of [6], study the well-posedness of the problem (11-13) for any  $x^* \in \Omega_{\text{PML}}$ .

The fundamental solution of the anisotropic Helmholtz equation (11) is (cf. [28])

$$\widetilde{G}(x;y) := G(\widetilde{x};\widetilde{y}) := \frac{\mathbf{i}}{4} H_0^{(1)} \Big( k\rho(\widetilde{x};\widetilde{y}) \Big), \tag{14}$$

where  $\tilde{y} := (y_1, \tilde{y}_2)$ , where the complexified distance function  $\rho$  is defined to be

$$\rho(\tilde{x}, \tilde{y}) := \left[ (x_1 - y_1)^2 + (\tilde{x}_2 - \tilde{y}_2)^2 \right]^{1/2} ,$$

and where the half-power operator  $z^{1/2}$  is chosen to be the branch of  $\sqrt{z}$  with nonnegative real part for  $z \in \mathbb{C} \setminus (-\infty, 0]$  such that  $\arg(z^{1/2}) \in [0, \pi)$ . The special choice of the function  $\sigma$  in (10) ensures  $\widetilde{G}(x; y) = \widetilde{G}(x; y_{\text{imag}})$  for any  $x \in \Gamma_{H+L}$ , whenever  $y := (y_1, y_2)$  and  $y_{\text{imag}} := (y_1, 2(H+L) - y_2)$ , the mirror image of y w.r.t. the line  $\Gamma_{H+L}$ , are sufficiently close to  $\Gamma_{H+L}$  so that  $\rho(\tilde{x}; \tilde{y}) = \rho(\tilde{x}; \tilde{y}_{\text{imag}})$ .

To remove the singularity of the right-hand side of (11), we introduce

$$\tilde{u}_r^{\mathrm{og}}(x;x^*) := \tilde{u}^{\mathrm{og}}(x;x^*) - \chi(x;x^*)\tilde{u}^{\mathrm{inc}}(x;x^*),$$

with the same cut-off function  $\chi$  as in (8), where  $\tilde{u}^{\rm inc}(x;x^*) := u^{\rm inc}(\tilde{x};\tilde{x}^*)$ . Then,  $\tilde{u}_r^{\rm og}$  satisfies

$$\nabla \cdot (\mathbf{A} \nabla \tilde{u}_r^{\text{og}}) + k^2 \alpha \tilde{u}_r^{\text{og}} = \tilde{g}^{\text{inc}}, \quad \text{on} \quad \Omega_{\text{PML}}, \tag{15}$$

$$\tilde{u}_r^{\mathrm{og}} = 0, \quad \text{on } \Gamma,$$
 (16)

$$\tilde{u}_r^{\text{og}} = 0, \quad \text{on} \quad \Gamma_{H+L},$$
(17)

where  $\tilde{g}^{\text{inc}} := [\nabla \cdot (\mathbf{A}\nabla) + k^2 \alpha](1 - \chi(x;x^*))\tilde{u}^{\text{inc}}(x;x^*) \in L^2(\Omega_{\text{PML}})$  with  $\operatorname{supp} \tilde{g}^{\text{inc}}$  included in  $\overline{\Omega}_{\text{PML}} = \overline{S}_H \cup \overline{S}_H^L$ . Taking into account that  $x^*$  can be located in  $S_H^L$ , the support  $\operatorname{supp} \tilde{g}^{\text{inc}}$  may not completely lie in the physical domain  $S_H$ . To establish a Dirichlet-to-Neumann map on  $\Gamma_H$  like (6), we need to study the following boundary value problem in the PML strip  $S_H^L$ : Given  $q \in H^{1/2}(\Gamma_H)$ ,  $s \in H^{1/2}(\Gamma_{H+L})$ , and  $\tilde{g}_{\text{PML}}^{\text{inc}} := \tilde{g}^{\text{inc}}|_{S_H^L} \in L^2(S_H^L)$  with  $\operatorname{supp} \tilde{g}_{\text{PML}}^{\text{inc}} \subset \overline{S}_H^L$ , find  $v \in H^1(S_H^L)$  such that

$$\nabla \cdot (\mathbf{A} \nabla v) + k^2 \alpha v = \tilde{g}_{\text{PML}}^{\text{inc}}, \text{ on } S_H^L,$$
$$v = q, \quad \text{ on } \Gamma_H,$$
$$v = s, \quad \text{ on } \Gamma_{H+L}$$

Let

$$v_0(x) := v(x) - v_{\mathrm{PML}}^{\mathrm{inc}}(x), \quad v_{\mathrm{PML}}^{\mathrm{inc}}(x) := \int_{S_H^L} \left[ \widetilde{G}(x;y) - \widetilde{G}(x;y_{\mathrm{imag}}) \right] \widetilde{g}_{\mathrm{PML}}^{\mathrm{inc}}(y) \mathrm{d}y,$$

where we recall that  $y_{\text{imag}}$  is the mirror image of y w.r.t. the line  $\Gamma_{L+H}$ . Thus,  $v_0$  satisfies

$$\nabla \cdot (\mathbf{A} \nabla v_0) + k^2 \alpha v_0 = 0, \quad \text{on} \quad S_H^L, \qquad (18)$$

$$v_0 = q_n, \quad \text{on} \quad \Gamma_H, \\ v_0 = s_n, \quad \text{on} \quad \Gamma_{H+L}, \qquad (18)$$
where  $q_n := q - v_{\text{PML}}^{\text{inc}}|_{\Gamma_H} \in H^{1/2}(\Gamma_H)$  and  $s_n := s - v_{\text{PML}}^{\text{inc}}|_{\Gamma_{H+L}} \in H^{1/2}(\Gamma_{H+L}).$ 

Now (18) is equivalent to the complexified Helmholtz equation and the Fourier transform of this w.r.t.  $\tilde{x}_1$  is a simple ordinary differential equation w.r.t.  $\tilde{x}_2$ . The general solution of this differential equation is in the span of two exponential functions of  $\tilde{x}_2$ , i.e. of two complexified generalized plane waves. In

other words, looking for  $\mathcal{F}v_0$  in terms of these plane waves, we get

$$\hat{v}_0(x_2;\xi) := [\mathcal{F}v_0](x_2;\xi) = A(\xi) \exp\left(z(\xi)(\tilde{x}_2 - H)\right) + B(\xi) \exp\left(-z(\xi)(\tilde{x}_2 - H)\right), \quad (19)$$

where we recall that z has been defined in (5),

$$A(\xi) = \frac{\hat{s}_n(\xi) - \exp\left(-z(\xi)L_c\right)\hat{q}_n(\xi)}{\exp\left(z(\xi)L_c\right) - \exp\left(-z(\xi)L_c\right)}, \quad B(\xi) = \frac{-\hat{s}_n(\xi) + \exp\left(z(\xi)L_c\right)\hat{q}_n(\xi)}{\exp\left(z(\xi)L_c\right) - \exp\left(-z(\xi)L_c\right)},$$

 $\hat{s}_n(\xi) = [\mathcal{F}s_n](H + L; \xi)$  and  $\hat{q}_n(\xi) = [\mathcal{F}q_n](H; \xi)$ . Here, to make A and B well-defined, we could let  $\xi$  travel through a Sommerfeld integral path  $-\infty + 0\mathbf{i} \to 0 \to \infty - 0\mathbf{i}$  instead of through  $\mathbb{R}$  (cf. [25]) such that  $z \neq 0$ . Consequently,

$$-\frac{\partial \hat{v}_0}{\partial x_2}\Big|_{x_2=H} = z \frac{-2}{\exp(zL_c) - \exp(-zL_c)} \hat{s}_n + z \frac{\exp(zL_c) + \exp(-zL_c)}{\exp(zL_c) - \exp(-zL_c)} \hat{q}_n.$$

Now define two bounded operators  $\mathcal{T}_p: H^{1/2}(\Gamma_H) o H^{-1/2}(\Gamma_H)$  by

$$\mathcal{F}[\mathcal{T}_p q_n](H;\xi) = z(\xi) \frac{\exp(z(\xi)L_c) + \exp(-z(\xi)L_c)}{\exp(z(\xi)L_c) - \exp(-z(\xi)L_c)} \hat{q}_n(\xi),$$

and  $\mathcal{N}_p: H^{1/2}(\Gamma_{H+L}) \to H^{-1/2}(\Gamma_H)$  by

$$\mathcal{F}[\mathcal{N}_p s_n](H+L;\xi) = z(\xi) \frac{-2}{\exp(z(\xi)L_c) - \exp(-z(\xi)L_c)} \hat{s}_n(\xi).$$

Note that the above definitions allow  $\xi \in \mathbb{R}$  now, since limits can be considered when z = 0. Returning back to the PML-truncated problem (15-17), we reformulate it as an equivalent boundary value problem on the physical region  $S_H$ : Find  $\tilde{u}_r^{\text{og}} \in V_H$  that satisfies

$$\nabla \cdot (\mathbf{A} \nabla \tilde{u}_r^{\mathrm{og}}) + k^2 \alpha \tilde{u}_r^{\mathrm{og}} = \tilde{g}^{\mathrm{inc}}|_{S_H}, \quad \text{on} \quad S_H, \\ \partial_{\nu} \tilde{u}_r^{\mathrm{og}} = -\mathcal{T}_p \left( \tilde{u}_r^{\mathrm{og}}|_{\Gamma_H} \right) + f_p, \text{ on} \quad \Gamma_H,$$

where

$$f_p := \mathcal{N}_p \left( v_{\mathrm{PML}}^{\mathrm{inc}} |_{\Gamma_{H+L}} \right) + \mathcal{T}_p \left( v_{\mathrm{PML}}^{\mathrm{inc}} |_{\Gamma_H} \right) + \partial_\nu v_{\mathrm{PML}}^{\mathrm{inc}} |_{\Gamma_H} \in H^{-1/2}(\Gamma_H).$$

The associated variational formulation reads as follows: Find  $\tilde{u}_r^{\mathrm{og}} \in V_H$ , such that for any  $\psi \in V_H$ ,

$$b_p(\tilde{u}_r^{\text{og}}, \psi) = -\int_{S_H} \tilde{g}^{\text{inc}}|_{S_H} \bar{\psi} \, \mathrm{d}x + \int_{\Gamma_H} f_p \bar{\psi} \, \mathrm{d}s,$$
(20)

DOI 10.20347/WIAS.PREPRINT.2866

Berlin 2021

where the sesquilinear form  $b_p(\cdot,\cdot):V_H\times V_H\to \mathbb{C}$  is given by

$$b_p(\phi,\psi) := \int_{S_H} \left( \nabla \phi \cdot \nabla \bar{\psi} - k^2 \phi \bar{\psi} \right) \mathrm{d}x + \int_{\Gamma_H} \bar{\psi} \mathcal{T}_p \phi \, \mathrm{d}s.$$
(21)

As in [6], we define the k-dependent norm

$$||\phi||_{H^s(\mathbb{R})}^2 = \int_{\mathbb{R}} (k^2 + \xi^2)^s |[\mathcal{F}\phi](\xi)|^2 \mathrm{d}\xi$$

for  $H^s(\mathbb{R}).$  Then, the following lemma roughly characterizes the difference of  $\mathcal{T}_p$  and  $\mathcal{T}.$ 

**Lemma 3.1** For any L with  $kS_cL > 0$ , we have

$$||\mathcal{T} - \mathcal{T}_p|| \le \frac{1}{kS_cL} + \frac{2}{\sqrt{3}} \frac{\exp(-[2kL])}{[2kL]}.$$

Proof. By a simple analysis, it can be seen that

$$\begin{aligned} ||\mathcal{T} - \mathcal{T}_{p}|| &= \sup_{\xi \in \mathbb{R}} \frac{|z(\xi)|}{\sqrt{k^{2} + \xi^{2}}} |1 - \coth(z(\xi)L_{c})| = \max\{S_{1}, S_{2}\}, \\ S_{1} &:= \sup_{\xi \in \mathbb{R}: |\xi| \le k} \frac{2 \left| z(\xi) \exp\left( - 2z(\xi)L_{c} \right) \right|}{\sqrt{k^{2} + \xi^{2}} \left| 1 - \exp\left( - 2z(\xi)L_{c} \right) \right|}, \\ S_{2} &:= \sup_{\xi \in \mathbb{R}: |\xi| \ge k} \frac{2 \left| z(\xi) \exp\left( - 2z(\xi)L_{c} \right) \right|}{\sqrt{k^{2} + \xi^{2}} \left| 1 - \exp\left( - 2z(\xi)L_{c} \right) \right|}. \end{aligned}$$

Recalling  $L_c = L + \mathbf{i} S_c L$  and setting  $t := \sqrt{|k^2 - x^2|}/k$ , we arrive at

$$S_{1} = \sup_{0 \le t \le 1} \frac{2t \exp(-2tkS_{c}L)}{\sqrt{2 - t^{2}}\sqrt{1 + \exp(-4tkL) - 2\cos(2tkS_{c}L)\exp(-2tkL)}}$$
  
= 
$$\sup_{0 \le t \le 1} \frac{2t \exp(-2tkS_{c}L)}{\sqrt{2 - t^{2}}\sqrt{(1 - \exp(-2tkS_{c}L))^{2} + 4\exp(-2tkS_{c}L)\sin^{2}(tkL)}}$$

and

$$S_2 = \sup_{t \ge 1} \frac{2t \exp(-2tkL)}{\sqrt{2+t^2}\sqrt{(1-\exp(-2tkL))^2 + 4\exp(-2tkL)\sin^2(tkS_cL)}}$$

Since, for  $t \ge 0$ , the function

$$f(t) = \frac{t \exp(-2tkS_cL)}{1 - \exp(-2tkS_cL)}$$

is nonincreasing, it is easy to see that  $S_1 \leq 2f(0) = 1/[kS_cL]$ . One similarly shows that

$$S_2 \le \frac{2}{\sqrt{3}} \frac{\exp(-2kL)}{2kL}.$$

We do not intend to study the further relation of  $||\mathcal{T}_p - \mathcal{T}||$  and parameter  $S_c$  as done in [6], where the performance of the PML is optimized. For our purpose, the estimate in Lemma 3.1 is enough. Clearly, the sesquilinear form  $b_p$  in (21) defines a bounded linear functional  $\mathcal{A}_p : V_H \to V_H^*$  such that, for any  $\phi \in V_H$ ,

$$\left\langle (\mathcal{A} - \mathcal{A}_p)\phi, \psi \right\rangle = b(\phi, \psi) - b_p(\phi, \psi) = \int_{\Gamma_H} \bar{\psi}(\mathcal{T} - \mathcal{T}_p)\phi \, \mathrm{d}s.$$

Analogous to [6, Sec. 3], we see immediately that

$$||\mathcal{A} - \mathcal{A}_p|| \le 2||\mathcal{T} - \mathcal{T}_p|| \le \frac{2}{kS_cL} + \frac{2}{\sqrt{3}} \frac{\exp(-[2kL])}{[2kL]}.$$

Consequently, since A is invertible,  $A_p$  has a bounded inverse provided that  $S_cL$  and L are sufficiently large. Since the right-hand side of (20) defines a bounded functional in  $V_H^*$ , we in fact have justified the following well-posedness result.

**Theorem 3.2** Provided that L and  $S_cL$  (note  $S_cL \ge SL$ ) are sufficiently large, the PML-truncated problem (11), (12) and (13) admits a unique solution  $\tilde{u}^{\mathrm{og}}(x;x^*) = \tilde{u}_r^{\mathrm{og}}(x;x^*) + \chi(x;x^*)\tilde{u}^{\mathrm{inc}}(x;x^*)$  with  $\tilde{u}_r^{\mathrm{og}} \in H_0^1(\Omega_{\mathrm{PML}}) := \{\phi \in H^1(\Omega_{\mathrm{PML}}) : \phi|_{\Gamma \cup \Gamma_{H+L}} = 0\}$  for any  $x^* \in \Omega_{\mathrm{PML}}$ . Moreover, there holds the estimate  $||\tilde{u}_r^{\mathrm{og}}(\cdot;x^*)||_{H^1(\Omega_{\mathrm{PML}})} \le C||\tilde{g}^{\mathrm{inc}}||_{L^2(\Omega_{\mathrm{PML}})}$ .

**Remark 3.3** The well-posedness in Theorem 3.2 holds in general for any Lipschitz curve satisfying (GC1).

Since, for any  $\phi \in V_H$ ,

$$b_p(\phi,\psi) = b(\phi,\psi) - \int_{\Gamma_H} \bar{\psi}(\mathcal{T}-\mathcal{T}_p)\phi \,\mathrm{d}s,$$

the inf-sup condition (9) of b implies the inf-sup condition of  $b_p$ : For any  $\phi \in V_H$ ,

$$\sup_{\psi \in V_{H}} \frac{|b_{p}(\phi,\psi)|}{||\psi||_{V_{H}}} \geq \sup_{\psi \in V_{H}} \frac{|b(\phi,\psi)|}{||\psi||_{V_{H}}} - \left\{ \frac{2}{kS_{c}L} + \frac{2}{\sqrt{3}} \frac{\exp(-[2kL])}{[2kL]} \right\} ||\phi||_{V_{H}} \quad (22)$$

$$\geq \left( \gamma - \frac{2}{kS_{c}L} - \frac{2}{\sqrt{3}} \frac{\exp(-[2kL])}{[2kL]} \right) ||\phi||_{V_{H}},$$

provided  $S_cL$  and L are sufficiently large. As a consequence of (22), we immediately obtain the prior error estimate for the PML truncation if  $x^* \in S_H$ .

**Corollary 3.4** Provided that  $S_cL$  and L are sufficiently large, there holds

$$||u^{\rm og}(\cdot;x^*) - \tilde{u}^{\rm og}(\cdot;x^*)||_{V_H} \le \frac{2}{\gamma \min\left\{[kS_cL], \sqrt{3}[2kL]\exp([2kL])\right\} - 2} ||u^{\rm og}_r(\cdot;x^*)||_{V_H}$$
(23)

whenever  $x^* \in S_H$ .

*Proof.* Since, for  $x^* \in S_H$ ,  $(u^{\text{og}} - \tilde{u}^{\text{og}})|_{S_H} = (u_r^{\text{og}} - \tilde{u}_r^{\text{og}})|_{S_H} \in V_H$ , we have, for any  $\phi \in V_H$ ,

$$b_p(u_r^{\mathrm{og}} - \tilde{u}_r^{\mathrm{og}}, \phi) = -\int_{\Gamma_H} \bar{\phi}(\mathcal{T} - \mathcal{T}_p) u_r^{\mathrm{og}} \mathrm{d}s$$

so that by the inf-sup condition (22),

$$\begin{aligned} ||u_{r}^{\text{og}} - \tilde{u}_{r}^{\text{og}}||_{V_{H}} \\ &\leq \left(\gamma - \frac{2}{kS_{c}L} - \frac{2}{\sqrt{3}} \frac{\exp(-[2kL])}{[2kL]}\right)^{-1} \sup_{\phi \in V_{H}} \frac{|b_{p}(u_{r}^{\text{og}} - \tilde{u}_{r}^{\text{og}}, \phi)|}{||\phi||_{V_{H}}} \\ &= \left(\gamma - \frac{2}{kS_{c}L} - \frac{2}{\sqrt{3}} \frac{\exp(-[2kL])}{[2kL]}\right)^{-1} \sup_{\phi \in V_{H}} \frac{|\int_{\Gamma_{H}} \bar{\phi}(\mathcal{T} - \mathcal{T}_{p})u_{r}^{\text{og}} \mathrm{d}s|}{||\phi||_{V_{H}}} \\ &\leq \left(\gamma - \frac{2}{kS_{c}L} - \frac{2}{\sqrt{3}} \frac{\exp(-[2kL])}{[2kL]}\right)^{-1} \left(\frac{2}{kS_{c}L} + \frac{2}{\sqrt{3}} \frac{\exp(-[2kL])}{[2kL]}\right) ||u_{r}^{\text{og}}||_{V_{H}}. \end{aligned}$$

## 4 Semi-waveguide problems

Unlike the exponential convergence results in [10, 34], (23) indicates only a poor convergence of the PML method over  $S_H$ . However, we believe that exponential convergence can be realized in a compact subset of  $S_H$ , which is indeed true if  $\Gamma$  is flat [6]. Then the vertical PML truncation is efficient and the next essential question is how to accurately truncate  $\Omega_{PML}$  in the lateral  $x_1$ -direction. To address this question, as inspired by [21] and as illustrated in Figure 2 (a), we shall consider the following two semi-waveguide problems:

$$(P^{\pm}): \begin{cases} \nabla \cdot (\mathbf{A}\nabla \tilde{u}) + k^{2}\alpha \tilde{u} = 0, & \text{on } \Omega_{\text{PML}}^{\pm} := \Omega_{\text{PML}} \quad \cap \left\{ x : \pm x_{1} > \frac{T}{2} \right\}, \\ \tilde{u} = 0, & \text{on } \Gamma^{\pm} := \Gamma \quad \cap \left\{ x : \pm x_{1} > \frac{T}{2} \right\}, \\ \tilde{u} = 0, & \text{on } \Gamma_{L+H}^{\pm} := \Gamma_{L+H} \quad \cap \left\{ x : \pm x_{1} > \frac{T}{2} \right\}, \\ \partial_{\nu_{c}} \tilde{u} = g^{\pm}, & \text{on } \Gamma_{0}^{\pm} := \Omega_{\text{PML}} \quad \cap \left\{ x : x_{1} = \pm \frac{T}{2} \right\}, \end{cases}$$

for given Neumann data  $g^{\pm} \in H^{-1/2}(\Gamma_0^{\pm})$ , where  $\nu_c := \mathbf{A}\nu$  denotes the co-normal vector with  $\nu$  pointing towards  $\Omega_{PML}^{\pm}$ ,  $\tilde{u}$  denotes a generic field, and we note that  $\Gamma^{\pm} \subset \Gamma_T$  does not contain the defected part  $\Gamma_0$ . In this section, we shall study the well-posedness of the semi-waveguide problems  $(P^{\pm})$ . By Theorem 3.2, the following uniqueness result is easy to obtain.

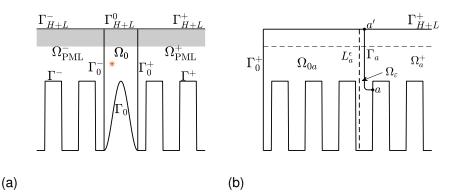


Figure 2: (a) Region  $\Omega_{\rm PML}$  is divided into three regions  $\Omega_{\rm PML}^-$ ,  $\Omega_0$  and  $\Omega_{\rm PML}^+$ ; semi-waveguide problems  $(P^{\pm})$  are defined in  $\Omega^{\pm}$  bounded by  $\Gamma_{H+L}^{\pm}$ ,  $\Gamma_0^{\pm}$  and  $\Gamma^{\pm}$ . (b) Domain  $\Omega_{\rm PML}^+$  is further truncated onto  $\Omega_{0a}$  by a smooth curve  $\Gamma_a$  intersecting  $\Gamma^+$  and  $\Gamma_{H+L}^+$  perpendicularly at a and a', respectively.  $\Omega_a^+ = \Omega_{\rm PML}^+ \backslash \overline{\Omega}_{0a}$  and the auxiliary line  $L_a^{\epsilon}$  is chosen such that the domain  $\Omega_{\epsilon}$  is sufficiently narrow.

**Lemma 4.1** Provided that  $S_cL$  and L are sufficiently large, problem  $(P^{\pm})$  has at most one solution in  $H^1(\Omega_{PML}^{\pm})$ .

*Proof.* Suppose  $\tilde{u} \in H^1(\Omega_{PML}^+)$  satisfies  $(P^+)$  with  $g^+ = 0$ . Let

$$\Omega^{e}_{\rm PML} = \left\{ x \in \mathbb{R}^{2} : (x_{1}, x_{2}) \in \Omega^{+}_{\rm PML} \text{ or } (T - x_{1}, x_{2}) \in \Omega^{+}_{\rm PML} \right\} \cup \Gamma^{+}_{0}, 
\Gamma^{e} = \left\{ x \in \mathbb{R}^{2} : (x_{1}, x_{2}) \in \Gamma^{+} \text{ or } (T - x_{1}, x_{2}) \in \Gamma^{+} \text{ or } (T/2, x_{2}) \in \Gamma \right\}.$$

Then,

$$\tilde{u}^e(x_1, x_2) = \begin{cases} \tilde{u}(x_1, x_2), & x_1 \ge T/2, \\ -\tilde{u}(T - x_1, x_2), & x_1 < T/2, \end{cases}$$

in  $H^1(\Omega^e_{\text{PML}})$  satisfies problem (15-17) with  $\tilde{g}$ ,  $\Omega_{\text{PML}}$  and  $\Gamma$  replaced by 0,  $\Omega^e_{\text{PML}}$  and  $\Gamma^e$ , respectively. Theorem 3.2 and Remark 3.3 imply that  $\tilde{u}^e = 0$  on  $\Omega^e_{\text{PML}}$  so that  $\tilde{u} = \tilde{u}^e|_{\Omega^+_{\text{PML}}} = 0$ . The uniqueness of problem  $(P^-)$  can be established similarly.  $\Box$ 

We are ready to study the well-posedness of problem  $(P^{\pm})$  by Fredholm's alternative. Without loss of generality, we shall study  $(P^{+})$  only. To make use of the Fredholm theory, we need first to truncate  $\Omega_{\rm PML}^{+}$  by an exact transparent boundary condition. Under condition (GC2), there exists a line segment  $L_a \subset \Gamma_T \cap \Gamma$  with the midpoint  $a := (a_1, a_2) \in L_a$  for  $a_1 > T/2$ . For a small fixed constant  $\epsilon > 0$ , we can find a vertical line segment  $L_a^{\epsilon}$  and a simple and smooth curve  $\Gamma_a \subset \Omega_{\rm PML}$  connecting  $L_a$  and  $\Gamma_{H+L}^{+}$  such that the distance of  $L_a^{\epsilon}$  and  $\Gamma_a$  is  $\epsilon$ , that  $\{x \in \Gamma_a : x_2 > H\}$  is on a vertical line and that  $\Gamma_a$  intersects  $L_a$  perpendicularly at a (cf. Figure 2 (b)). Let  $\Omega_{\epsilon}$  be the domain bounded by  $\Gamma_a, L_a^{\epsilon}$ ,  $L_a$  and  $\Gamma_{H+L}^{+}$ , and  $\Omega_a^{+}$  be the unbounded domain bounded by  $\Gamma_a, \Gamma_{H+L}^{+}$  and  $\Gamma$ . For sufficiently small  $\epsilon$ , the above choice of  $L_a^{\epsilon}$  and  $\Gamma_a$  guarantees that  $k^2 > 0$  is not an eigenvalue of

$$-\nabla \cdot (\mathbf{A}\nabla \tilde{u}) = k^2 \alpha \tilde{u}, \quad \text{on} \quad \Omega_{\epsilon}, \tag{24}$$

$$\tilde{u} = 0, \quad \text{on } \partial \Omega_{\epsilon}.$$
 (25)

Now, for the unbounded domain  $\Omega_{\epsilon}^+ = \Omega_{\epsilon} \cup \Gamma_a \cup \Omega_a^+$ , by a symmetrical reflection w.r.t. the line containing  $L_a^{\epsilon}$ , the partial boundary  $\partial \Omega_{\epsilon}^+ \cap \Gamma$  can be extended to a Lipschitz boundary, denoted by  $\Gamma_{\epsilon}$ , satisfying (GC1). Then, Theorem 3.2, with  $\Gamma_{\epsilon}$  in place of  $\Gamma$ , can help to construct the Dirichlet Green function of  $\Omega_{\epsilon}^+$  by  $\widetilde{G}_{\mathrm{D}}(x;y) := \widetilde{u}^{\mathrm{og}}(x;y) - \widetilde{u}^{\mathrm{og}}(x;y_{\mathrm{imag}}^{\epsilon})$  satisfying  $\widetilde{G}_{\mathrm{D}}(\cdot;y)|_{\partial \Omega_{\epsilon}^+} = 0$ , where  $y_{\mathrm{imag}}^{\epsilon}$  is the mirror image of the source point y w.r.t. the line  $L_a^{\epsilon}$ . Choosing  $\Gamma_a$  in such a special way, the following local regularity property of  $\widetilde{G}_{\mathrm{D}}(x;y)$  can be ensured.

**Proposition 4.2** Under the geometrical conditions (GC1) and (GC2), for sufficiently large values of L,  $S_cL$  and m in (10),  $\tilde{G}_D(x; y)$  admits the decomposition

$$\widetilde{G}_{\mathrm{D}}(x;y) = \widetilde{G}(x;y) - \widetilde{G}(x;y_{\mathrm{imag}}^{l}) + R_{l}(x,y), \ y \in \overline{\Omega}_{l}, \ l = a, a', \ x \in \overline{\Omega}_{l} \cup \overline{\Omega}_{c}$$

such that  $R_l(x, y)$  is a sufficiently smooth function of  $(x, y) \in \overline{\Omega}_l \cup \overline{\Omega}_c \times \overline{\Omega}_l$ , where  $\widetilde{G}$  is defined by (14),  $\Omega_l$  is a sufficiently small neighborhood of point l in  $\Omega_{\epsilon}^+$ , and  $\Omega_c$  can be any bounded subset of  $\Omega_{\epsilon}^+$ . Furthermore,  $y_{imag}^a$  and  $y_{imag}^{a'}$  are the mirror images of y w.r.t. the line containing the segment  $L_a$  and the line  $\Gamma_{L+H}$ , respectively.

*Proof.* We consider y close to point a' only. Define

$$\tilde{u}_{a'}(x;y) := \tilde{u}^{\mathrm{og}}(x;y) - \chi_{a'}(x) \left[ \widetilde{G}(x;y) - \widetilde{G}(x;y_{\mathrm{imag}}^{a'}) \right],$$

where the cut-off function  $\chi_{a'} \equiv 1$  in a neighborhood of a' and has a small support that is independent of y. Then, it can be seen that  $\tilde{u}_{a'}$  satisfies (15-17) with  $\tilde{g}^{\text{inc}}$  replaced by

$$\left[\nabla \cdot (\mathbf{A}\nabla) + k^2 \alpha\right] \left(1 - \chi_{a'}(x)\right) \left[\widetilde{G}(x;y) - \widetilde{G}(x;y_{\text{imag}}^{a'})\right] \in C^{m-1}_{\text{comp}}(\Omega_{\text{PML}} \times \overline{\Omega}_l),$$

where  $C_{\text{comp}}^{m-1}$  consists of m-1 times differentiable functions with compact supports. Note that m defined in (10) determines the smoothness of  $\sigma$ . By arguing the same way as in [18, Lem 2.4] and by choosing m sufficiently large,  $R_{a'}(x, y) = \tilde{u}_{a'}(x; y) - \tilde{u}^{\text{og}}(x; y_{\text{imag}}^{\epsilon})$  becomes a sufficiently smooth function for  $(x, y) \in \overline{\Omega}_l \cup \overline{\Omega}_c \times \overline{\Omega}_l$ .  $\Box$ 

On  $\Gamma_a,$  we now define the following two integral operators:

$$[\mathcal{S}_a \phi](x) = 2 \int_{\Gamma_a} \widetilde{G}_{\mathrm{D}}(x; y) \phi(y) \,\mathrm{d}s(y),$$
  
$$[\mathcal{K}_a \phi](x) = 2 \int_{\Gamma_a} \partial_{\nu_c(y)} \widetilde{G}_{\mathrm{D}}(x; y) \phi(y) \,\mathrm{d}s(y),$$

Note that, in general, the complexified double layer operator is a strongly singular operator, which is to be defined by principle value integrals. In our case, however, the complexification is given by (10) and the curve  $\Gamma_a$  is vertical in the complexified region above  $\Gamma_H$ . From the local asymptotics of the Hankel function it is not hard to conclude that the double layer kernel is weakly singular. Proposition 4.2 reveals that classic mapping properties hold for the above two integral operators on the open arc  $\Gamma_a$ .

Lemma 4.3 We can uniquely extend the operator  $S_a$  as a bounded operator from  $H^{-1/2}(\Gamma_a)$  to  $\widetilde{H}^{1/2}(\Gamma_a)$ , operator  $\mathcal{K}_a$  as a compact (and certainly bounded) operator from  $\widetilde{H}^{1/2}(\Gamma_a)$  to  $\widetilde{H}^{1/2}(\Gamma_a)$ . Moreover, we have the decomposition  $S_a = S_{p,a} + \mathcal{L}_{p,a}$  such that  $\mathcal{L}_{p,a} \colon H^{-1/2}(\Gamma_a) \to \widetilde{H}^{1/2}(\Gamma_a)$  is compact and  $S_{p,a} \colon H^{-1/2}(\Gamma_a) \to \widetilde{H}^{1/2}(\Gamma_a)$  has a positive definite real part, i.e., for some constant c > 0,

$$\operatorname{Re}\left(\int_{\Gamma_a} \mathcal{S}_{p,a} \phi \bar{\phi} \,\mathrm{d}s\right) \ge c ||\phi||_{H^{-1/2}(\Gamma_a)}^2,$$

for any  $\phi \in H^{-1/2}(\Gamma_a)$ .

*Proof.* By Proposition 4.2, the proof follows from arguments similar as in [18, Sec. 2.3]. It relies on the Fredholmness of the single-layer potential and the compactness of the double-layer potential of kernels corresponding to the uncomplexified Green function G, as has been studied in [30, Thm. 7.6] and [18]. Note that the strongly singular contribution in the kernel function of the double-layer potential corresponding to the complexified Green function  $\widetilde{G}$ , the fundamental solution of (11), vanishes in the kernel  $\partial_{\nu_c} \widetilde{G}_D$ . We omit the details.  $\Box$ 

Analogous to [23, Lem. 5.1], one gets the Green's representation

$$\tilde{u}(x) = \int_{\Gamma_a} \left[ \partial_{\nu_c(y)} \widetilde{G}_{\mathcal{D}}(x; y) \widetilde{u}(y) - \widetilde{G}_{\mathcal{D}}(x; y) \partial_{\nu_c(y)} \widetilde{u}(y) \right] \mathrm{d}s(y).$$
<sup>(26)</sup>

By the jump relations [23, Thm. 5.1], letting x approach  $\Gamma_a$ , we get the transparent boundary condition (TBC)

$$\tilde{u}|_{\Gamma_a} - \mathcal{K}_a(\tilde{u}|_{\Gamma_a}) = -\mathcal{S}_a(\partial_{\nu_c}\tilde{u}|_{\Gamma_a}).$$
<sup>(27)</sup>

As indicated in Figure 2 (b), let  $\Omega_{0a}$  be the domain bounded by  $\Gamma_0^+$ ,  $\Gamma_a$ ,  $\Gamma_{H+L}^+$  and  $\Gamma_+$ ,

$$H_D^1(\Omega_{0a}) := \left\{ v|_{H^1(\Omega_{0a})} : v \in H^1(\Omega_{\text{PML}}^+), v|_{\Gamma_+} = 0, v|_{\Gamma_{H+L}} = 0 \right\},\$$

and  $V_a := H_D^1(\Omega_{0a}) \times H^{-1/2}(\Gamma_a)$  be equipped with the natural cross-product norm. Problem  $(P^+)$  can be equivalently formulated as the following boundary value problem: Find  $(\tilde{u}, \phi) \in V_a$  satisfying

$$\nabla \cdot (\mathbf{A}\nabla \tilde{u}) + k^2 \alpha \tilde{u} = 0, \quad \text{on} \quad \Omega_{0a},$$
  
$$\partial_{\nu_c} \tilde{u}|_{\Gamma_0^+} = g^+, \quad \text{on} \quad \Gamma_0^+,$$
  
$$\partial_{\nu_c} \tilde{u}|_{\Gamma_a} = \phi, \quad \text{on} \quad \Gamma_a,$$
  
$$\tilde{u}|_{\Gamma_a} - \mathcal{K}_a(\tilde{u}|_{\Gamma_a}) = -\mathcal{S}_a \phi, \quad \text{on} \quad \Gamma_a.$$

An equivalent variational formulation reads: Find  $(\tilde{u}, \phi) \in V_a$  such that

$$b_{ps}\Big((\tilde{u},\phi),(v,\psi)\Big) = \int_{\Gamma_0^+} g^+ \bar{v} \,\mathrm{d}s,\tag{28}$$

for all  $(v, \psi) \in V_a$ , where the sesquilinear form  $b_{ps}(\cdot, \cdot) : V_a \times V_a \to \mathbb{C}$  is given by

$$b_{ps}\Big((\tilde{u},\phi),(v,\psi)\Big) := \int_{\Omega_{0a}} [(\mathbf{A}\nabla\tilde{u})^T \overline{\nabla v} - k^2 \alpha \tilde{u} \bar{v}] \,\mathrm{d}x - \int_{\Gamma_a} [\phi \bar{v} - (\tilde{u} - \mathcal{K}_a \tilde{u} + \mathcal{S}_a \phi) \,\bar{\psi}] \,\mathrm{d}s.$$

We are now ready to establish the well-posedness of problems  $(P^+)$ .

**Theorem 4.4** Under the geometrical conditions (GC1) and (GC2), provided that L and  $S_cL$  are sufficiently large, the semi-waveguide problem  $(P^{\pm})$  has a unique solution  $\tilde{u} \in H^1(\Omega_{\text{PML}}^{\pm})$ . There holds  $||\tilde{u}||_{H^1(\Omega_{\text{PML}}^{\pm})} \leq C||g^{\pm}||_{H^{-1/2}(\Gamma_0^{\pm})}$  for any  $g^{\pm} \in H^{-1/2}(\Gamma_0^{\pm})$ , where C is independent of  $g^{\pm}$ .

*Proof.* We study  $(P^+)$  only. For the variational problem (28), we can decompose  $b_{ps} = b_1 + b_2$  where

$$b_1\Big((\tilde{u},\phi),(v,\psi)\Big) := \int_{\Omega_{0a}} \left[ (\mathbf{A}\nabla\tilde{u})^T \overline{\nabla v} + \tilde{u}\bar{v} \right] \mathrm{d}x - \int_{\Gamma_a} \left[ \phi\bar{v} - \tilde{u}\bar{\psi} - \mathcal{S}_{p,a}\phi\bar{\psi} \right] \mathrm{d}s$$
$$b_2\Big((\tilde{u},\phi),(v,\psi)\Big) := -\int_{\Omega_{0a}} \left[ 1 + k^2\alpha \right] \tilde{u}\bar{v} \,\mathrm{d}x + \int_{\Gamma_a} \left[ \mathcal{L}_{p,a}\phi - \mathcal{K}_a\tilde{u} \right] \bar{\psi} \,\mathrm{d}s.$$

According to Lemma 4.3,  $b_1$  is coercive on V as

$$\operatorname{Re} b_1 \Big( (\tilde{u}, \phi), (\tilde{u}, \phi) \Big) = \int_{\Omega_{0a}} \Big[ |\tilde{u}_{x_1}|^2 + (1 + \sigma^2(x_2))^{-1} |\tilde{u}_{x_2}|^2 + |\tilde{u}|^2 \Big] \, \mathrm{d}x + \operatorname{Re} \left( \int_{\Gamma_a} \mathcal{S}_{p,a} \phi \bar{\phi} \mathrm{d}s \right) \\ \geq c ||\tilde{u}||_{H^1(\Omega_{0a})}^2 + c ||\phi||_{H^{-1/2}(\Gamma_a)}^2,$$

and the bounded linear operator associated with  $b_2$  is compact. Consequently,  $b_{ps}$  is Fredholm of index zero [30, Thm. 2.34].

Now, we prove  $\tilde{u} = 0$  and  $\phi = 0$  when  $g^+ = 0$ . By (26), we can directly define

$$\tilde{u}^{\text{ext}}(x) := \int_{\Gamma_a} \left[ \partial_{\nu_c(y)} \widetilde{G}_{\mathcal{D}}(x; y) \widetilde{u}(y) - \widetilde{G}_{\mathcal{D}}(x; y) \partial_{\nu_c(y)} \widetilde{u}(y) \right] \mathrm{d}s(y), \ x \in \Omega_{\epsilon}^+ \backslash \overline{\Gamma}_a.$$

Then, the TBC (27) implies  $\gamma^+(\tilde{u}^{\text{ext}}|_{\Omega_a^+}) = \tilde{u}|_{\Gamma_a}$  so that, by the jump relations,  $\gamma^-(\tilde{u}^{\text{ext}}|_{\Omega_{\epsilon}}) = 0$ , where  $\gamma^+$  and  $\gamma^-$  are the trace operators of  $\tilde{u}^{\text{ext}}$  onto  $\Gamma_a$  from  $\Omega_a^+$  and  $\Omega_{\epsilon}$ , respectively. Using that

the Green function  $\tilde{G}_D$  vanishes on  $\partial \Omega_{\epsilon}^+$ , function  $\tilde{u}^{\text{ext}}|_{\Omega_{a}^+}$  is zero on  $\partial \Omega_{\epsilon}^+$  too. Thus,  $\tilde{u}^- = \tilde{u}^{\text{ext}}|_{\Omega_{\epsilon}}$  satisfies (24) and (25). However, the special choice of  $\epsilon$  and  $\Omega_{\epsilon}$  ensures  $\tilde{u}^- \equiv 0$  on  $\Omega_{\epsilon}$  so that the trace of  $\partial_{\nu_c} \tilde{u}^{\text{ext}}$  taken from  $\Omega_{\epsilon}$  is 0. The jump conditions then imply that the trace of  $\partial_{\nu_c} \tilde{u}^{\text{ext}}$  taken from  $\Omega_{a}^+$  is  $\phi$ . Consequently,

$$w(x) := \begin{cases} \tilde{u}(x), & x \in \Omega_{0a}, \\ \tilde{u}^{\text{ext}}(x), & x \in \Omega_a^+, \end{cases}$$

belongs to  $H^1(\Omega^+_{\rm PML})$  and satisfies  $(P^+)$  with  $g^+ = 0$ . However, Lemma 4.1 already justifies that w must be 0 on  $\Omega^+_{\rm PML}$ , which indicates that  $\tilde{u} = 0$  and  $\phi = 0$ .

Finally, the assertion of the theorem follows from Fredholm's alternative and the fact that the right-hand side of (28) defines a bounded anti-linear functional in  $V_a^*$ .  $\Box$ 

**Remark 4.5** Like Theorem 3.2, Theorem 4.4 also holds for any Lipschitz curve  $\Gamma^{\pm}$ , which are not necessarily periodic, satisfying the geometrical conditions (GC1) and (GC2).

### 5 Lateral boundary conditions

According to Theorem 3.2,  $\partial_{\nu_c} \tilde{u}^{\mathrm{og}}(\cdot; x^*)|_{\Gamma_0}^{\pm} \in H^{-1/2}(\Gamma_0^{\pm})$  for any  $x^* \in S_H$  with  $|x_1^*| < T/2$ . Thus,  $\tilde{u} = \tilde{u}^{\mathrm{og}}(\cdot; x^*)|_{\Omega_{\mathrm{PML}}^{\pm}}$  satisfies  $(P^{\pm})$  with  $g^{\pm} = \partial_{\nu_c} \tilde{u}^{\mathrm{og}}(\cdot; x^*)|_{\Gamma_0^{\pm}}$  in the distributional sense. By Theorem 4.4 we define vertical Neumann-to-Dirichlet (vNtD) operators  $\mathcal{N}^{\pm} : H^{-1/2}(\Gamma_0^{\pm}) \to \tilde{H}^{1/2}(\Gamma_0^{\pm})$  satisfying  $\tilde{u}^{\mathrm{og}}|_{\Gamma_0^{\pm}} = \mathcal{N}^{\pm}(\partial_{\nu_c} \tilde{u}^{\mathrm{og}}|_{\Gamma_0^{\pm}})$ . Such transparent boundary conditions can serve as exact *lateral boundary conditions* to terminate the  $x_1$ -variable for the PML-truncated problem (11) and (12). Consequently, the original problem (1) and (2) over an unbounded domain and equipped with the hSRC condition (7) can be truncated onto the perturbed cell  $\Omega_0 := \Omega_{\mathrm{PML}} \cap \left\{x : |x_1| < \frac{T}{2}\right\}$  and be reformulated as the following boundary value problem:

$$(\text{BVP1}): \begin{cases} \nabla \cdot (\mathbf{A} \nabla \tilde{u}^{\text{og}}) + k^2 \alpha \tilde{u}^{\text{og}} = -\delta(x - x^*), & \text{on } \Omega_0, \\ \tilde{u}^{\text{og}} = 0, & \text{on } \Gamma_0 = \Gamma \quad \cap \{x \colon |x_1| < T/2\}, \\ \tilde{u}^{\text{og}} = 0, & \text{on } \Gamma_{H+L}^0 = \Gamma_{H+L} \cap \{x \colon |x_1| < T/2\}, \\ \tilde{u}^{\text{og}} = \mathcal{N}^{\pm}(\partial_{\nu_c} \tilde{u}^{\text{og}}), & \text{on } \Gamma_0^{\pm}. \end{cases}$$

Theorems 3.2 and 4.4 directly imply that (BVP1) admits the following unique solution

$$\tilde{u}^{\mathrm{og}}(\cdot;x^*) = \tilde{u}_r^{\mathrm{og}}(\cdot;x^*)|_{\Omega_0} + \chi(\cdot;x^*)|_{\Omega_0}\tilde{u}^{\mathrm{inc}}(x;x^*)|_{\Omega_0},$$

with  $\tilde{u}_r^{\text{og}}$  defined in Theorem 3.2. Nevertheless, it is challenging to get  $\mathcal{N}^{\pm}$  by directly solving the unbounded problem  $(P^{\pm})$  in practice. To overcome this difficulty, in this section, we shall define two closely related Neumann-marching operators, derive the governing Riccati equations, and design an efficient RDP to accurately approximate  $\mathcal{N}^{\pm}$ .

### 5.1 Neumann-marching operators $\mathcal{R}_p^{\pm}$

Now let

$$\begin{split} \Gamma_j^{\pm} &:= \left\{ (x_1 \pm jT, x_2) \colon x = (x_1, x_2) \in \Gamma_0^{\pm} \right\}, \\ \Omega_{\mathrm{PML}, j}^{\pm} &:= \left\{ x \in \Omega_{\mathrm{PML}}^{\pm} \colon \pm x_1 > \frac{T}{2} + (j-1)T \right\}, \\ \Omega_j^{\pm} &:= \Omega_{\mathrm{PML}, j}^{\pm} \backslash \overline{\Omega}_{\mathrm{PML}, j+1}^{\pm} \end{split}$$

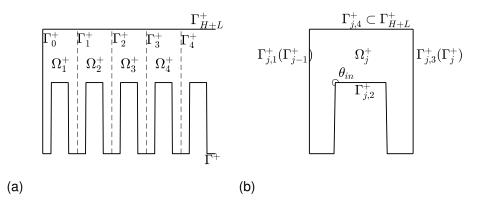


Figure 3: (a) The semi-waveguide region  $\Omega_{PML}^+$  is divided into domains  $\Omega_j^+$ ,  $j = 1, \cdots$  of the same shape. The operator  $\mathcal{R}_p^+$  can then march Neumann data through the vertical line segments  $\Gamma_j^+$ ,  $j = 0, \cdots$ . (b) The boundary of  $\Omega_j$  consists of four parts:  $\Gamma_{j,1}^+$  (left),  $\Gamma_{j,2}^+$  (bottom),  $\Gamma_{j,3}^+$  (right),  $\Gamma_{j,4}^+$  (top). Here,  $\theta^{\text{in}}$  indicates the interior angle at a corner, as will be used in (45).

for  $j \in \mathbb{N}^*$ , as illustrated in Figure 3(a) for the notation with superscript +.

As inspired by [21], the well-posedness of  $(P^{\pm})$  well defines two bounded Neumann-marching operators  $\mathcal{R}_p^{\pm} : H^{-1/2}(\Gamma_0^{\pm}) \to H^{-1/2}(\Gamma_1^{\pm})$  such that  $\partial_{\nu_c^{\pm}} \tilde{u}^{\mathrm{og}}|_{\Gamma_1^{\pm}} = \mathcal{R}_p^{\pm}(\partial_{\nu_c^{\pm}} \tilde{u}^{\mathrm{og}}|_{\Gamma_0^{\pm}})$ , where  $\nu_c^{\pm} := \mathbf{A}\nu^{\pm}$  with  $\nu^{\pm} := (\pm 1, 0)^T$ . We have the following properties of  $\mathcal{R}_p^{\pm}$ , analogous to [21, Thm. 3.1].

**Proposition 5.1** Under the conditions that (GC2) holds and  $kS_cL$  as well as kL are sufficiently large, we can choose  $\Gamma_0^{\pm}$  intersecting  $\Gamma$  at a smooth point such that  $\mathcal{R}_p^{\pm}$  are compact operators and

$$\partial_{\nu_c^{\pm}} \tilde{u}^{\text{og}}|_{\Gamma_{j+1}^{\pm}} = \mathcal{R}_p^{\pm}(\partial_{\nu_c^{\pm}} \tilde{u}^{\text{og}}|_{\Gamma_j^{\pm}})$$
<sup>(29)</sup>

holds for any  $j \ge 0$ . Furthermore,

$$\rho(\mathcal{R}_p^{\pm}) < 1, \tag{30}$$

where  $\rho$  denotes the spectral radius.

*Proof.* We study only the property of  $\mathcal{R}_p^+$ . The choice of  $\Gamma_0^+$  and the interior regularity theory of elliptic operators directly imply the compactness of  $\mathcal{R}_p^+$ .

It is clear that (29) holds for j = 0. We need only justify the case j = 1 as all others can be done by induction. Consider the semi-waveguide problem  $(P^+)$  with  $g^+ = -\partial_{\nu_c^+} \tilde{u}^{\mathrm{og}}|_{\Gamma_1^+}$ , where the negative sign appears since  $\nu_c^+ = -\nu_c$ . Theorem 4.4 implies that  $\tilde{u}_n^{\mathrm{og}}(x) = \tilde{u}^{\mathrm{og}}(x_1 + T, x_2)$  for  $x \in \Omega_{\mathrm{PML}}^+$  is the unique solution. Then  $\partial_{\nu_c^+} \tilde{u}_n^{\mathrm{og}}|_{\Gamma_1^+} = \mathcal{R}_p^+(\partial_{\nu_c^+} \tilde{u}_n^{\mathrm{og}}|_{\Gamma_0^+})$ , which reads exactly as  $\partial_{\nu_c^+} \tilde{u}^{\mathrm{og}}|_{\Gamma_2^+} = \mathcal{R}_p^+(\partial_{\nu_c^+} \tilde{u}^{\mathrm{og}}|_{\Gamma_1^+})$ .

To prove (30) we take an arbitrary eigenfunction  $0 \neq g \in H^{-1/2}(\Gamma_0^+)$  such that  $\mathcal{R}_p^+ g = \lambda_0 g$ . Suppose the function  $\tilde{u}$  satisfies  $(P^+)$  with  $g^+ = g$  on  $\Gamma_0^+$ . Then, for any  $v \in H^1(\Omega_{\text{PML}}^+)$ , we arrive at  $v(\cdot - jT, \cdot) \in H^1(\Omega_{\text{PML}, j+1}^+)$  for any  $j \geq 0$ , and, by Green's identity, that

$$\begin{aligned} |\lambda_{0}|^{j} \left| \int_{\Gamma_{0}} g\bar{v} \,\mathrm{d}s \right| &= \left| \int_{\Gamma_{j}} (\mathcal{R}_{p}^{+})^{j} g\overline{v(\cdot - jT, \cdot)} \,\mathrm{d}s \right| \\ &= \left| \int_{\Omega_{\mathrm{PML}, j+1}} \left[ (\mathbf{A}\nabla \tilde{u}^{\mathrm{og}})^{T} \overline{\nabla v(\cdot - jT, \cdot)} - k^{2} \alpha \tilde{u}^{\mathrm{og}} \overline{v(\cdot - jT, \cdot)} \right] \mathrm{d}x \\ &\leq C ||\tilde{u}^{\mathrm{og}}||_{H^{1}(\Omega_{\mathrm{PML}, j+1}^{+})} ||v||_{H^{1}(\Omega_{\mathrm{PML}}^{+})} \to 0, \quad j \to \infty. \end{aligned}$$

Choosing v such that  $\int_{\Gamma_0} g\bar{v} \, ds \neq 0$ , we get  $|\lambda_0| < 1$  and (30) follows.  $\Box$ 

By the identity  $\rho(\mathcal{R}_p^{\pm}) = \lim_{j \to \infty} ||(\mathcal{R}_p^{\pm})^j||^{1/j}$  (cf. e.g. [22]), there exists a sufficiently large integer  $N_0 > 0$  such that  $(\mathcal{R}_p^{\pm})^{N_0}$  is contracting, i.e.,  $||(\mathcal{R}_p^{\pm})^{N_0}|| < 1$ . Let  $\Omega_j^{\pm,N_0}$  be the interior of  $N_0$  consecutive cells  $\cup_{j'=1}^{N_0} \overline{\Omega}_{(j-1)N_0+j'}^{\pm}$ . From the corresponding recursion (29) with  $\Omega_j^{\pm}$  and  $\mathcal{R}_p^{\pm}$  replaced by  $\Omega_j^{\pm,N_0}$  and  $(\mathcal{R}_p^{\pm})^{N_0}$ , respectively, we get that  $\tilde{u}^{\text{og}}$  decays exponentially at infinity of the strip.

**Corollary 5.2** If (GC2) holds and kL as well as  $kS_cL$  are sufficiently large, then

$$||\tilde{u}^{\mathrm{og}}(\cdot; x^*)||_{H^1(\Omega_j^{\pm, N_0})} \le C||(\mathcal{R}_p^{\pm})^{N_0}||^{j-1}||\tilde{g}^{\mathrm{inc}}||_{L^2(\Omega_{\mathrm{PML}})},\tag{31}$$

where we recall  $\tilde{g}^{\text{inc}} := [\nabla \cdot (\mathbf{A}\nabla) + k^2 \alpha](1 - \chi(x; x^*))\tilde{u}^{\text{inc}}(x; x^*)$  and where the constant C is independent of  $j \ge 0$ . In other words, for any  $x^* \in \Omega_{\text{PML}}$ , the PML truncated solution  $\tilde{u}^{\text{og}}(x; x^*)$  decays exponentially fast to 0 in the strip as  $|x_1| \to \infty$ .

**Remark 5.3** The authors in [6] have revealed a similar result as (31) for  $\Gamma$  being a flat surface. The above corollary indicates that such an exponentially decaying property for the PML truncated solution holds even for locally defected periodic curves. As a consequence, this reveals that the PML truncation cannot realize an exponential convergence to the true solution for numerical solutions at regions sufficiently away from the source or local defects since the true solution is expected to decay only of an algebraic rate at infinity: [7] has indicated that  $u^{\text{og}}$  behaves as  $\mathcal{O}(x_1^{-3/2})$  as  $x_1 \to \infty$ .

Though Corollary 5.2 provides disappointing results, we point out that (31) holds for L being fixed but  $j \to \infty$ . If, on the contrary, j is fixed but  $L \to \infty$ , we believe exponential convergence can still be achieved. Applying the Neumann-marching operators  $\mathcal{R}_p^{\pm}$  in the numerical algorithm, we need a description of  $\mathcal{R}_p^{\pm}$  more efficient than that in [21]. Take  $\mathcal{R}_p^+$  as an example. Recall that  $\Omega_j^+$  denotes the j-th unit cell on the right of  $\Gamma_0^+$  (cf. Figure 3(b)), which is unperturbed for  $j \ge 1$ . To simplify the presentation, we further denote the four boundaries of  $\Omega_j^+$  by  $\Gamma_{j,1} = \Gamma_{j-1}^+, \Gamma_{j,3} = \Gamma_j^+, \Gamma_{j,2} = \overline{\Omega}_j^+ \cap \Gamma$ , and  $\Gamma_{j,4} = \overline{\Omega}_j^+ \cap \Gamma_{H+L}^+$ . Consider the following boundary value problem for a generic field  $\tilde{u}$ :

$$(\text{BVP2}): \begin{cases} \nabla \cdot (\mathbf{A}\nabla \tilde{u}) + k^2 \alpha \tilde{u} = 0, & \text{on } \Omega_j^+, \\ \tilde{u} = 0, & \text{on } \Gamma_{j,2} \cup \Gamma_{j,4}, \\ \partial_{\nu_c} \tilde{u} = g_i, & \text{on } \Gamma_i^+, i = j - 1, j, \end{cases}$$

for  $g_i \in H^{-1/2}(\Gamma_i^+), i = j - 1, j$ . We have the following well-posedness theorem.

**Theorem 5.4** Provided that  $kT/\pi \notin \mathcal{E} := \{i'/2^{j'} : j' \in \mathbb{N}, i' \in \mathbb{N}^*\}$  and that L as well as  $S_cL$  are sufficiently large, (BVP2) is well-posed. The well-posedness even holds with  $\Omega_j^+$  replaced by the interior domain of  $2^l$  consecutive cells, say  $\bigcup_{i=1}^{2^l} \overline{\Omega}_i^+$ , for any number  $l \ge 0$ .

Proof. It is clear that only uniqueness is needed [30, Thm. 4.10]. Suppose j = 1 and  $g_i = 0, i = 0, 1$ . Then, by first an even extension over  $\Gamma_0^+$  and then a 2T-periodic extension, we get a 2T-periodic solution  $\tilde{u}^e$  (corresponding to normal incidence) in a strip bounded in the  $x_2$ -direction by  $\Gamma_{H+L}$  and a 2T-periodic grating surface, possibly different from  $\Gamma$ . However, according to the well-posedness theory [5, Cor. 5.2] for the half-space scattering by the grating, the PML convergence theory in [10, Thm. 2.4] can be readily adapted here to show that  $\tilde{u}^e \equiv 0$ , taking into account that  $kT/\pi \notin \mathcal{E}$  has excluded horizontally propagating Bloch modes.  $\Box$  **Remark 5.5** We note that the condition  $kT/\pi \notin \mathcal{E}$  is not necessary for the well-posedness of (BVP2). Alternatively, if  $kT/\pi \in \mathcal{E}$ , one may impose zero Neumann condition on  $\Gamma_{4,j}$  to guarantee the uniqueness of the modified (BVP2) (cf. [29, 34]).

Taking into account the just proved Theorem 5.4, we can introduce the bounded Neumann-to-Dirichlet operator  $\mathcal{N}^{(0)} \colon H^{-1/2}(\Gamma_{j-1}^+) \times H^{-1/2}(\Gamma_j^+) \to \widetilde{H}^{1/2}(\Gamma_{j-1}^+) \times \widetilde{H}^{1/2}(\Gamma_j^+)$  such that

$$\begin{bmatrix} \tilde{u}|_{\Gamma_{j-1}^{+}} \\ \tilde{u}|_{\Gamma_{j}^{+}} \end{bmatrix} = \mathcal{N}^{(0)} \begin{bmatrix} \partial_{\nu_{c}^{-}} \tilde{u}|_{\Gamma_{j-1}^{+}} \\ \partial_{\nu_{c}^{+}} \tilde{u}|_{\Gamma_{j}^{+}} \end{bmatrix},$$
(32)

for all  $j \ge 1$ . Due to the shape invariance of  $\Omega_j^+$  w.r.t. j, operator  $\mathcal{N}^{(0)}$  is in fact independent of j. Suppose j = 1. Then, by the linearity principle, operator  $\mathcal{N}^{(0)}$  can be rewritten in the matrix form

$$\mathcal{N}^{(0)} = \left[ \begin{array}{cc} \mathcal{N}_{00}^{(0)} & \mathcal{N}_{01}^{(0)} \\ \mathcal{N}_{10}^{(0)} & \mathcal{N}_{11}^{(0)} \end{array} \right]$$

where the bounded entry  $\mathcal{N}_{i'j'}^{(0)} \colon H^{-1/2}(\Gamma_{j'}^+) \to \widetilde{H}^{1/2}(\Gamma_{i'})$  maps  $\partial_{\nu_c} \tilde{u}|_{\Gamma_{j'}^+} = g_{j'}$  to  $\tilde{u}|_{\Gamma_{i'}^+}$  if  $g_{1-j'} = 0$  for i', j' = 0, 1 in (BVP2). Due to the shape invariance of  $\Gamma_j^+$ , we shall identify  $H^{-1/2}(\Gamma_j^+)$  for all  $j \ge 0$  with the space  $H^{-1/2}(\Gamma_0^+)$ , and similarly  $\widetilde{H}^{1/2}(\Gamma_j^+)$  with the dual space of  $H^{-1/2}(\Gamma_0^+)$ .

Returning back to the semi-waveguide problems ( $P^{\pm}$ ), we have, by the definition of  $\mathcal{R}_p^+$  and (32) for j=1 and 2, that

$$\mathcal{N}_{10}^{(0)}\partial_{\nu_{c}^{-}}\tilde{u}^{\mathrm{og}}|_{\Gamma_{0}^{+}} - \mathcal{N}_{11}^{(0)}\mathcal{R}_{p}^{+}\partial_{\nu_{c}^{-}}\tilde{u}^{\mathrm{og}}|_{\Gamma_{0}^{+}} = \tilde{u}^{\mathrm{og}}|_{\Gamma_{1}^{+}} = \mathcal{N}_{00}^{(0)}\mathcal{R}_{p}^{+}\partial_{\nu_{c}^{-}}\tilde{u}^{\mathrm{og}}|_{\Gamma_{0}^{+}} - \mathcal{N}_{01}^{(0)}(\mathcal{R}_{p}^{+})^{2}\partial_{\nu_{c}^{-}}\tilde{u}^{\mathrm{og}}|_{\Gamma_{0}^{+}} = \mathcal{N}_{00}^{(0)}\mathcal{R}_{p}^{+}\partial_{\nu_{c}^{-}}\tilde{u}^{\mathrm{og}}|_{\Gamma_{0}^{+}} - \mathcal{N}_{01}^{(0)}(\mathcal{R}_{p}^{+})^{2}\partial_{\nu_{c}^{-}}\tilde{u}^{\mathrm{og}}|_{\Gamma_{0}^{+}} = \mathcal{N}_{00}^{(0)}\mathcal{R}_{p}^{+}\partial_{\nu_{c}^{-}}\tilde{u}^{\mathrm{og}}|_{\Gamma_{0}^{+}} - \mathcal{N}_{01}^{(0)}(\mathcal{R}_{p}^{+})^{2}\partial_{\nu_{c}^{-}}\tilde{u}^{\mathrm{og}}|_{\Gamma_{0}^{+}} = \mathcal{N}_{00}^{(0)}\mathcal{R}_{p}^{+}\partial_{\nu_{c}^{-}}\tilde{u}^{\mathrm{og}}|_{\Gamma_{0}^{+}} - \mathcal{N}_{01}^{(0)}\mathcal{R}_{p}^{+}\partial_{\nu_{c}^{-}}\tilde{u}^{\mathrm{og}}|_{\Gamma_{0}^{+}} - \mathcal{N}_{01}^{(0)}\mathcal{R}_{p}^{+} - \mathcal{N}_{01}^{(0)}\mathcal{R}_{p}^{+} - \mathcal{N}_{01}^{(0)}\mathcal{R}_{p}^{+} - \mathcal{N}_{01}^{(0)}\mathcal{R}_{p}^{+} - \mathcal{N}_{01}^{(0)}\mathcal{R}_{p}^{+} - \mathcal{N}_{01}^{(0)}\mathcal{R}_{p}^$$

Here and in the following, the product of two operators should be regarded as their composition. Thus,

$$\left[\mathcal{N}_{10}^{(0)} + \mathcal{N}_{11}^{(0)}\mathcal{R}_p^+ + \mathcal{N}_{00}^{(0)}\mathcal{R}_p^+ + \mathcal{N}_{01}^{(0)}(\mathcal{R}_p^+)^2\right] \left(\partial_{\nu_c} \tilde{u}^{\mathrm{og}}|_{\Gamma_0^+}\right) = 0,$$

for any  $\partial_{\nu_c} \tilde{u}^{\mathrm{og}}|_{\Gamma_0^+} \in H^{-1/2}(\Gamma_0^+)$ , so that we end up with the following Riccati equation for  $\mathcal{R}_p^+$ :

$$\mathcal{N}_{10}^{(0)} + \left[\mathcal{N}_{11}^{(0)} + \mathcal{N}_{00}^{(0)}\right] \mathcal{R}_p^+ + \mathcal{N}_{01}^{(0)} (\mathcal{R}_p^+)^2 = 0.$$
(33)

One similarly obtains the governing equation for  $\mathcal{R}_p^-$ :

$$\mathcal{N}_{01}^{(0)} + \left[\mathcal{N}_{11}^{(0)} + \mathcal{N}_{00}^{(0)}\right] \mathcal{R}_p^- + \mathcal{N}_{10}^{(0)} (\mathcal{R}_p^-)^2 = 0.$$
(34)

Analogous to [21], the previous results in fact indicate that the two Riccati equations (33) and (34) must be uniquely solvable under the condition that  $\rho(R_p^{\pm}) < 1$ . The vNtD operators  $\mathcal{N}^{\pm}$  mapping  $\partial_{\nu_c} \tilde{u}^{\mathrm{og}}|_{\Gamma_0^{\pm}}$  to  $\tilde{u}^{\mathrm{og}}|_{\Gamma_0^{\pm}}$  are respectively given by

$$\mathcal{N}^{+} = \mathcal{N}_{00}^{(0)} - \mathcal{N}_{01}^{(0)} \mathcal{R}_{p}^{+}, \quad \mathcal{N}^{-} = \mathcal{N}_{11}^{(0)} - \mathcal{N}_{10}^{(0)} \mathcal{R}_{p}^{-}$$

However, due to the nonlinearity of the Riccati equations (33) and (34), it is not that easy to get  $\mathcal{N}^{\pm}$  in practice [21]. To tackle this difficulty, we shall develop an RDP to effectively approximate  $\mathcal{R}_n^{\pm}$ .

#### 5.2 Recursive doubling procedure

Take  $\mathcal{R}_p^+$  as an example. We first study the NtD operator

$$\mathcal{N}^{(l)} = \left[ egin{array}{cc} \mathcal{N}^{(l)}_{00} & \mathcal{N}^{(l)}_{01} \ \mathcal{N}^{(l)}_{10} & \mathcal{N}^{(l)}_{11} \end{array} 
ight]$$

on the boundary of  $\bigcup_{j=1}^{2^l} \overline{\Omega}_j^+$  for  $l \ge 1$ , where  $\mathcal{N}_{i'j'}^{(l)}$  is bounded from  $H^{-1/2}(\Gamma_0^+)$  to  $\widetilde{H}^{1/2}(\Gamma_0)$  for i', j' = 0, 1. If l = 1, we need to compute  $\mathcal{N}^{(1)}$  on the boundary of  $\overline{\Omega}_1^+ \cup \overline{\Omega}_2^+$ . Using (32) for j = 1 and 2 and eliminating  $\widetilde{u}^{\text{og}}$  and  $\partial_{\nu_c} \widetilde{u}^{\text{og}}$  by the continuity condition on  $\Gamma_1^+$ , one gets

$$\left(\mathcal{N}_{00}^{(l-1)} + \mathcal{N}_{11}^{(l-1)}\right) \left(\partial_{\nu_c^+} \tilde{u}^{\mathrm{og}}|_{\Gamma_1^+}\right) = -\mathcal{N}_{10}^{(l-1)} \left(\partial_{\nu_c^-} \tilde{u}^{\mathrm{og}}|_{\Gamma_0^+}\right) + \mathcal{N}_{01}^{(l-1)} \left(\partial_{\nu_c^+} \tilde{u}^{\mathrm{og}}|_{\Gamma_2^+}\right). \tag{35}$$

By Theorem 5.4, the well-posedness of the modified (BVP2) for l = 1, indicates that there exist two bounded operators  $\mathcal{A}_{l-1}, \mathcal{B}_{l-1}: H^{-1/2}(\Gamma_0^+) \to H^{-1/2}(\Gamma_0^+)$  such that

$$\partial_{\nu_c^+} \tilde{u}^{\mathrm{og}}|_{\Gamma_1^+} = -\mathcal{A}_{l-1}(\partial_{\nu_c^-} \tilde{u}^{\mathrm{og}}|_{\Gamma_0^+}) + \mathcal{B}_{l-1}(\partial_{\nu_c^+} \tilde{u}^{\mathrm{og}}|_{\Gamma_2^+}).$$

Equation (35) implies that

$$\mathcal{A}_{l-1} = \left(\mathcal{N}_{00}^{(l-1)} + \mathcal{N}_{11}^{(l-1)}\right)^{-1} \mathcal{N}_{10}^{(l-1)}, \quad \mathcal{B}_{l-1} = \left(\mathcal{N}_{00}^{(l-1)} + \mathcal{N}_{11}^{(l-1)}\right)^{-1} \mathcal{N}_{01}^{(l-1)}$$

where  $(\mathcal{N}_{00}^{(l-1)} + \mathcal{N}_{11}^{(l-1)})^{-1}$  is a generalized inverse from  $\widetilde{H}^{1/2}(\Gamma_0)$  to  $H^{-1/2}(\Gamma_0^+)$ . Thus, one obtains

$$\mathcal{N}_{00}^{(l)} = \mathcal{N}_{00}^{(l-1)} - \mathcal{N}_{01}^{(l-1)} \mathcal{A}_{l-1}, \quad \mathcal{N}_{01}^{(l)} = \mathcal{N}_{01}^{(l-1)} \mathcal{B}_{l-1}, \tag{36}$$

$$\mathcal{N}_{10}^{(l)} = \mathcal{N}_{10}^{(l-1)} \mathcal{A}_{l-1}, \qquad \qquad \mathcal{N}_{11}^{(l)} = \mathcal{N}_{11}^{(l-1)} - \mathcal{N}_{10}^{(l-1)} \mathcal{B}_{l-1}.$$
(37)

Equations (36-37) can be recursively applied to get  $\mathcal{N}^{(l)}$  for all  $l \geq 1$ , and the number of consecutive cells  $\{\Omega_j\}$  doubles after each iteration, which form the origin of the term "*recursive doubling procedure*" (RPD) in the literature (cf. [33, 15]). In the following, we shall see that RDP provides a simple approach for solving (33) and (34).

Now, analogously to (33) and (34), we obtain from  $\mathcal{N}^{(l)}$  and (29) the following equations

$$\mathcal{N}_{10}^{(l)} + \left[ \mathcal{N}_{11}^{(l)} + \mathcal{N}_{00}^{(l)} \right] (\mathcal{R}_p^+)^{2^l} + \mathcal{N}_{01}^{(l)} (\mathcal{R}_p^+)^{2^{(l+1)}} = 0, \qquad (38)$$
$$\mathcal{N}^+ = \mathcal{N}_{00}^{(l)} - \mathcal{N}_{01}^{(l)} (\mathcal{R}_p^+)^{2^l}.$$

Since  $||(\mathcal{R}_p^+)^{N_0}|| < 1$ , the third term in (38) is expected to be exponentially small for  $l \gg \log_2 N_0$ , so that we approximate

$$(\mathcal{R}_{p}^{+})^{2^{l}} \approx - \left[ \mathcal{N}_{11}^{(l)} + \mathcal{N}_{00}^{(l)} \right]^{-1} \mathcal{N}_{10}^{(l)},$$
  
 
$$\mathcal{N}^{+} \approx \mathcal{N}_{00}^{(l)} + \mathcal{N}_{01}^{(l)} \left[ \mathcal{N}_{11}^{(l)} + \mathcal{N}_{00}^{(l)} \right]^{-1} \mathcal{N}_{10}^{(l)},$$
 (39)

and we get  $\mathcal{R}_p^+$  iteratively from

$$(\mathcal{R}_p^+)^{2^j} = -\left[\mathcal{N}_{11}^{(j)} + \mathcal{N}_{00}^{(j)}\right]^{-1} \left[\mathcal{N}_{10}^{(j)} - \mathcal{N}_{01}^{(j)} (\mathcal{R}_p^+)^{2^{j+1}}\right], \quad j = l-1, \cdots, 0.$$
(40)

One similarly obtains  $\mathcal{N}^-$  and  $\mathcal{R}^-_p$  from

$$(\mathcal{R}_{p}^{-})^{2^{l}} \approx - \left[ \mathcal{N}_{11}^{(l)} + \mathcal{N}_{00}^{(l)} \right]^{-1} \mathcal{N}_{01}^{(l)},$$
  
 
$$\mathcal{N}^{-} \approx \mathcal{N}_{11}^{(l)} + \mathcal{N}_{01}^{(l)} \left[ \mathcal{N}_{11}^{(l)} + \mathcal{N}_{00}^{(l)} \right]^{-1} \mathcal{N}_{10}^{(l)},$$
 (41)

$$\left(\mathcal{R}_{p}^{-}\right)^{2^{j}} = -\left[\mathcal{N}_{11}^{(j)} + \mathcal{N}_{00}^{(j)}\right]^{-1} \left[\mathcal{N}_{01}^{(j)} - \mathcal{N}_{10}^{(j)}(\mathcal{R}_{p}^{-})^{2^{j+1}}\right], \quad j = l - 1, \cdots, 0.$$
(42)

From the above, it can be seen that the essential step to approximate  $\mathcal{N}^{\pm}$  is to get the NtD operator  $\mathcal{N}^{(0)}$  on the boundary of the unit cell  $\Omega_1^{\pm}$ . As no information of the field  $\tilde{u}^{\text{og}}$  in  $\Omega_1^{\pm}$  is required, it is clear that the BIE method is an optimal choice since it treats only the boundary of  $\Omega_1^{\pm}$ . Since PML is involved in domain  $\Omega_1^{\pm}$ , the high-accuracy PML-based BIE method developed in our previous work [28] straightforwardly provides an accurate approximation of  $\mathcal{N}^{(0)}$ , so as to effectively drive RDP to get  $\mathcal{N}^{\pm}$ . We shall present the details in the next section.

### 6 The PML-based BIE method

In this section, we shall first review the PML-based BIE method of [28] to approximate the NtD operator on the boundary of a unit cell, one with perturbation and one without, by an NtD matrix. Then, we shall use these NtD matrices to approximate the two vNtD operators  $\mathcal{N}^{\pm}$  on  $\Gamma_0^{\pm}$  and to solve (BVP1) finally. From now on, we shall assume that the scattering surface  $\Gamma$  is piecewise smooth and satisfies (GC1), but not necessarily (GC2). Though the previous well-posedness theory relies on (GC2), our numerical solver does not rely on such an assumption, and we believe (GC2) can be weakened such that at least piecewise smooth curves are admitted. We shall investigate this in a future work.

#### 6.1 Approximating $\mathcal{N}^{\pm}$

Without loss of generality, consider (BVP2) in an unperturbed cell, say  $\Omega_1^+$ , and we need to approximate  $\mathcal{N}^{(0)}$  first. According to [28], for any  $\tilde{u}$  satisfying

$$\nabla \cdot (\mathbf{A}\nabla \tilde{u}) + k^2 \alpha \tilde{u} = 0, \tag{43}$$

on  $\Omega_1^+$ , we have the Green's representation

$$\tilde{u}(x) = \int_{\partial\Omega_1^+} \left\{ \widetilde{G}(x,y)\partial_{\nu_c} \tilde{u}(y) - \partial_{\nu_c} \widetilde{G}(x,y)\tilde{u}(y) \right\} \mathrm{d}s(y), \tag{44}$$

for all  $x \in \Omega_1^+$ . We recall that  $\nu$  denotes the outer unit normal vector on  $\partial \Omega_1^+$  and  $\nu_c := A\nu$  the co-normal vector. Moreover, as x approaches  $\partial \Omega_1^+ = \bigcup_{j=1}^4 \overline{\Gamma}_{j,1}$ , the usual jump conditions imply  $\mathcal{K}[\tilde{u}](x) - \mathcal{K}_0[1](x)\tilde{u}(x) = \mathcal{S}[\partial_{\nu_c}\tilde{u}](x)$  (cf. [28]), where we use the following integral operators

$$\begin{split} \mathcal{S}[\phi](x) &:= 2 \int_{\partial \Omega_1^+} \widetilde{G}(x, y) \phi(y) \, \mathrm{d}s(y), \\ \mathcal{K}[\phi](x) &:= 2 \int_{\partial \Omega_1^+} \partial_{\nu_c} \widetilde{G}(x, y) \phi(y) \, \mathrm{d}s(y), \\ \mathcal{K}_0[\phi](x) &:= 2 \int_{\partial \Omega_1^+} \partial_{\nu_c} \widetilde{G}_0(x, y) \phi(y) \, \mathrm{d}s(y), \end{split}$$

where  $\widetilde{G}_0(x,y) := -\log \rho(\tilde{x},\tilde{y})/(2\pi)$  is the fundamental solution of the complexified Laplace equation  $\Delta \tilde{u}(x) = 0$ . Note that

$$\mathcal{K}_0[1](x) = -\frac{\theta^{\rm in}(x)}{\pi},\tag{45}$$

where  $\theta^{in}(x)$  is defined as the interior angle at x (cf. Figure 3(b)). However, numerically evaluating  $\mathcal{K}_0[1]$  near corners is more advantageous as has been illustrated in the literature (cf. [13, 27]). Thus,  $\tilde{u} = (\mathcal{K} - \mathcal{K}_0[1])^{-1} \mathcal{S} \partial_{\nu_c} \tilde{u}$  on  $\partial \Omega_1^+$ . Consequently, the NtD operator  $\mathcal{N}_u$  for any unperturbed domain can be defined as  $\mathcal{N}_u := (\mathcal{K} - \mathcal{K}_0[1])^{-1} \mathcal{S}$ . Here and in the following, we denote the operator of multiplication with a function by the same symbol as that for the function. Note that some authors denote the operator of multiplication by  $\mathcal{K}_0[1]$  in the last formula by  $\mathcal{K}_0[1]I$  and would prefer to write  $\mathcal{N}_u := (\mathcal{K} - \mathcal{K}_0[1]I)^{-1}\mathcal{S}$ .

To approximate  $\mathcal{N}_u$ , we need to discretize the two integral operators and the multiplication operator on the right-hand side of the definition. Suppose now the piecewise smooth boundary curve  $\partial \Omega_1^+$  is parameterized by  $\{x(s) := (x_1(s), x_2(s)) | 0 \le s \le L_1\}$ , which is close to the arclength parameterization. Since corners may exist,  $\tilde{u}(x(s))$  can have corner singularities in its derivatives at corners. To smoothen  $\tilde{u}$ , we introduce a grading function  $s = w(t), 0 \le t \le 1$ . For a smooth segment of  $\partial \Omega_1^+$ corresponding to  $s \in [s^0, s^1]$  and  $t \in [t^0, t^1]$  such that  $s^i = w(t^i)$  for i = 0, 1, where  $s^0$  and  $s^1$  are the parameters of the corners, we take (cf. [13, Eq. (3.104)])

$$s = w(t) := \frac{s^0 w_1^p + s^1 w_2^p}{w_1^p + w_2^p}, \quad t \in [t^0, t^1],$$

where the positive integer p ensures that the derivatives of w(t) up to order p vanish at the corners,

$$w_1 := \left(\frac{1}{2} - \frac{1}{p}\right)\xi^3 + \frac{\xi}{p} + \frac{1}{2}, \quad w_2 := 1 - w_1, \quad \xi := \frac{2t - (t^0 + t^1)}{t^1 - t^0}.$$

To simplify notation, we shall use x(t) to denote x(w(t)), and x'(t) to denote  $\frac{dx}{ds}(w(t))w'(t)$  in the following. Assume that [0, 1] is uniformly sampled by N grid points  $t_j := jh, j = 1, \cdots, N$  with even N and grid size h := 1/N and that the grid points contain all the corner points. Thus,  $\mathcal{S}[\partial_{\nu_c} \tilde{u}]$  at point  $x = x(t_j)$  can be parameterized by

$$\mathcal{S}[\partial_{\nu_c} \tilde{u}] \Big( x(t_j) \Big) = \int_0^1 S(t_j, t) \phi^{\mathrm{s}}(t) \,\mathrm{d}t, \tag{46}$$

where the kernel is  $S(t_j, t) := i/2 \ H_0^{(1)}(k\rho(x(t_j), x(t)))$  and where the scaled co-normal vector  $\phi^{\rm s}(t) := \partial_{\nu_c} \tilde{u}(x(t))|x'(t)|$  is introduced to regularize the approximation of  $\mathcal{N}_u$ . Indeed,  $\phi^{\rm s}(t)$  is smoother than  $\partial_{\nu_c} \tilde{u}(x(t))$ .

Considering the logarithmic singularity of  $S(t_j, t)$  at  $t = t_j$ , we can discretize the integral in (46) by Alpert's 6th-order hybrid Gauss-trapezoidal quadrature rule (cf. [1]) and then get, by trigonometric interpolation,

$$\mathcal{S}[\partial_{\nu_c} \tilde{u}^s] \left[ \begin{array}{c} x(t_1) \\ \vdots \\ x(t_N) \end{array} \right] \approx \mathbf{S} \left[ \begin{array}{c} \phi^{\mathrm{s}}(t_1) \\ \vdots \\ \phi^{\mathrm{s}}(t_N) \end{array} \right],$$

where the  $N \times N$  matrix **S** approximates S. Similarly, the integrals  $\mathcal{K}[\tilde{u}](x(t_j))$  and  $\mathcal{K}_0[1](x(t_j))$  for  $j = 1, \dots, N$  are approximated so that we obtain, on the boundary of  $\partial \Omega_1^+$ ,

$$\begin{bmatrix} \boldsymbol{u}_{1,1} \\ \boldsymbol{u}_{1,2} \\ \boldsymbol{u}_{1,3} \\ \boldsymbol{u}_{1,4} \end{bmatrix} = \boldsymbol{N}_u \begin{bmatrix} \boldsymbol{\phi}_{1,1}^{\mathrm{s}} \\ \boldsymbol{\phi}_{1,2}^{\mathrm{s}} \\ \boldsymbol{\phi}_{1,3}^{\mathrm{s}} \\ \boldsymbol{\phi}_{1,4}^{\mathrm{s}} \end{bmatrix},$$
(47)

where  $\boldsymbol{u}_{1,j'}$  and  $\boldsymbol{\phi}_{1,j'}^s$  for j' = 1, 2, 3, 4 represent  $N_{j'} \times 1$  column vectors of  $\tilde{u}$  and  $\boldsymbol{\phi}^s$ , respectively, at the  $N_{j'}$  grid points of  $\Gamma_{1,j'}$ . Note that  $N = \sum_{j'=1}^4 N_{j'}$  and the grid points on  $\Gamma_{1,3}$  are obtained by horizontally shifting the grid points on  $\Gamma_{1,1}$  to  $\Gamma_{1,3}$  so that  $N_1 = N_3$ . Clearly, the  $N \times N$  matrix  $N_u$  approximates the scaled NtD operator  $\mathcal{N}_u^s$  related to  $\mathcal{N}_u$  by  $\mathcal{N}_u \partial_{\nu_c} \tilde{u} = \mathcal{N}_u^s \phi^s$ . Now, using  $\tilde{u}|_{\Gamma_{1,2}\cup\Gamma_{1,4}} = 0$ , we eliminate the vectors  $\boldsymbol{u}_{1,2}$ ,  $\boldsymbol{u}_{1,4}$ ,  $\boldsymbol{\phi}_{1,2}$  and  $\boldsymbol{\phi}_{1,4}$  in (47) so that we obtain two  $2N_1 \times 2N_1$  matrices  $\boldsymbol{N}^{(0)}$  and  $\boldsymbol{T}$  that satisfy

$$\begin{bmatrix} \boldsymbol{u}_{1,1} \\ \boldsymbol{u}_{1,3} \end{bmatrix} = \boldsymbol{N}^{(0)} \begin{bmatrix} \boldsymbol{\phi}_{1,1}^{\mathrm{s}} \\ \boldsymbol{\phi}_{1,3}^{\mathrm{s}} \end{bmatrix}, \begin{bmatrix} \boldsymbol{\phi}_{1,2}^{\mathrm{s}} \\ \boldsymbol{\phi}_{1,4}^{\mathrm{s}} \end{bmatrix} = \boldsymbol{T} \begin{bmatrix} \boldsymbol{\phi}_{1,1}^{\mathrm{s}} \\ \boldsymbol{\phi}_{1,3}^{\mathrm{s}} \end{bmatrix},$$
(48)

where we denote

$$oldsymbol{N}^{(0)} = \left[egin{array}{ccc} oldsymbol{N}_{00}^{(0)} & oldsymbol{N}_{01}^{(0)} \ oldsymbol{N}_{10}^{(0)} & oldsymbol{N}_{11}^{(0)} \end{array}
ight]$$

with  $N_{ij}^{(0)} \in \mathbb{C}^{N_1 \times N_1}$ . On the continuous level, a representation like (48) follows from the wellposedness of (BVP2) in Theorem 5.4. So we presume that the elimination leading to (48) on the discretized level is stable. Note that, different from [28], we no longer simultaneously assume  $\tilde{u} = \phi^s = 0$ on  $\Gamma_{1,2} \cup \Gamma_{1,4}$ , which could cause pronounced errors in numerical results. Now compare (32) and (48). Like  $N_u$ , the matrix  $N^{(0)}$  approximates the scaled NtD operator  $\mathcal{N}^{(0),s}$  on  $\Gamma_1^+ \cup \Gamma_3^+$  related to  $\mathcal{N}^{(0)}$ by  $\mathcal{N}^{(0)}\partial_{\nu_c}\tilde{u} = \mathcal{N}^{(0),s}\phi^s$ .

Consequently, the previously developed RDP can be easily adapted here in terms of formally replacing  $\mathcal{N}$  by N for the equations (36-42), so that we get two  $N_1 \times N_1$  matrices  $\mathbf{R}_p^+$  and  $\mathbf{N}^+$  approximating the (scaled) Neumann-marching operator  $\mathcal{R}_p^+$  and the (scaled) vNtD operator  $\mathcal{N}^+$  such that  $\phi_{1,3}^{\rm s} = -\mathbf{R}_p^+\phi_{1,1}^{\rm s}$  and  $\mathbf{u}_{1,1} = \mathbf{N}^+\phi_{1,1}^{\rm s}$ . One similarly obtains two  $N_1 \times N_1$  matrices  $\mathbf{R}_p^-$  and  $\mathbf{N}^-$  approximating  $\mathcal{R}_p^-$  and  $\mathcal{N}^-$ , respectively.

#### 6.2 Solving (BVP1)

We are now ready to use the PML-based BIE method to solve the main problem (BVP1). Fix  $x^* \in \Omega_0$ . To eliminate the  $\delta$  function, we consider  $\tilde{u}^{\rm sc}(x;x^*) := \tilde{u}^{\rm og}(x;x^*) - \tilde{u}^{\rm inc}(x;x^*)$ , satisfying (43). For simplicity, we denote  $\Gamma_{0,1} := \Gamma_0^-$ ,  $\Gamma_{0,2} := \Gamma_0$ ,  $\Gamma_{0,3} := \Gamma_0^+$ , and  $\Gamma_{0,4} := \Gamma_0^{H+L}$  (cf. Fig.2 (a)). Then, analogous to (47), on the four boundaries  $\Gamma_{0,j}$ , j = 1, 2, 3, 4, we apply the PML-based BIE method of the previous section to approximate the NtD operator for  $\tilde{u}^{\rm sc}$  and  $\partial_{\nu_c} \tilde{u}^{\rm sc}$  on the boundary of the perturbed cell  $\Omega_0$  by a matrix  $N_p$ ,

$$\begin{bmatrix} \boldsymbol{u}_{0,1}^{\mathrm{sc}} \\ \boldsymbol{u}_{0,2}^{\mathrm{sc}} \\ \boldsymbol{u}_{0,3}^{\mathrm{sc}} \\ \boldsymbol{u}_{0,4}^{\mathrm{sc}} \end{bmatrix} = \boldsymbol{N}_p \begin{bmatrix} \boldsymbol{\phi}_{0,1}^{\mathrm{sc},\mathrm{s}} \\ \boldsymbol{\phi}_{0,2}^{\mathrm{sc},\mathrm{s}} \\ \boldsymbol{\phi}_{0,3}^{\mathrm{sc},\mathrm{s}} \\ \boldsymbol{\phi}_{0,4}^{\mathrm{sc},\mathrm{s}} \end{bmatrix},$$
(49)

where  $u_{0,j}^{sc}$  and  $\phi_{0,j}^{sc,s}$  for j = 1, 2, 3, 4 represent column vectors of  $\tilde{u}^{sc}$  and  $\partial_{\nu_c} \tilde{u}^{sc} |x'|$ , respectively, at the grid points of  $\Gamma_{0,j}$ . Rewriting the above in terms of  $\tilde{u}^{og}$  and  $\partial_{\nu_c} \tilde{u}^{og}$ , we get

$$\begin{bmatrix} \boldsymbol{u}_{0,1}^{\text{og}} \\ \boldsymbol{u}_{0,2}^{\text{og}} \\ \boldsymbol{u}_{0,3}^{\text{og}} \\ \boldsymbol{u}_{0,4}^{\text{og}} \end{bmatrix} = \boldsymbol{N}_p \begin{bmatrix} \boldsymbol{\phi}_{0,1}^{\text{og,s}} \\ \boldsymbol{\phi}_{0,2}^{\text{og,s}} \\ \boldsymbol{\phi}_{0,3}^{\text{og,s}} \\ \boldsymbol{\phi}_{0,3}^{\text{og,s}} \\ \boldsymbol{\phi}_{0,4}^{\text{og,s}} \end{bmatrix} + \begin{bmatrix} \boldsymbol{u}_{0,1}^{\text{inc}} \\ \boldsymbol{u}_{0,2}^{\text{inc}} \\ \boldsymbol{u}_{0,3}^{\text{inc}} \\ \boldsymbol{u}_{0,4}^{\text{inc}} \end{bmatrix} - \boldsymbol{N}_p \begin{bmatrix} \boldsymbol{\phi}_{0,1}^{\text{inc,s}} \\ \boldsymbol{\phi}_{0,2}^{\text{inc,s}} \\ \boldsymbol{\phi}_{0,3}^{\text{inc,s}} \\ \boldsymbol{\phi}_{0,4}^{\text{inc,s}} \end{bmatrix},$$
(50)

where  $u_{0,j}^{\text{inc}}$  and  $\phi_{0,j}^{\text{inc},\text{s}}$  represent column vectors of  $\tilde{u}^{\text{inc}}(x;x^*)$  and  $\partial_{\nu_c}\tilde{u}^{\text{inc}}(x;x^*)|x'|$ , respectively, at the grid points of  $\Gamma_{0,j}$ , etc. The boundary conditions in (BVP1) imply

Solving the linear system (50-51), we get  $\tilde{u}^{\text{og}}(x;x^*)$  and  $\partial_{\nu_c}\tilde{u}^{\text{og}}(x;x^*)$  on all grid points of  $\partial\Omega_0$ .

Finally, we discuss how to evaluate  $\tilde{u}^{\text{og}}(x; x^*)$  in the physical domain  $S_H$ . We distinguish two cases:

- 1.  $x \in \Omega_0$ : Since  $\tilde{u}^{sc}$  and  $\partial_{\nu_c} \tilde{u}^{sc} |x'|$  are available on the grid points of  $\partial \Omega_0$ , we use Green's representation formula (44) with  $\partial \Omega_1^+$  replaced by  $\partial \Omega_0$  to compute  $\tilde{u}^{sc}(x;x^*)$  in  $\Omega_0$  so that  $\tilde{u}^{og}(x;x^*)$  becomes available in  $\Omega_0$ .
- 2.  $x \in \Omega_j^{\pm}$ : Consider  $\Omega_1^+$  first. Suppose  $\boldsymbol{u}_{1,j'}^{\mathrm{og}}$  and  $\phi_{1,j'}^{\mathrm{og},\mathrm{s}}$  represent column vectors of  $\tilde{u}^{\mathrm{og}}$  and  $\partial_{\nu_c} \tilde{u}^{\mathrm{og}} |x'|$  at the grid points of  $\Gamma_{1,j'}$  for  $1 \leq j' \leq 4$ . By the continuity of  $\partial_{\nu_c} \tilde{u}^{\mathrm{og}}$  on the curves  $\Gamma_{1,1} = \Gamma_{0,3} = \Gamma_0^+$ , we obtain  $\phi_{1,1}^{\mathrm{og},\mathrm{s}} = -\boldsymbol{\phi}_{0,3}^{\mathrm{og},\mathrm{s}}$ . Since  $\phi_{1,3}^{\mathrm{og},\mathrm{s}} = -\boldsymbol{R}_p^+ \boldsymbol{\phi}_{1,1}^{\mathrm{og},\mathrm{s}}$ , we get  $\boldsymbol{u}_{1,j'}^{\mathrm{og}}$  for j' = 1, 3 by (48), and  $\boldsymbol{\phi}_{1,j'}^{\mathrm{og},\mathrm{s}}$  for j' = 2, 4. Using  $\boldsymbol{u}_{1,2}^{\mathrm{og}} = \boldsymbol{u}_{1,4}^{\mathrm{og}} = 0$ , the functions  $\tilde{u}^{\mathrm{og}}(x; x^*)$  and  $\partial_{\nu_c} \tilde{u}^{\mathrm{og}} |x'|$  on  $\partial\Omega_1^+$  become available. Hence, the Green's representation formula (44) applies and provides  $\tilde{u}^{\mathrm{og}}(x; x^*)$  in  $\Omega_1^+$ . Repeating the same procedure, one obtains  $\tilde{u}^{\mathrm{og}}(x; x^*)$  in  $\Omega_j^+$  for  $j \geq 2$ . The case for  $x \in \Omega_j^-$  can be handled similarly.

Consequently,  $u^{\text{tot}}(x;x^*) \approx \tilde{u}^{\text{og}}(x;x^*)$  becomes available for  $x \in S_H \subset \overline{\Omega}_0 \cup \left[ \cup_{j=1}^{\infty} \overline{\Omega}_{j,+} \cup \overline{\Omega}_{j,-} \right]$ .

## 6.3 Computing $u^{\text{tot}}$ for plane-wave incidence

To close this section, we briefly discuss how to compute  $u^{\text{tot}}$  for an incident plane wave given by  $u^{\text{inc}} = e^{\mathbf{i}k(\cos\theta x_1 - \sin\theta x_2)}$  with  $\theta \in (0, \pi)$ . First, we consider the unperturbed case  $\Gamma = \Gamma_T$  so that  $u^{\text{tot}}$  becomes the reference solution  $u^{\text{tot}}_{\text{ref}}$ . It is clear that  $u^{\text{sc}}_{\text{ref}} = u^{\text{tot}}_{\text{ref}} - u^{\text{inc}}$  satisfies the following quasi-periodic boundary condition

$$u_{\rm ref}^{\rm sc}\left(-\frac{T}{2}, x_2\right) = \gamma \, u_{\rm ref}^{\rm sc}\left(\frac{T}{2}, x_2\right),\tag{52}$$

$$\partial_{x_1} u_{\text{ref}}^{\text{sc}} \left( -\frac{T}{2}, x_2 \right) = \gamma \, \partial_{x_1} u_{\text{ref}}^{\text{sc}} \left( \frac{T}{2}, x_2 \right),$$
(53)

where  $\gamma = e^{\mathbf{i}k\cos\theta T}$ . On  $\Gamma$ , Equation (2) implies

$$u_{\rm ref}^{\rm sc} = -u^{\rm inc}.$$
 (54)

Due to quasi-periodicity we could express  $u_{ref}^{sc}$  above  $\Gamma_H$  in terms of a Fourier series, i.e.,

$$u_{\text{ref}}^{\text{sc}}(x_1, x_2) = \sum_{j=-\infty}^{\infty} R_j e^{\mathbf{i}\alpha_j x_1 + \mathbf{i}\beta_j x_2}, \quad x_2 \ge H,$$

where  $\alpha_j = k \cos \theta + \frac{2\pi j}{T}$ , where  $\beta_j = \sqrt{k^2 - \alpha_j^2}$  for  $|\alpha_j| \le k$  and  $\beta_j = \mathbf{i} \sqrt{\alpha_j^2 - k^2}$  otherwise, and where  $R_j$  denotes the *j*-th Rayleigh coefficient of the reflected wave. Thus, on the PML boundary  $\Gamma_{H+L}$  the complexified field  $\tilde{u}_{ref}^{sc}(x_1, x_2) = u_{ref}^{sc}(x_1, \tilde{x}_2)$  satisfies

$$\tilde{u}_{\rm ref}^{\rm sc}(x_1, H+L) = \sum_{j=-\infty}^{\infty} R_j e^{\mathbf{i}\alpha_j x_1 + \mathbf{i}\beta_j (H+L) - \beta_j S_c L}.$$

For simplicity, we assume that all  $\beta_j$  are sufficiently away from 0 so that, provided L and  $S_cL$  are sufficiently large, we can directly impose the Dirichlet boundary condition

$$\tilde{u}_{\text{ref}}^{\text{sc}}(x_1, H+L) = 0.$$
(55)

If  $\beta_j$  is quite close to 0, alternative accurate boundary conditions can be developed. We refer the readers to [26, 29, 34] for details. On the other hand,  $\tilde{u}_{ref}^{sc}$  satisfies the quasi-periodic conditions (52) and (53) and the boundary condition (54), but with u replaced by  $\tilde{u}$ .

On the boundary  $\partial \Omega_0$ , the PML-BIE method gives, analogous to (49),

$$\begin{bmatrix} \boldsymbol{u}_{1}^{\mathrm{sc}} \\ \boldsymbol{u}_{2}^{\mathrm{sc}} \\ \boldsymbol{u}_{3}^{\mathrm{sc}} \\ \boldsymbol{u}_{4}^{\mathrm{sc}} \end{bmatrix} = \boldsymbol{N}_{p} \begin{bmatrix} \boldsymbol{\phi}_{1}^{\mathrm{sc}} \\ \boldsymbol{\phi}_{2}^{\mathrm{sc}} \\ \boldsymbol{\phi}_{3}^{\mathrm{sc}} \\ \boldsymbol{\phi}_{4}^{\mathrm{sc}} \end{bmatrix},$$
(56)

where  $\boldsymbol{u}_{j'}^{\mathrm{sc}}$  and  $\phi_{j'}^{\mathrm{sc}}$  for  $1 \leq j' \leq 4$  represent vectors of values of  $\tilde{u}_{\mathrm{ref}}^{\mathrm{sc}}$  and  $\partial_{\nu_c} \tilde{u}_{\mathrm{ref}}^{\mathrm{sc}} |w'|$ , respectively, at the grid points of  $\Gamma_{0,j'}$ . Note that  $\boldsymbol{N}_p$  is the same as  $\boldsymbol{N}_u$  in (47) since  $\Gamma = \Gamma_T$ . Equation (55) directly implies  $\boldsymbol{u}_4^{\mathrm{sc}} = 0$ . The quasi-periodic conditions (52) and (53) imply  $\boldsymbol{u}_3^{\mathrm{sc}} = \gamma \boldsymbol{u}_1^{\mathrm{sc}}$  and  $\phi_3^{\mathrm{sc}} = -\gamma \phi_1^{\mathrm{sc}}$ . The boundary condition (54) indicates

$$\boldsymbol{u}_{2}^{\mathrm{sc}} = -\boldsymbol{u}_{2}^{\mathrm{inc}},\tag{57}$$

where  $u_2^{\text{inc}}$  represents the vector of values of  $u^{\text{inc}}$  at the grid points of  $\Gamma_{0,2}$ . Solving the linear system (56-57) gives rise to values of  $\tilde{u}_{\text{ref}}^{\text{sc}}$  and  $\partial_{\nu_c} \tilde{u}_{\text{ref}}^{\text{sc}} |w'|$  on  $\partial \Omega_0$ . Green's representation formula (44) can help to compute  $\tilde{u}_{\text{ref}}^{\text{sc}}$  in  $\Omega_0$ . The quasi-periodicity helps to construct  $\tilde{u}_{\text{ref}}^{\text{sc}}$  in any other cell  $\Omega_j^{\pm}$  for  $j \in \mathbb{N}^*$ . Consequently,  $u_{\text{ref}}^{\text{tot}}$  becomes available in the whole physical domain  $S_H$ .

Now, if  $\Gamma$  is a local perturbation of  $\Gamma_T$ , then  $\tilde{u}_{ref}^{sc}$  is available by the above arguments, and one follows the same approach developed in Section 6.2 to get  $\tilde{u}^{og} = \tilde{u}^{sc} - \tilde{u}_{ref}^{sc}$  in any unperturbed cell and, therewith,  $u^{tot}$  in the complete physical region  $S_H$ . We omit the details here.

### 7 Numerical examples

In this section, we will carry out four numerical experiments to validate the performance of the PMLbased BIE method and the proposed theory. In all examples, we set the period to T = 1 and the wavelength to the free-space value  $\lambda = 1$  so that  $k_0 = 2\pi$ . We consider two types of incidence: (1) a cylindrical incidence excited at the source point  $x^* = (0, 1.5)$ ; (2) a plane-wave incidence of angle  $\theta$  to be specified. We suppose that only one unit cell of the periodic background structure is perturbed. To setup the PML, we choose m = 0 in (10) to simplify the definition of  $\sigma$ . In the RDP iterations (39), (40), (41) and (42), we take l = 20. Furthermore, we choose H = 3 and set the computational domain to be  $[-5.5, 5.5] \times [-2, 3]$ , which contains 11 cells. To validate the accuracy of our method, we compute the relative error

$$E_{\rm rel} := \frac{||(\phi_{2,0}^{\rm sc,s})^{\rm num} - (\phi_{2,0}^{\rm sc,s})^{\rm exa}||_{\infty}}{||(\phi_{2,0}^{\rm sc,s})^{\rm exa}||_{\infty}}$$

for  $\phi_{2,0}^{\mathrm{sc,s}}$  representing the scaled normal derivative  $|w'|\partial_{\nu}u^{\mathrm{sc}}$  on  $\Gamma_{2,0}$ , the perturbed part of  $\Gamma$ , and for different values of S and L in the setup of the PML. Here superscript "num" indicates numerical solution and superscript "exa" a sufficiently accurate numerical solution or, if available, the exact solution.

#### 7.1 Example 1: A flat curve

In the first example, we assume that  $\Gamma$  is the straight line  $\{x : x_2 = 0\}$ . Certainly, we can regard such a simple structure as a periodic structure with period equal to one wavelength. Formally, we regard the line segment  $\Gamma_{0,2} \subset \Gamma$  between  $x_1 = -0.5$  and  $x_1 = 0.5$  as the "perturbed" part. For the cylindrical incidence, the total wave field  $u^{\text{tot}}$  is given by

$$u^{\text{tot}}(x;x^*) = \frac{\mathbf{i}}{4} \left[ H_0^{(1)}(k|x-x^*|) - H_0^{(1)}(k|x-x^*_{\text{imag}}|) \right],$$

where the image source point is  $x_{imag}^* = (0, -1.5)$ . Using this to compute the scaled co-normal derivative on segment  $\Gamma_{0,2}$ , we get the reference solution and can check the accuracy of our method. We discretize each smooth segment of the "perturbed"/unperturbed unit cell by 600 grid points. To check how the wavenumber condition in Theorem 5.4 affects the accuracy of our numerical solver, we consider two values of the refractive index n in  $\Omega$ : (1) n = 1.03 so that  $kT/\pi = 2.06 \notin \mathcal{E}$ ; (2) n = 1 so that  $kT/\pi = 2 \in \mathcal{E}$ . For both cases, we compare results of Dirichlet and Neumann boundary conditions on  $\Gamma_{H+L}$ .

For n = 1.03, Figure 4 (a) and (b) compare the exact solution and our numerical solution for L = 2.2 and S = 2.8. The two solutions are indistinguishable. To give a detailed comparison, Figure 4 (c)

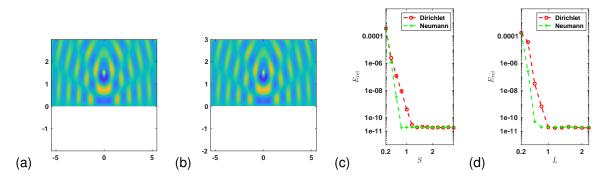


Figure 4: Example 1: Real part of  $u^{\text{tot}}$  in  $[-5.5, 5.5] \times [-2.0, 3.0]$  excited by a source at y = (0, 1.5): (a) exact solution; (b) numerical solution. Convergence history of relative error  $E_{\text{rel}}$  versus: (c) PML absorbing constant S; (d) Thickness of the PML L, for both Dirichlet and Neumann conditions on  $\Gamma_{H+L}$ .

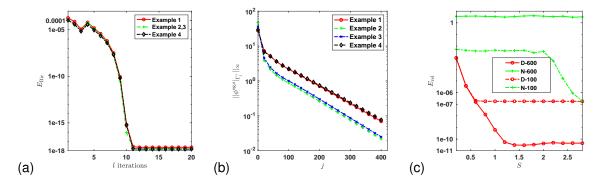


Figure 5: All four examples: (a) Convergence history of  $E_{\text{Ric}}$  in (58) against the number of iterations l; (b) Radiation behavior of  $\phi^{\text{og},s}|_{\Gamma_j^+}$  as  $j \to \infty$ . (c) Performance of Dirichlet and Neumann conditions in Example 1 for n = 1, at which  $kT/\pi \in \mathcal{E}$ ; here 'D' stands for Dirichlet and 'N' for Neumann, and 100 indicates 100 grid points are used to discretize each smooth segment of the unit cells, etc.

and (d) show how the relative error  $E_{\rm rel}$  decays as one of the two PML parameters, the absorbing constant S and the thickness L, increases for either zero Dirichlet or zero Neumann condition on  $\Gamma_{H+L}$ . In Figure 4(c), we take L = 2.2 and let S vary between 0.2 and 2.8, while in Figure 4(d), we take S = 2.8 and let the PML thickness L vary between 0.2 and 2.2. In both figures, the vertical axis is logarithmically scaled so that the vertical dashed lines indicate that the relative error  $E_{\rm rel}$  decays exponentially as L or S increases for both conditions. On the other hand, Neumann condition gives faster convergence rates than Dirichlet condition. The convergence curves indicate that nearly 11 significant digits are revealed by the proposed PML-based BIE method. The 'o' lines in Figure 5(a) show the convergence curve of

$$E_{\rm Ric} = \left\| \boldsymbol{N}_{10}^{(0)} + \left[ \boldsymbol{N}_{11}^{(0)} + \boldsymbol{N}_{00}^{(0)} \right] \boldsymbol{R}_p^+ + \boldsymbol{N}_{01}^{(0)} (\boldsymbol{R}_p^+)^2 \right\|_{\infty}$$
(58)

against the number of iterations l. It can be seen that, after no more than 11 iterations,  $\mathbf{R}_p^+$  satisfies its governing Riccati equation (33) up to round-off errors. The 'o' lines in Figure 5(b) show the curve of  $||\phi^{\mathrm{og},\mathrm{s}}|_{\Gamma_j^+}||_{\infty}$  against j. It can be seen that  $\phi^{\mathrm{og},\mathrm{s}}$  and, hence,  $\partial_{\nu_c^+} u^{\mathrm{og}}$  indeed decays exponentially as j or  $x_1$  increases, as has been claimed in Corollary 5.2.

In Figure 5(c), we compare Dirichlet and Neumann conditions for n = 1. We take L = 2.2 and let S vary from 0.2 to 2.8. Among the four convergence curves, solid lines indicate 600 grid points chosen on each smooth segment of each unit cell, while dashed lines indicate 100 grid points. Graphs with '+' indicates Neumann condition on  $\Gamma_{H+L}$ , while 'o' indicates Dirichlet condition. If 100 grid points are used, the error  $E_{\rm rel}$  for Neumann condition starts decreasing after  $S \ge 2$ , whereas  $E_{\rm rel}$  for Dirichlet condition has already reached its minimal error. If 600 grid points are used, Neumann condition does not lead to a converging  $E_{\rm rel}$  for  $S \in [0.2, 2.8]$ , but Dirichlet condition still shows the same convergence rate and accuracy as in the case n = 1.03. Consequently, Dirichlet condition outperforms Neumann condition for n = 1.

#### 7.2 Example 2: A sine curve

In the second example, we assume that the boundary  $\Gamma$  is the sine curve  $\{x : x_2 = \sin(2\pi x_1 + \pi)\}$  (cf. Figure 6(a)) and that n = 1.03 to obtain  $kT/\pi \notin \mathcal{E}$ . For the cylindrical incidence, we discretize each smooth segment of any unit cell by 600 grid points, and compare results of Dirichlet and Neumann boundary conditions on  $\Gamma_{H+L}$ . Taking S = 2.8 and L = 2.2, we evaluate the wave field in the

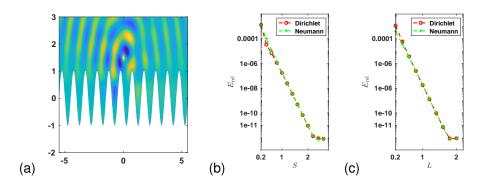


Figure 6: Example 2: (a) Numerical solution of real part of the total wave field u in the domain  $[-5.5, 5.5] \times [-2.0, 3.0]$  excited by a point source at y = (0, 1.5). Convergence history of relative error  $E_{\rm rel}$  versus: (b) PML absorbing constant S for fixed PML thickness L = 2, (c) PML thickness L for fixed PML absorbing constant S = 2.8; vertical axes are logarithmically scaled.

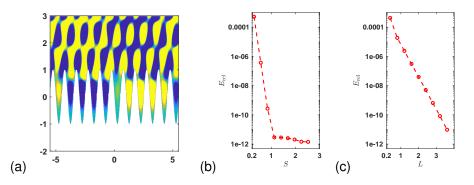


Figure 7: Example 2: (a) Numerical solution of real part of the total wave field u in the domain  $[-5.5, 5.5] \times [-2.0, 3.0]$  excited by a plane-wave incidence of angle  $\theta = \pi/3$ . Convergence history of relative error  $E_{\rm rel}$  versus: (b) PML absorbing constant S for fixed PML thickness L = 4, (c) PML thickness L for fixed PML absorbing constant S=2.8; vertical axes are logarithmically scaled.

domain  $[-5.5, 5.5] \times [-2.0, 3.0]$  and use this as the reference solution since the exact solution is no longer available. In Figure 6, (a) shows the field pattern of the reference solution, and (b) and (c) show the convergence history of relative error  $E_{\rm rel}$  versus one of the two PML parameters S and L, respectively. Again, we observe that  $E_{\rm rel}$  decays exponentially as S or L increases. Unlike the flat surface in Example 1, we no longer observe a faster convergence rate of Neumann condition, but find that both conditions share the same convergence rate and accuracy. Comparing with the bad result for Dirichlet condition and  $kT/\pi \in \mathcal{E}$  and the impressive improvement for  $kT/\pi \notin \mathcal{E}$ , we conclude that, in the sine curve example, Neumann condition is less superior than Dirichlet condition, and thus we shall only use the latter one in the remaining experiments. With Dirichlet condition, the '+' lines in Figure 5 (a) show the curve of  $||\phi^{\mathrm{og},\mathrm{s}}|_{\Gamma_{+}^{+}}||_{\infty}$  against j.

For the plane-wave incidence, we take  $\theta = \pi/3$ . Employing the method in Section 6.3, we discretize each smooth segment of any unit cell by 700 grid points. Taking S = 2.8 and L = 4, we evaluate the wave field in  $[-5.5, 5.5] \times [-2.0, 3.0]$  and use this as the reference solution. In Figure 7, (a) shows the field pattern, and (b) and (c) show the convergence history of relative error  $E_{\rm rel}$  versus one of the two PML parameters S and L, respectively. For both incidences, the convergence curves in Figures 6 and 7 decay exponentially, indicating that nearly 12 significant digits are revealed by the proposed

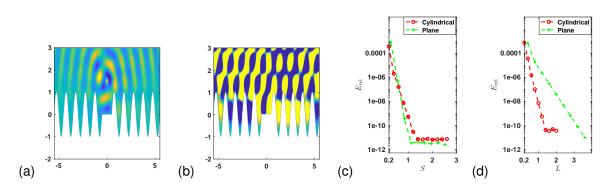


Figure 8: Example 3: Numerical solution of real part of the total wave field u in  $[-5.5, 5.5] \times [-2.0, 3.0]$  excited by: (a) a cylindrical wave with source at y = (0, 1.5); (b) a plane-wave incidence of angle  $\theta$  set to  $\pi/3$ . Convergence history of relative error  $E_{rel}$  versus: (c) PML thickness L for fixed PML absorbing constant S = 2.8 for both incidences; (d) PML absorbing constant S for fixed PML thickness L = 2.2 (4.0) for cylindrical (plane-wave) incidence.

PML-based BIE method.

#### 7.3 Example 3: A locally perturbed sine curve

In the third example, we assume that  $\Gamma$  is the sine curve  $\Gamma_T := \{x : x_2 = \sin(2\pi x_1 + \pi)\}$  locally perturbed such that the part between  $x_1 = -0.5$  and  $x_1 = 0.5$  is replaced by the line segment  $\{(x_1, 0) : x_1 \in [-0.5, 0.5]\}$  (cf. Figure 8 (a)). For the cylindrical incidence, we discretize each smooth segment of any unit cell by 600 grid points. Taking S = 2.8 and L = 2.2, we evaluate the wave field in  $[-5.5, 5.5] \times [-2.0, 3.0]$  and use this as the reference solution, the field pattern of which is shown in Figure 8 (a). The 'x' lines in Figure 5 (b) show the curve of  $||\phi^{\text{og},s}|_{\Gamma^+}||_{\infty}$  against j.

For the plane-wave incidence, we take  $\theta = \pi/3$  and discretize each smooth segment of any unit cell by 700 grid points. Taking S = 2.8 and L = 4, we evaluate the wave field in  $[-5.5, 5.5] \times [-2.0, 3.0]$  and use this as the reference solution, the field pattern of which is shown in Figure 8 (b).

For both incidences, Figure 8 (c) and (d) show the convergence history of relative error  $E_{\rm rel}$  versus one of the two PML parameters S and L, respectively. The convergence curves decay exponentially and indicate that nearly 11 significant digits are revealed by the proposed PML-based BIE method.

### 7.4 Example 4: A locally perturbed binary grating

In the last example, we assume that the boundary  $\Gamma$  consists of periodic rectangular grooves of depth 0.5 and width 0.25, with the part between  $x_1 = -0.5$  and  $x_1 = 0.5$  replaced by the line segment  $\{(x_1, 0) : x_1 \in [-0.5, 0.5]\}$  (cf. Figure 9(a)). For the cylindrical incidence, we discretize each smooth segment of any unit cell by 600 grid points. Taking S = 2.8 and L = 2.2, we evaluate the wave field in  $[-5.5, 5.5] \times [-2.0, 3.0]$  and use this as the reference solution, the field pattern of which is shown in Figure 9 (a). The ' $\Diamond$ ' lines in Figure 5 (a) show the convergence curve of  $E_{\rm Ric}$  in (58) against the number of iterations l. The ' $\Diamond$ ' lines in Figure 5 (b) show the curve of  $||\phi^{\rm og,s}|_{\Gamma^+_{\rm r}}||_{\infty}$  against j.

For the plane-wave incidence, we take  $\theta = \pi/6$  and discretize each smooth segment of any unit cell by 600 grid points. Taking S = 2.8 and L = 3, we evaluate the wave field in  $[-5.5, 5.5] \times [-2.0, 3.0]$  and use this as the reference solution, the field pattern of which is shown in Figure 9 (b).

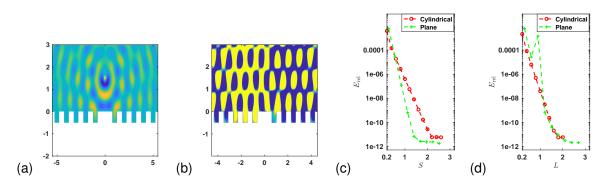


Figure 9: Example 4: Numerical solution of real part of the total wave field u in  $[-5.5, 5.5] \times [-2.0, 3.0]$  excited by: (a) a cylindrical wave with source at y = (0, 1.5); (b) a plane wave of incident angle  $\theta = \pi/6$ . Convergence history of relative error  $E_{\rm rel}$  versus: (c) PML thickness L for fixed PML absorbing constant S = 2.8 for both incidences; (d) PML absorbing constant S for fixed PML thickness L = 2.2 (3.0) for cylindrical (plane-wave) incidence.

For both incidences, Figure 9 (c) and (d) show the convergence history of relative error  $E_{\rm rel}$  versus one of the two PML parameters S and L, respectively. The convergence curves decay exponentially and indicate that nearly 12 significant digits are revealed by the proposed PML-based BIE method.

## 8 Conclusion

This paper studied the perfectly-matched-layer (PML) theory for wave scattering in a half space of homogeneous medium bounded by a two-dimensional, perfectly conducting, and locally defected periodic surface, and developed a high-accuracy boundary-integral-equation (BIE) solver. By placing a PML in the vertical direction to truncate the unbounded domain to a strip, we proved that the PML solution converges to the true solution in the physical subregion of the strip at an algebraic order of the PML thickness. Laterally, the unbounded strip is divided into three regions: a region containing the defect and two semi-waveguide regions of periodic subsurfaces, separated by two vertical line segments. We proved the well-posedness of an associated scattering problem in both semi-waveguides. Based on this, we defined Neumann-to-Dirichlet (NtD) operators on the associated vertical segments. The two NtD operators, serving as exact lateral boundary conditions, reformulate the unbounded strip problem as a boundary value problem over the defected region. Each NtD operator is closely related to a Neumann-marching operator, governed by a nonlinear Riccati equation, which was efficiently solved by an RDP method and a high-accuracy PML-based BIE method so that the boundary value problem on the defected region can be solved finally. Our future research plan shall focus on the following two aspects:

- (1). Extend the current work to study locally defected periodic structures of stratified media. In such case, propagating Bloch modes may exist so that the related Neumann marching operators  $\mathcal{R}_p^{\pm}$  may not be contracting.
- (2). Rigorously justify that the PML solution converges exponentially to the true solution in any compact subset of the strip, as has been demonstrated by numerical experiments.

### References

- B.K. Alpert. Hybrid Gauss-trapezoidal quadrature rules. SIAM Journal on Scientific Computing, 20(5):1551–1584, 1999.
- [2] T. Arens and T. Hohage. On radiation conditions for rough surface scattering problems. *IMA Journal of Applied Mathematics*, 70(6):839–847, 2005.
- [3] G. Bao, D.C. Dobson, and J.A. Cox. Mathematical studies in rigorous grating theory. *J. Opt. Soc. Am. A*, 12(5):1029–1042, 1995.
- [4] J.-P. Bérenger. A perfectly matched layer for the absorption of electromagnetic waves. J. Comput. Phys., 114(2):185 – 200, 1994.
- [5] S.N. Chandler-Wilde and J. Elschner. Variational approach in weighted Sobolev spaces to scattering by unbounded rough surfaces. SIAM J. Math. Anal., 42:2554–2580, 2010.
- [6] S.N. Chandler-Wilde and P. Monk. The PML for rough surface scattering. *Applied Numerical Mathematics*, 59:2131–2154, 2009.
- [7] S.N. Chandler-Wilde and P. Monk. Existence, uniqueness and variational methods for scattering by unbounded rough surfaces. SIAM J. Math. Anal., 37:598–618, 2005.
- [8] S.N. Chandler-Wilde, C.R. Ross, and B. Zhang. Scattering by infinite one-dimensional rough surfaces. Proc. R. Soc. Lon. A, 455:3767–3787, 1999.
- [9] S.N. Chandler-Wilde and B. Zhang. Electromagnetic scattering by an inhomogeneous conducting or dielectric layer on a perfectly conducting plate. *Proc. Roy. Soc. London A*, 454:519âĂŞ542, 1998.
- [10] Z. Chen and H. Wu. An adaptive finite element method with perfectly matched absorbing layers for the wave scattering by periodic structures. *SIAM J. Numer. Anal.*, 41(3):799–826, 2003.
- [11] W.C. Chew. Waves and fields in inhomogeneous media. IEEE PRESS, New York, 1995.
- [12] W.C. Chew and W.H. Weedon. A 3D perfectly matched medium for modified Maxwell's equations with stretched coordinates. *Microwave and Optical Technology Letters*, 7(13):599–604, 1994.
- [13] D. Colton and R. Kress. Inverse Acoustic and Electromagnetic Scattering Theory (3rd Edition). Springer, 2013.
- [14] J.A. DeSanto and P.A. Martin. On angular-spectrum representations for scattering by infinite rough surfaces. *Wave Motion*, 24:421–433, 1996.
- [15] M. Ehrhardt, H. Han, and C. Zheng. Numerical simulation of waves in periodic structures. Commun. in Comp. Phys., 5(5):849–870, 2009.
- [16] A. Fichtner. Full Seismic Waveform Modelling and Inversion. Springer, 2011.
- [17] S.F. Helfert and R. Pregla. Efficient analysis of periodic structure. J. Lightwave Technol., 16:1694– 1702, 1998.
- [18] G. Hu, W. Lu, and A. Rathsfeld. Time-harmonic acoustic scattering from locally perturbed periodic curves. *submitted*, 2020.

- [19] Z. Hu and Y.Y. Lu. Efficient numerical method for analyzing coupling structures of photonic crystal waveguides. *IEEE Photon. Tech. Lett.*, 21(23):1737–1739, 2009.
- [20] S. Johnson. Notes on Perfectly Matched Layers (PMLs), http://www-math.mit.edu/stevenj/18.369/spring09/pml.pdf. Unpublished, 2008.
- [21] P. Joly, J-R. Li, and S. Fliss. Exact boundary conditions for periodic waveguides containing a local perturbation. *Commun. in Comp. Phys.*, 1(6):945–973, 2006.
- [22] T. Kato. *Perturbation Theory for Linear Operators*. Classics in Mathematics, Springer-Verlag, Berlin, 1995, Reprint of the 1980 edition.
- [23] M. Lassas and E. Somersalo. Analysis of the PML equations in general convex geometry. *Proceedings of the Royal Society of Edinburgh: Section A Mathematics*, 131(5):1183–1207, 2001.
- [24] A. Lechleiter and R. Zhang. A Floquet-Bloch transform based numerical method for scattering from locally perturbed periodic surfaces. SIAM J. Sci. Comput., 39(5):B819–B839, 2017.
- [25] W. Lu. Mathematical analysis of wave radiation by a step-like surface. *SIAM J. Appl. Math.*, 81(2):666–693, 2021.
- [26] W. Lu and Y.Y. Lu. High order integral equation method for diffraction gratings. J. Opt. Soc. Am. A, 29(5):734–740, 2012.
- [27] W. Lu and Y.Y. Lu. Efficient high order waveguide mode solvers based on boundary integral equations. *J. Comput. Phys.*, 272:507 525, 2014.
- [28] W.Lu, Y.Y.Lu, and J. Qian. Perfectly matched layer boundary integral equation method for wave scattering in a layered medium. SIAM J. Appl. Math., 78(1):246–265, 2018.
- [29] W. Lu, Y. Y. Lu, and D. Song. A numerical mode matching method for wave scattering in a layered medium with a stratified inhomogeneity. SIAM J. Sci. Comput., 41(2):B274–B294, 2019.
- [30] W. McLean. *Strongly Elliptic Systems and Boundary Integral Equations*. Cambridge University Press, New York, NY, 2000.
- [31] P. Monk. Finite Element Methods for Maxwell's Equations. Oxford University Press, 2003.
- [32] J. Sun and C. Zheng. Numerical scattering analysis of te plane waves by a metallic diffraction grating with local defects. *J. Opt. Soc. Am. A*, 26(1):156–162, 2009.
- [33] L. Yuan and Y.Y. Lu. A recursive doubling Dirichlet-to-Neumann map method for periodic waveguides. *J. Lightwave Technol.*, 25:3649–3656, 2007.
- [34] W. Zhou and H. Wu. An adaptive finite element method for the diffraction grating problem with PML and few-mode DtN truncations. *J. Sci. Comput.*, 76:1813–1838, 2018.