SIAM J. NUMER. ANAL.
Vol. 53, No. 3, pp. 1585–1607


# AN ADAPTIVE FINITE ELEMENT METHOD FOR THE DIFFRACTION GRATING PROBLEM WITH TRANSPARENT BOUNDARY CONDITION*

ZHOUFENG WANG†, GANG BAO‡, JIAQING LI§, PEIJUN LI¶, AND HAIJUN WU‖

**Abstract.** The diffraction grating problem is modeled by a boundary value problem governed by a Helmholtz equation with transparent boundary conditions. An a posteriori error estimate is derived when the truncation of the nonlocal boundary operators takes place. To overcome the difficulty caused by the fact that the truncated Dirichlet-to-Neumann (DtN) mapping does not converge to the original DtN mapping in its operator norm, a duality argument without assuming more regularity than the weak solution is applied. The a posteriori error estimate consists of two parts, the finite element discretization error and the truncation error of boundary operators which decays exponentially with respect to the truncation parameter. Based on the a posteriori error control, a finite element adaptive strategy is established for the diffraction grating problem, such that the truncation parameter is determined through the truncation error and the mesh elements for local refinements are marked through the finite element discretization error. Numerical experiments are presented to illustrate the competitive behavior of the proposed adaptive algorithm.

**Key words.** Helmholtz equation, transparent boundary condition, a posteriori error estimates, adaptive algorithm, diffractive optics

**AMS subject classifications.** 65N12, 65N15, 65N30, 78A40

**DOI.** 10.1137/140969907

## 1. Introduction.

The diffraction grating problem is referred to as wave scattering by periodic structures. Scattering theory in periodic structures, which is crucial in application fields of microoptics (cf. [10]), has been studied extensively in the past several decades. We can refer to Petit [31], Dobson and Friedman [22], Ammari and Bao [2], Bao, Dobson, and Cox [11], Bao [6, 7], Bao, Cao, and Yang [8], Abboud [1], Ammari and Nédélec [3], Berenger [14], Dobson [23], and Yachin and Yasumoto [36] for both a good introduction and the existence, uniqueness, and numerical approximations of solutions to grating problems. This problem still receives considerable attention in the applied mathematical community. A broader review on the diffractive optics technology and Maxwell's equations can be found in Bao, Cowsar, and Masters [10] and Ammari and Bao [2]. Especially, Chen and Wu [21] proposed a new

*Received by the editors May 20, 2014; accepted for publication (in revised form) February 18, 2015; published electronically June 30, 2015.
http://www.siam.org/journals/sinum/53-3/96990.html

†Department of Mathematics, Nanjing University, Jiangsu 210093, People's Republic of China. Current address: School of Mathematics and Statistics, Henan University of Science and Technology, Luoyang 471023, China (zfwang801003@126.com).

‡Department of Mathematics, Zhejiang University, Hangzhou 310027, China (baog@zju.edu.cn). The research of this author was supported in part by a Key Project of the Major Research Plan of NSFC (91130004), an NSFC A3 Project (11421110002), an NSFC Tianyuan Project (11426235), and a special research grant from Zhejiang University.

§The 28th Research Institute of China Electronics Technology Group Corporation, Nanjing, Jiangsu 210007, People's Republic of China (ljqnju@gmail.com).

¶Department of Mathematics, Purdue University, West Lafayette, IN 47907 (lipeijun@math.purdue.edu). The research of this author was supported in part by NSF grant DMS-1151308.

‖Department of Mathematics, Nanjing University, Jiangsu 210093, People's Republic of China (hjw@nju.edu.cn). The research of this author was supported in part by the National Magnetic Confinement Fusion Science Program under grant 2011GB105003 and by the NSF of China under grant 91130004.

numerical approach with combinations of the adaptive finite element method and the perfectly matched layer (PML) technique for the one-dimensional (1D) grating problem. Based on this numerical tool, there also exist lots of results in the literature for the diffractive grating problem; we can refer to Bao, Chen, and Wu [9], Bao and Wu [13], Chen and Liu [20], and Chen and Chen [17]. One of the advantages of this approach is that PML can be used to deal with the difficulty in truncating the unbounded domain. On the other hand, the adaptive finite element method can very efficiently capture the local singularities.

A posteriori error estimates, which measure the actual discrete errors without knowledge of the limit solutions, are of crucial importance in designing algorithms for mesh modifications that equidistribute the computational effort and optimize the computation. The adaptive finite element methods based on the a posteriori error estimates have become a class of important numerical tools for solving differential equations, especially for those which have the physical feature of multiscale phenomenon. Ever since the pioneering work of Babuška and Rheinboldt [5], this method has received lots of attention and undergone intensive study by many researchers; see, e.g., [28, 29, 35, 24, 27, 9, 16, 21, 18, 19, 20]. For the convergence of adaptive finite element methods, we refer to Morin, Nochetto, and Siebert [29, 30], Chen and Dai [18], Dörfler [24], and Mekchay and Nochetto [26]. Studies on the quasi-optimality of adaptive finite element methods can be found in Cascon et al. [16], Binev, Dahmen and DeVore [15], and Stevenson [34]. The adaptive finite element method is very popular in grating problems (cf. [9, 17, 12, 21, 25]), largely because it greatly improves the convergence speed of numerical solution for problems with local singularities.

The aim of this paper is to develop an adaptive finite element method with truncated Dirichlet-to-Neumann (DtN) boundary condition for solving the 1D grating problem. In this approach, no extra artificial domain needs to be imposed to surround the computational domain, which is totally different from the perfectly matched absorbing layers technique. It also should be pointed out that the truncated boundary operators in numerical schemes are determined by taking sufficiently many terms of the corresponding infinite series expansions. Note that the truncated DtN mapping does not converge to the original DtN mapping in its operator norm. The a posteriori analysis of the adaptive PML method [9, 20, 21] cannot apply directly to our adaptive DtN case, since the fact was used that the DtN mapping of the truncated PML problem converges exponentially fast to the original DtN mapping. To overcome this difficulty, we develop a duality argument similar to that (or the so-called Schatz argument) for the a priori error estimates for indefinite problems [32, 6], but without assuming more regularity of the dual problem than $H^1$. Finally we obtain an a posteriori error estimate between the solution of the scattering problem and the finite element solution. The a posteriori error estimate, which consists of two parts, the finite element discretization error and the truncation error of boundary operators, is used to design the adaptive finite element algorithm to choose elements for refinements and to determine the truncation parameter $N$. We remark that the truncation error part decays exponentially with respect to $N$ and as a consequence the choice of the truncation parameter $N$ is insensitive to the given tolerance. The numerical experiments demonstrate a comparable behavior to [21] and show much more competitive efficiency by adaptively refining the mesh as compared with uniformly refining the mesh. Thus, the present work provides a viable alternative to the adaptive finite element method with PML for solving the same grating problem. The algorithm is also expected to be used to solve many other scattering problems and even more general partial differential equations where transparent boundary conditions are available but PML may not be implemented or cannot be applied.

The outline of this paper is as follows. In section 2, we briefly introduce the weak formulation for the transverse electric (TE) case of the 1D grating problem with the transparent boundary condition. In section 3 we introduce the finite element discretization. A crucial a posteriori estimate is also stated. Section 4 is devoted to the finite element analysis, including proving some important lemmas and the derivation of an a posteriori error estimate, which lays down the basis of the adaptive algorithm. In section 5 we discuss the implementation of the adaptive algorithm and present several numerical examples to illustrate the performance of the proposed method. In section 6 we summarize our research work in this paper and forecast future research directions.

**2. Problem formulation.** In this section, we introduce a mathematical model for the diffraction grating problem and its weak formulation by using the transparent boundary condition.

**2.1. Model problem.** The electromagnetic fields in the whole space are governed by the following time-harmonic Maxwell equations:

$$\nabla \times \mathbf{E} - i\omega\mu\mathbf{H} = 0, \tag{2.1}$$

$$\nabla \times \mathbf{H} + i\omega\varepsilon\mathbf{E} = 0, \tag{2.2}$$

where $\omega$ is the angular frequency, $\varepsilon$ is the dielectric permittivity, $\mu$ is the magnetic permeability and is defined as a positive constant everywhere, i.e., the media is assumed to be nonmagnetic, and $\mathbf{E}$ and $\mathbf{H}$ denote the electric field and the magnetic field in $\mathbb{R}^3$, respectively. In this paper, our attention is restricted to the two-dimensional setting, i.e., we only consider a 1D grating problem. The more sophisticated biperiodic diffraction grating, which belongs to the three-dimensional problem, will be discussed in a separate work. It can be assumed that the medium parameters and the grating settings are invariant in the $x_2$ direction. In the meantime we assume that the dielectric coefficient is periodic in the $x_1$ direction with period $L$:

$$\varepsilon(x_1 + nL, x_3) = \varepsilon(x_1, x_3) \quad \text{for all} \quad x_1, x_3 \in \mathbb{R}, \quad n \in \mathbb{N}.$$

Here we assume that $\varepsilon \in L^\infty(\mathbb{R}^2)$ with $\operatorname{Im}\varepsilon \geq 0$ and $\operatorname{Re}\varepsilon > 0$ whenever $\operatorname{Im}\varepsilon = 0$.

First, the problem geometry is defined as

$$\Omega_0 = \{(x_1, x_3) \in \mathbb{R}^2 : b_2 < x_3 < b_1\}$$

for some positive constants $b_2$ and $b_1$.

Figure 1 shows the structure of the problem geometry, where $s_1$ and $s_2$ are two simple curves embedded in the region $\Omega_0$. The medium is described on the whole as inhomogeneous, so specifically, the medium in the region $\Omega_0$ between $s_1$ and $s_2$ is inhomogeneous, yet the medium is homogeneous above the curve $s_1$ and below the curve $s_2$. Based on the characteristics of the medium, it is assumed that there exist positive constants $d_1$ and $d_2$ such that

$$\begin{aligned} \varepsilon(x_1, x_3) &= \varepsilon_1 \text{ in } \Omega_0^1 = \{(x_1, x_3) \in \mathbb{R}^2 : x_3 \geq b_1 - d_1\}, \\ \varepsilon(x_1, x_3) &= \varepsilon_2 \text{ in } \Omega_0^2 = \{(x_1, x_3) \in \mathbb{R}^2 : x_3 \leq b_2 + d_2\}, \end{aligned}$$

where $\varepsilon_1$ and $\varepsilon_2$ are constants. In practical applications, we have $\varepsilon_1 > 0$, but $\varepsilon_2$ may be complex, which depends on substrate material used in $\Omega_0^2$.

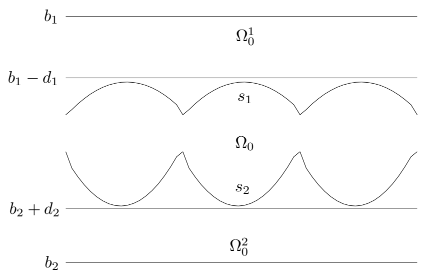


FIG. 1. *Geometry of the grating problem.*

The Maxwell equations (2.1) and (2.2) can be simplified by considering two fundamental polarizations: the TE polarization and the TM (transverse magnetic) polarization. Thus the vector Maxwell equations can be finally reduced to the scalar Helmholtz equation. In the TE case, $\mathbf{E} = (0, u, 0)^\top \in \mathbb{R}^3$ and $u = u(x_1, x_3)$ satisfies the Helmholtz equation

$$\Delta u(x) + k^2(x)u(x) = 0 \quad \text{in } \mathbb{R}^2, \tag{2.3}$$

where $k^2(x) = \omega^2\varepsilon(x)\mu$ is the magnitude of the wave vector. In the TM case, $\mathbf{H} = (0, u, 0)^\top \in \mathbb{R}^3$ and $u = u(x_1, x_3)$ satisfies the following equation:

$$\operatorname{div}(k^{-2}(x)\nabla u(x)) + u(x) = 0 \quad \text{in } \mathbb{R}^2. \tag{2.4}$$

**2.2. The weak formulation.** Denote by $k_j = \omega\sqrt{\varepsilon_j\mu}$ the wave number in $\Omega_0^j, j = 1, 2$. Let $u_I = e^{i\alpha x_1 - i\beta x_3}$ be the incident plane wave, where $\alpha = k_1 \sin\theta$, $\beta = k_1 \cos\theta$, and $-\frac{\pi}{2} < \theta < \frac{\pi}{2}$ is the incident angle. In the following, we will seek quasi-periodic solutions $u$ of (2.3), where quasi-periodic means that $u_\alpha = ue^{-i\alpha x_1}$ is periodic in the $x_1$ direction with period $L$.

As seen in Figure 2, due to the quasi-periodicity, the problem may be reduced into the bounded domain

$$\Omega = \{0 < x_1 < L,\ b_2 < x_3 < b_1\}.$$

Define the boundaries $\Gamma_1 = \{0 < x_1 < L, x_3 = b_1\}$, $\Gamma_2 = \{0 < x_1 < L, x_3 = b_2\}$. The curves $s_1$ and $s_2$ divide $\Omega$ into three connected components, Denote the component which meets $\Gamma_1$ by $\Omega^1$, the component which meets $\Gamma_2$ by $\Omega^2$, and let $\Omega^* = \Omega - \overline{\Omega}^1 - \overline{\Omega}^2$. Introduce the sets

$$\Gamma_1' = \{0 < x_1 < L,\ x_3 = b_1'\}, \quad \Gamma_2' = \{0 < x_1 < L,\ x_3 = b_2'\},$$

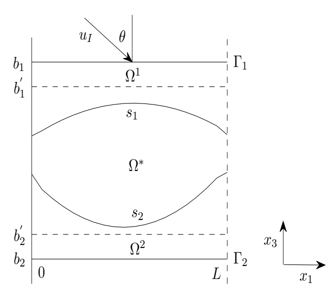


FIG. 2. *Geometry of the grating problem in one period L.*

with $b_2 < b_2' < b_1' < b_1$ such that $\Omega^* \subseteq \{0 < x_1 < L,\ b_2' < x_3 < b_1'\}$, $|b_j - b_j^{'}| \le d_j$, $j = 1, 2$, and

$$\varepsilon(x_1, x_3) = \varepsilon_1 \quad \text{for } x_3 \ge b_1',$$
$$\varepsilon(x_1, x_3) = \varepsilon_2 \quad \text{for } x_3 \le b_2'.$$

For any quasi-periodic function $f$ which has the expansion $f = \sum_{n\in\mathbb{Z}} f_\alpha^{(n)} e^{i(\alpha_n+\alpha)x_1}$, we introduce the DtN operator $\Gamma^{(j)}$ as defined in [6, 21]:

$$T^{(j)} f(x_1) = \sum_{n\in\mathbb{Z}} i\beta_j^n f_\alpha^{(n)} e^{i(\alpha_n+\alpha)x_1},\ 0 < x_1 < L,\ j = 1, 2, \tag{2.5}$$

where $\alpha_n = 2\pi n/L$ and

$$(\beta_j^n)^2 = k_j^2 - (\alpha_n + \alpha)^2, \quad \mathrm{Im}(\beta_j^n) \ge 0. \tag{2.6}$$

Note that $\beta_1^0 = \beta$ by definition. It follows from the Rayleigh expansions that we have the following transparent boundary conditions [6, 7]:

$$\frac{\partial(u - u_I)}{\partial n} - T^{(1)}(u - u_I) = 0 \ \ on\ \Gamma_1,\ \frac{\partial u}{\partial n} - T^{(2)} u = 0 \ \ on\ \Gamma_2, \tag{2.7}$$

where $n$ denotes the unit outer normal on $\Gamma_j$, $j = 1, 2$.

Introduce a subspace $X(\Omega)$ of $H^1(\Omega)$, which includes all the quasi-periodic functions in $H^1(\Omega)$:

$$X(\Omega) = \{w \in H^1(\Omega) : w(0, x_3) = e^{-i\alpha L} w(L, x_3) \ \text{ for } \ b_2 < x_3 < b_1\}.$$

Multiplying the complex conjugate of a test function $\psi$ in $X(\Omega)$, integrating over $\Omega$, and using Green formulas for (2.3) with the boundary conditions (2.7), we arrive at the weak formulation in the TE polarization: Giving an incident plane wave $u_I = e^{i\alpha x_1 - i\beta x_3}$, find $u \in X(\Omega)$ such that

$$a_{TE}(u, \psi) = -\int_{\Gamma_1} 2i\beta u_I \overline{\psi} dx_1 \ \text{ for all } \ \psi \in X(\Omega), \tag{2.8}$$

where a sesquilinear form $a_{TE}$:$X(\Omega) \times X(\Omega) \to C$ as follows:

$$a_{TE}(\varphi, \psi) = \int_\Omega (\nabla\varphi \cdot \nabla\overline{\psi} - k^2(x)\varphi\overline{\psi}) dx - \sum_{j=1}^{2} \int_{\Gamma_j} (T^{(j)}\varphi)\overline{\psi} dx_1. \tag{2.9}$$

Existence and uniqueness of the solution to (2.8) is strictly proved in Dobson [23], and the corresponding results of the TM case may be found in [7]. Throughout this paper, we assume that the variational problem (2.8) has a unique solution for any frequency $\omega$. The general theory in Babuška and Aziz [4] implies that there exists a constant $\gamma > 0$ such that the following inf-sup condition holds:

$$\sup_{0\ne\psi\in X(\Omega)} \frac{|a_{TE}(\varphi, \psi)|}{\|\psi\|_{H^1(\Omega)}} \ge \gamma \|\varphi\|_{H^1(\Omega)} \ \text{ for all } \ \varphi \in X(\Omega). \tag{2.10}$$

**3. The discrete approximation and the main result.** First, we do a truncation approximation to the nonlocal boundary operator from the corresponding infinite series expansion; then the finite element formulation of (2.8) is presented by using the proposed approximation of boundary operators. Based on the introduced notation, we will finally give the main result on the a posteriori error estimate for the case of the TE polarization.

**3.1. The discrete problem.** Let $\mathcal{M}_h$ be a regular triangulation of the domain $\Omega$. Every triangle $T \in \mathcal{M}_h$ is considered as closed. To define a finite element space whose functions are quasic-periodic in the $x_1$ direction, we require that if $(0, z)$ is a node on the left boundary, then $(L, z)$ also is a node on the right boundary, and vice versa. Let $V_h \subset X(\Omega)$ denote a conforming finite element space, that is,

$$V_h := \{v_h \in C(\overline{\Omega}) : v_h|_K \in P_p(K) \text{ for all } K \in \mathcal{M}_h,\\ v_h(0, x_3) = e^{-i\alpha L} v_h(L, x_3) \text{ for } b_2 < x_3 < b_1\},$$

where $p$ is a positive integer and $P_p(K)$ is the set of polynomials of degrees $\leq p$. The finite element approximation to problem (2.8) reads as follows: Find $u_h \in V_h$ such that

$$a_{TE}(u_h, \psi_h) = -\int_{\Gamma_1} 2i\beta u_I \overline{\psi}_h dx_1 \quad \text{for all} \quad \psi_h \in V_h. \tag{3.1}$$

In the above formulation, the DtN operators $T^{(j)}$ given by (2.5) are defined by an infinite series which is unrealistic in actual calculations; thus it is necessary to truncate the nonlocal operator by taking sufficiently many terms of the expansions so as to attain our feasible algorithm. Note from (2.6) that if $k_j$ is real, then the $n$'s satisfying $\frac{2\pi}{L}\left(n + \frac{\alpha L}{2\pi}\right) < k_j$ correspond to outgoing modes which should be contained in the truncated DtN mappings. Denote by $n_\alpha := \frac{\alpha L}{2\pi}$. We truncate the DtN mappings $T^{(j)}$ as follows:

$$T^{(j,N_j)} f(x_1) = \sum_{|n+n_\alpha| \leq N_j} i\beta_j^n f_\alpha^{(n)} e^{i(\alpha_n + \alpha)x_1}, \quad 0 \leq x_1 \leq L, \quad j = 1, 2. \tag{3.2}$$

Here $N_j$ is usually an integer greater than $\frac{k_j L}{2\pi}$ if $k_j$ is real. We are now ready to define the truncated finite element formulation which leads to the discrete schemes to (2.8): Find $u_h^N \in V_h$ such that

$$a_{TE}^N(u_h^N, \psi_h) = -\int_{\Gamma_1} 2i\beta u_I \overline{\psi}_h dx_1 \quad \text{for all} \quad \psi_h \in V_h, \tag{3.3}$$

where the sesquilinear form $a_{TE}^N : X(\Omega) \times X(\Omega) \to C$ is defined as follows:

$$a_{TE}^N(\varphi, \psi) = \int_\Omega (\nabla\varphi \cdot \nabla\overline{\psi} - k^2(x)\varphi\overline{\psi})dx - \sum_{j=1}^{2} \int_{\Gamma_j} (T^{(j,N_j)}\varphi)\overline{\psi} dx_1. \tag{3.4}$$

For sufficiently large $N_j$ and sufficiently small $h$, the discrete inf-sup condition of (3.4) may be established by a general argument of Schatz [32]. Based on the important condition and the general theory in [4], the existence and uniqueness of the solution of problem (3.3) may be obtained. In fact, one can also see Bao [6] for the well-posedness of the problem (3.3). In this paper, our research interest is focused on a posteriori error estimates and the associated adaptive algorithm. Thus we assume that the discrete problem (3.3) has a unique solution $u_h^N \in V_h$.

**3.2. The main result.** For any $T \in \mathcal{M}_h$, denote by $h_T$ its diameter. Let $\mathcal{B}_h^j$ denote the set of all the sides that lie on $\Gamma_j$, $j = 1, 2$, and let $\mathcal{B}_h$ denote the set of all the sides except $\mathcal{B}_h^j$ in $\Omega$. For any $e \in \mathcal{B}_h$ or $e \in \mathcal{B}_h^j$, $h_e$ stands for its length. For any interior side $e \in \mathcal{B}_h$ which is the common side of $T_1$ and $T_2 \in \mathcal{M}_h$, we define the jump residual across $e$ as

$$J_e = -(\nabla u_h^N|_{T_1} \cdot n_1 + \nabla u_h^N|_{T_2} \cdot n_2), \tag{3.5}$$

where $n_j$ is the unit outward normal vector to the boundary of $T_j$, $j = 1, 2$.

Define $\Gamma_{left} = \{x_1 = 0, b_2 < x_3 < b_1\}$ and $\Gamma_{right} = \{x_1 = L, b_2 < x_3 < b_1\}$. If $e = \Gamma_{left} \bigcap \partial T$ for some element $T \in \mathcal{M}_h$ and $e'$ be a corresponding side on $\Gamma_{right}$ which also belongs to some element $T'$, then we define the jump residual as

$$J_e = \frac{\partial}{\partial x_1}\left(u_h^N|_T\right) - e^{-i\alpha L}\frac{\partial}{\partial x_1}\left(u_h^N|_{T'}\right), \tag{3.6}$$

$$J_{e'} = e^{i\alpha L}\frac{\partial}{\partial x_1}\left(u_h^N|_T\right) - \frac{\partial}{\partial x_1}\left(u_h^N|_{T'}\right). \tag{3.7}$$

For any $e \in \mathcal{B}_h^1$ and $e' \in \mathcal{B}_h^2$, define the jump residual as follows:

$$J_e = 2\Big(T^{(1,N_1)}u_h^N(x_1, b_1) - \frac{\partial u_h^N}{\partial x_3}(x_1, b_1) - 2i\beta e^{i\alpha x_1 - i\beta b_1}\Big), \tag{3.8}$$

$$J_{e'} = 2\Big(T^{(2,N_2)}u_h^N(x_1, b_2) + \frac{\partial u_h^N}{\partial x_3}(x_1, b_2)\Big). \tag{3.9}$$

For any $T \in \mathcal{M}_h$, denote by $\eta_T$ the local error estimator, which is defined by

$$\eta_T = h_T\|\mathcal{L}u_h^N\|_{L^2(T)} + \left(\frac{1}{2}\sum_{e\subset\partial T} h_e\|J_e\|^2_{L^2(e)}\right)^{\frac{1}{2}}, \tag{3.10}$$

where the residual operator $\mathcal{L}=\triangle + k^2(x)$.

We now state the main conclusion, which will be importantly the theoretical basis of numerical calculation listed in section 5.

THEOREM 3.1. *Let $u$ and $u_h^N$ denote the solutions of* (2.8) *and* (3.3)*, respectively. Then there exist two integers $N_{j0}, j = 1, 2$, independent of $h$ and satisfying $(2\pi N_{j0}/L)^2 > \operatorname{Re} k_j^2$ such that for $N_j \geq N_{j0}$ the following a posteriori error estimate holds:*

$$\|u - u_h^N\|_{H^1(\Omega)} \leq \widetilde{C}\left(\left(\sum_{T\in\mathcal{M}_h}\eta_T^2\right)^{\frac{1}{2}} + \sum_{j=1}^{2} e^{-|b_j - b_j'|\sqrt{(2\pi N_j/L)^2 - \operatorname{Re} k_j^2}}\right),$$

*where the constant $\widetilde{C}$ is independent of $h$, $N_1$, and $N_2$ (cf.* (4.14)*).*

We remark that the first term on the right-hand side of the above estimate indicates the finite element discretization error, while the second term demonstrates the truncation error of the transparent boundary operators. Clearly the second term is exponentially decaying with respect to $N_j$ and the distances from $\Gamma_j, j = 1, 2$, to the grating.

**4. The a posteriori error analysis.** In this section we prove the a posteriori error estimate for the case of TE polarization in Theorem 3.1 by a duality argument similar to that (or the so-called Schatz argument) for the a priori error estimates for indefinite problems [32, 6]. Since the analysis on the TM polarization is similar, the corresponding results are directly stated without the relevant proof. In the end of this section, we will give the parallel results in the $\operatorname{Im}\varepsilon_2 > 0$ case for the sake of general consideration.

Denote the error by $\xi := u - u_h^N$. Introduce the following dual problem to the original scattering problem: Find $w \in X(\Omega)$ such that

$$a_{TE}(v, w) = (v, \xi) \quad \text{for all} \;\; v \in X(\Omega). \tag{4.1}$$

Some simple calculations yield that $w$ is the weak solution of the problem

$$\triangle w + \overline{k}^2 w = -\xi \;\; \text{in} \;\; \Omega, \tag{4.2}$$

$$\frac{\partial w}{\partial n} - T^{(j,*)} w = 0 \quad \text{on} \;\; \Gamma_j, \tag{4.3}$$

where $j = 1, 2$, $\overline{k}^2$ is a conjugate complex number of $k^2(x)$, and the dual operators take the following form:

$$T^{(j,*)} v = -\sum_{n\in\mathbb{Z}} i\, \overline{\beta_j^n} v_\alpha^{(n)} e^{i(\alpha_n+\alpha)x_1}.$$

We remark that the existence of solutions for (4.1) can be obtained from the Fredholm theory and the related proof of [7]. Here we shall not elaborate on this issue, and we assume this problem has a unique (weak) solution. Then we have

$$\|w\|_{H^1(\Omega)} \le C_0 \, \|\xi\|_{L^2(\Omega)} . \tag{4.4}$$

Note that, unlike the duality argument for a priori error estimates, we do not assume the $H^2$ regularity of the dual problem.

**4.1. Error representation formulae.** The following lemma gives some relations on the error $\xi = u - u_h^N$, which is the start point for the a posteriori error analysis.

LEMMA 4.1. *Let $u, u_h^N$, and $w$ be the solutions to problems* (2.8), (3.3), *and* (4.1), *respectively. Then*

$$\|\xi\|_{H^1(\Omega)}^2 = \operatorname{Re}\left( a_{TE}(\xi,\xi) + \sum_{j=1}^{2} \int_{\Gamma_j} \big(T^{(j)} - T^{(j,N_j)}\big)\xi\overline{\xi} dx_1 \right) + \sum_{j=1}^{2} \operatorname{Re} \int_{\Gamma_j} T^{(j,N_j)} \xi\overline{\xi} dx_1 + \int_\Omega \operatorname{Re}\big(k^2(x)+1\big)\,|\xi|^2\, dx, \tag{4.5}$$

$$\|\xi\|_{L^2(\Omega)}^2 = a_{TE}(\xi, w) + \sum_{j=1}^{2} \int_{\Gamma_j} (T^{(j)} - T^{(j,N_j)})\xi\overline{w} dx_1 - \sum_{j=1}^{2} \int_{\Gamma_j} \big(T^{(j)} - T^{(j,N_j)}\big)\xi\overline{w} dx_1, \tag{4.6}$$

$$a_{TE}(\xi,\psi) + \sum_{j=1}^{2} \int_{\Gamma_j} \big(T^{(j)} - T^{(j,N_j)}\big)\xi\overline{\psi} dx_1 = -\int_{\Gamma_1} 2i\beta u_I (\overline{\psi - \psi_h}) dx_1 - a_{TE}^N(u_h^N, \psi - \psi_h) + \sum_{j=1}^{2} \int_{\Gamma_j} \big(T^{(j)} - T^{(j,N_j)}\big) u\overline{\psi} dx_1 \quad \text{for all } \psi_h \in V_h, \psi \in X(\Omega). \tag{4.7}$$

*Proof.* Equation (4.5) follows from the definition of $a_{TE}$ in (2.9) and (4.6) follows by taking $v = \xi$ in (4.1). It remains to prove (4.7). From (2.8) and (3.3),

$$
\begin{aligned}
&a_{TE}(\xi, \psi) \\
&\quad = a_{TE}(u - u_h^N, \psi - \psi_h) + a_{TE}(u - u_h^N, \psi_h) \\
&\quad = -\int_{\Gamma_1} 2i\beta u_I(\overline{\psi - \psi_h})dx_1 - a_{TE}^N(u_h^N, \psi - \psi_h) \\
&\qquad + \big(a_{TE}^N(u_h^N, \psi - \psi_h) - a_{TE}(u_h^N, \psi - \psi_h)\big) + a_{TE}(u - u_h^N, \psi_h) \\
&\quad = -\int_{\Gamma_1} 2i\beta u_I(\overline{\psi - \psi_h})dx_1 - a_{TE}^N(u_h^N, \psi - \psi_h) \\
&\qquad + \big(a_{TE}^N(u_h^N, \psi - \psi_h) - a_{TE}(u_h^N, \psi - \psi_h)\big) + \big(a_{TE}^N(u_h^N, \psi_h) - a_{TE}(u_h^N, \psi_h)\big) \\
&\quad = -\int_{\Gamma_1} 2i\beta u_I(\overline{\psi - \psi_h})dx_1 - a_{TE}^N(u_h^N, \psi - \psi_h) + \big(a_{TE}^N(u_h^N, \psi) - a_{TE}(u_h^N, \psi)\big) \\
&\quad = -\int_{\Gamma_1} 2i\beta u_I(\overline{\psi - \psi_h})dx_1 - a_{TE}^N(u_h^N, \psi - \psi_h) \\
&\qquad - \sum_{j=1}^{2}\int_{\Gamma_j}(T^{(j)} - T^{(j,N_j)})\xi\overline{\psi}dx_1 + \sum_{j=1}^{2}\int_{\Gamma_j}(T^{(j)} - T^{(j,N_j)})u\overline{\psi}dx_1,
\end{aligned}
$$

which implies (4.7). This completes the proof of the lemma. ∎

In order to prove Theorem 3.1, we need to estimate (4.7) and the last term in (4.6).

**4.2. Estimation of (4.7).** We first state two simple lemmas. The following trace property, found in Chen and Wu [21], will be useful.

LEMMA 4.2. *For any $\psi \in X(\Omega)$, we have*

$$\|\psi\|_{H^{1/2}(\Gamma_j)} \le \widehat{C}\|\psi\|_{H^1(\Omega)}$$

*with $\widehat{C} = \sqrt{1 + (b_1 - b_2)^{-1}}$ and $j = 1, 2$. Here if $\psi(x_1, b_j) = \sum_{n\in\mathbb{Z}} \psi_\alpha^{(n)}(b_j)e^{i(\alpha_n+\alpha)x_1}$ on $\Gamma_j$,*

$$\|\psi\|_{H^{1/2}(\Gamma_j)} = \left(L\sum_{n\in\mathbb{Z}}\left(1 + |\alpha_n + \alpha|^2\right)^{1/2}|\psi_\alpha^{(n)}(b_j)|^2\right)^{1/2}.$$

The following lemma is crucial in deriving the truncation error (cf. [6]).

LEMMA 4.3. *Let $u$ be the solution to (2.8) and $u_\alpha^{(n)}(x_3) := \frac{1}{L}\int_0^L u(x_1, x_3)$ $e^{-i(\alpha_n+\alpha)x_1}dx_1$. Suppose that $(\alpha_n + \alpha)^2 \ge \operatorname{Re} k_j^2$. Then*

$$|u_\alpha^{(n)}(b_j)| \le |u_\alpha^{(n)}(b_j')|e^{-|b_j - b_j'|\sqrt{(\alpha_n+\alpha)^2 - \operatorname{Re} k_j^2}} \quad \textit{for} \quad j = 1, 2.$$

*Proof.* Clearly, the Rayleigh expansions (see, e.g., [21, (2.4) and (2.5)]) hold for $x_3 \ge b_1'$ and $x_3 \le b_2'$, so we have

$$|u_\alpha^{(n)}(b_j)| = |u_\alpha^{(n)}(b_j')|e^{-|b_j - b_j'|\operatorname{Im}\beta_j^n} \quad \text{for} \quad j = 1, 2,$$

where $\operatorname{Im}\beta_j^n$ can be bounded by using (2.6) and the fact that $\operatorname{Im}\left((\beta_j^n)^2\right) = \operatorname{Im} k_j^2 \geq 0$ as follows:

$$\begin{aligned}\operatorname{Im}\beta_j^n &= \frac{1}{\sqrt{2}}\left(\sqrt{\operatorname{Im}\left((\beta_j^n)^2\right)^2 + \operatorname{Re}\left((\beta_j^n)^2\right)^2} - \operatorname{Re}\left((\beta_j^n)^2\right)\right)^{\frac{1}{2}} \\ &\geq \sqrt{-\operatorname{Re}\left((\beta_j^n)^2\right)} = \sqrt{(\alpha_n+\alpha)^2 - \operatorname{Re} k_j^2}.\end{aligned}$$

This completes the proof of the lemma. ∎

Next we estimate (4.7).

LEMMA 4.4. *There exist integers $N_{j1}$ independent of $h$ and satisfying $(2\pi N_{j1}/L)^2 > \operatorname{Re} k_j^2$, $j = 1, 2$, such that for any $N_j \geq N_{j1}$ and $\psi \in X(\Omega)$ we have*

$$\begin{aligned}&\left| a_{TE}(\xi,\psi) + \sum_{j=1}^{2}\int_{\Gamma_j}(T^{(j)} - T^{(j,N_j)})\xi\overline{\psi}dx_1\right| \\ &\leq C_1\left(\left(\sum_{T\in\mathcal{M}_h}\eta_T^2\right)^{1/2} + \sum_{j=1}^{2} e^{-|b_j-b_j'|\sqrt{(2\pi N_j/L)^2 - \operatorname{Re} k_j^2}}\right)\|\psi\|_{H^1(\Omega)},\end{aligned}$$

*where $C_1$ is a constant independent of $h$ and $N_j$.*

*Proof.* Denote by

$$\begin{aligned}\mathbb{J}^1 &:= -\int_{\Gamma_1} 2i\beta u_I(\overline{\psi - \psi_h})dx_1 - a_{TE}^N(u_h^N, \psi - \psi_h), \\ \mathbb{J}^2 &:= \sum_{j=1}^{2}\int_{\Gamma_j}(T^{(j)} - T^{(j,N_j)})u\overline{\psi}dx_1.\end{aligned}$$

Then from (4.7),

$$a_{TE}(\xi,\psi) + \sum_{j=1}^{2}\int_{\Gamma_j}(T^{(j)} - T^{(j,N_j)})\xi\overline{\psi}dx_1 = \mathbb{J}^1 + \mathbb{J}^2.$$

By using the definition of the sesquilinear form (3.4), $\mathbb{J}^1$ can be rewritten as follows:

$$\begin{aligned}\mathbb{J}^1 = \sum_{T\in\mathcal{M}_h}&\Bigg(-\int_T\left(\nabla u_h^N\cdot\nabla(\overline{\psi-\psi_h}) - k^2(x)u_h^N(\overline{\psi-\psi_h})\right)dx \\ &\quad + \sum_{j=1}^{2}\sum_{e\subset\partial T\bigcap\Gamma_j}\int_e T^{(j,N_j)}u_h^N(\overline{\psi-\psi_h})dx_1\Bigg) \\ &+ \sum_{T\in\mathcal{M}_h}\left(-\sum_{e\subset\partial T\bigcap\Gamma_1}\int_e 2i\beta e^{i\alpha x_1 - i\beta b_1}(\overline{\psi-\psi_h})dx_1\right).\end{aligned}$$

Integration by parts yields

$$\begin{aligned}\mathbb{J}^1 = \sum_{T\in\mathcal{M}_h}&\Bigg(\int_T\left(\triangle u_h^N + k^2(x)u_h^N\right)(\overline{\psi-\psi_h})dx - \sum_{e\subset\partial T}\int_e\nabla u_h^N\cdot n(\overline{\psi-\psi_h})ds \\ &+ \sum_{e\subset\partial T\bigcap\Gamma_1}\int_e\left(T^{(1,N_1)} - 2i\beta e^{i\alpha x_1 - i\beta b_1}\right)u_h^N(\overline{\psi-\psi_h})dx_1 \\ &+ \sum_{e\subset\partial T\bigcap\Gamma_2}\int_e T^{(2,N_2)}u_h^N(\overline{\psi-\psi_h})dx_1\Bigg).\end{aligned}$$

Combining with (3.5)–(3.9) implies

$$\mathbb{J}^1 = \sum_{T\in\mathcal{M}_h} \left( \int_T \mathcal{L}u_h^N \overline{(\psi - \psi_h)} dx + \sum_{e\subset\partial T} \frac{1}{2} \int_e J_e \overline{(\psi - \psi_h)} ds \right).$$

Now we take $\psi_h = \Pi_h \psi \in V_h$. Here $\Pi_h$ is the Scott–Zhang interpolation operator adopted in Chen and Wu [21] first introduced in Scott–Zhang [33], which has the following interpolation estimates:

$$\|v - \Pi_h v\|_{L^2(T)} \le C h_T \|\nabla v\|_{L^2(\widetilde{T})}, \quad \|v - \Pi_h v\|_{L^2(e)} \le C h_e^{1/2} \|\nabla v\|_{L^2(\widetilde{e})},$$

where $\widetilde{T}$ and $\widetilde{e}$ are the union of all the elements in $\mathcal{M}_h$ having nonempty intersection with the element $T$ and the side $e$, respectively. It follows from the Cauchy–Schwarz inequality and the interpolation estimates that

$$\begin{aligned} |\mathbb{J}^1| \le C \sum_{T\in\mathcal{M}_h} \Bigg( & h_T \|\mathcal{L}u_h^N\|_{L^2(T)} \|\nabla\psi\|_{L^2(\widetilde{T})} \\ & + \sum_{e\subset\partial T} h_e^{1/2} \|J_e\|_{L^2(e)} \|\nabla\psi\|_{L^2(\widetilde{e})} \Bigg) \\ \le C \sum_{T\in\mathcal{M}_h} & \left( h_T \|\mathcal{L}u_h^N\|_{L^2(T)} + \left( \sum_{e\subset\partial T} h_e \|J_e\|^2_{L^2(e)} \right)^{1/2} \right) \|\psi\|_{H^1(\Omega)}, \end{aligned}$$

which leads to

$$|\mathbb{J}^1| \le C \left( \sum_{T\in\mathcal{M}_h} \eta_T^2 \right)^{1/2} \|\psi\|_{H^1(\Omega)}. \tag{4.8}$$

It remains to estimate $\mathbb{J}^2$. By noting that $N_{j1}$ is a sufficiently large integer, such that $(2\pi N_{j1}/L)^2 > \operatorname{Re} k_j^2$, combining with Lemmas 4.3 and 4.2, then for $N_j \ge N_{j1}$ we can get

$$\begin{aligned} |\mathbb{J}^2| = & \left| \sum_{j=1}^{2} L \sum_{|n+n_\alpha|>N_j} i\beta_j^n u_\alpha^{(n)}(b_j) \overline{\psi_\alpha^{(n)}(b_j)} \right| \\ \le & \sum_{j=1}^{2} L \sum_{|n+n_\alpha|>N_j} |\beta_j^n| e^{-|b_j - b_j'|\sqrt{(\alpha_n+\alpha)^2 - \operatorname{Re} k_j^2}} |u_\alpha^{(n)}(b_j')| |\psi_\alpha^{(n)}(b_j)| \\ \le & \sum_{j=1}^{2} e^{-|b_j - b_j'|\sqrt{(2\pi N_j/L)^2 - \operatorname{Re} k_j^2}} \left( L \sum_{|n+n_\alpha|>N_j} |\beta_j^n| |u_\alpha^{(n)}(b_j')|^2 \right)^{1/2} \\ & \times \left( L \sum_{|n+n_\alpha|>N_j} |\beta_j^n| |\psi_\alpha^{(n)}(b_j)|^2 \right)^{1/2} \\ \le & C \sum_{j=1}^{2} e^{-|b_j - b_j'|\sqrt{(2\pi N_j/L)^2 - \operatorname{Re} k_j^2}} \|u\|_{H^{1/2}(\Gamma_j')} \cdot \|\psi\|_{H^{1/2}(\Gamma_j)} \end{aligned}$$

$$\leq C\sum_{j=1}^{2} e^{-|b_j-b_j'|\sqrt{(2\pi N_j/L)^2-\mathrm{Re}\,k_j^2}}\|u\|_{H^1(\Omega)}\cdot\|\psi\|_{H^1(\Omega)}$$

Further, using (2.8) and (2.10), we can obtain

$$\|u\|_{H^1(\Omega)} \leq \frac{1}{\gamma}\sup_{0\neq\psi\in X(\Omega)}\frac{|a_{TE}(u,\psi)|}{\|\psi\|_{H^1(\Omega)}} \leq \frac{2|\beta|\widehat{C}}{\gamma}\|u_I\|_{L^2(\Gamma_1)}.$$

Therefore,

$$|\mathbb{J}^2| \leq C\sum_{j=1}^{2} e^{-|b_j-b_j'|\sqrt{(2\pi N_j/L)^2-\mathrm{Re}\,k_j^2}}\|u_I\|_{L^2(\Gamma_1)}\cdot\|\psi\|_{H^1(\Omega)}. \tag{4.9}$$

This proof follows from (4.8) and (4.9). □

**4.3. Approximation property of the truncated DtN operator.** In this subsection we estimate the last term in (4.6).

LEMMA 4.5. *For the solution $w$ of* (4.1)*, there exists an integer $N_{j2}$ independent of $h$ and satisfying $(2\pi N_{j2}/L)^2 > \mathrm{Re}\,k_j^2$, $j = 1, 2$, such that for $N_j \geq N_{j2}$, we have the following estimate:*

$$\left|\int_{\Gamma_j}(T^{(j)} - T^{(j,N_j)})\xi\overline{w}dx_1\right| \leq C_2 N_j^{-2}\|\xi\|^2_{H^1(\Omega)},$$

*where $C_2$ is a constant independent of $h$ and $N_j$ (cf.* (4.13)*).*

*Proof.* Since the parallel result for the case in which $j = 2$ can be derived similarly, we can only prove the inequality for the case in which $j = 1$. From the Cauchy–Schwarz inequality and Lemma 4.2 we easily deduce that

$$\begin{aligned}
\left|\int_{\Gamma_1}(T^{(1)} - T^{(1,N_1)})\xi\overline{w}dx_1\right| &\leq L\sum_{|n+n_\alpha|>N_1}|\beta_1^n||\xi_\alpha^{(n)}(b_1)||w_\alpha^{(n)}(b_1)| \\
&= L\sum_{|n+n_\alpha|>N_1}(|\alpha_n+\alpha||\beta_1^n|^3)^{-1/2}|\alpha_n+\alpha|^{1/2}|\xi_\alpha^{(n)}(b_1)||\beta_1^n|^{5/2}|w_\alpha^{(n)}(b_1)| \\
&\leq \max_{|n+n_\alpha|>N_1}(|\alpha_n+\alpha||\beta_1^n|^3)^{-1/2}\left(L\sum_{|n+n_\alpha|>N_1}|\alpha_n+\alpha||\xi_\alpha^{(n)}(b_1)|^2\right)^{1/2} \\
&\quad\times\left(L\sum_{|n+n_\alpha|>N_1}|\beta_1^n|^5|w_\alpha^{(n)}(b_1)|^2\right)^{1/2} \\
&\leq C(N_1N_1^3)^{-1/2}\|\xi\|_{H^{1/2}(\Gamma_1)}\left(L\sum_{|n+n_\alpha|>N_1}|\beta_1^n|^5|w_\alpha^{(n)}(b_1)|^2\right)^{1/2} \\
&\leq CN_1^{-2}\|\xi\|_{H^1(\Omega)}\left(L\sum_{|n+n_\alpha|>N_1}|\beta_1^n|^5|w_\alpha^{(n)}(b_1)|^2\right)^{1/2}.
\end{aligned} \tag{4.10}$$

Next, in order to estimate $w_\alpha^{(n)}(b_1)$ we consider the dual problem (4.2)–(4.3) in the following domain near $\Gamma_1$:

$$\widetilde{\Omega}^1 = \{0 < x_1 < L,\ b_1' < x_3 < b_1\}.$$

Substituting the series expansion of $w$ into the dual value problem, we can find the boundary value problem of the ordinary differential equation as follows:

$$\tag{4.11} \begin{cases} (w_\alpha^{(n)})''(x_3) - |\beta_1^n|^2 w_\alpha^{(n)}(x_3) = -\xi_\alpha^{(n)}(x_3) \quad \text{in} \quad (b_1', b_1), \\ (w_\alpha^{(n)})'(b_1) + |\beta_1^n| w_\alpha^{(n)}(b_1) = 0, \\ w_\alpha^{(n)}(b_1') = w_\alpha^{(n)}(b_1'), \end{cases}$$

where the absolute value of $n + n_\alpha$ is required to be greater than $N_1$.

According to the general theory of ordinary differential equations, the solution to (4.11) can be expressed as

$$\begin{aligned} w_\alpha^{(n)}(x_3) =& \frac{1}{2|\beta_1^n|}\bigg( - \int_{b_1}^{x_3} e^{|\beta_1^n|(x_3 - s)} \xi_\alpha^{(n)}(s) ds + \int_{b_1'}^{x_3} e^{|\beta_1^n|(s - x_3)} \xi_\alpha^{(n)}(s) ds \\ &- \int_{b_1'}^{b_1} e^{|\beta_1^n|(2b_1' - x_3 - s)} \xi_\alpha^{(n)}(s) ds + 2|\beta_1^n| e^{|\beta_1^n|(b_1' - x_3)} w_\alpha^{(n)}(b_1') \bigg), \end{aligned}$$

which leads to

$$\begin{aligned} |w_\alpha^{(n)}(b_1)| \le& \frac{1}{2|\beta_1^n|}\bigg( \int_{b_1'}^{b_1} \left(e^{|\beta_1^n|(s-b_1)} - e^{|\beta_1^n|(2b_1' - s - b_1)}\right) \cdot |\xi_\alpha^{(n)}(s)| ds \bigg) \\ &+ e^{-d|\beta_1^n|} |w_\alpha^{(n)}(b_1')| \\ \le& \frac{(1 - e^{-d|\beta_1^n|})^2}{2|\beta_1^n|^2} \|\xi_\alpha^{(n)}\|_{L^\infty([b_1', b_1])} + e^{-d|\beta_1^n|} |w_\alpha^{(n)}(b_1')| \\ \le& \frac{1}{2|\beta_1^n|^2} \|\xi_\alpha^{(n)}\|_{L^\infty([b_1', b_1])} + e^{-d|\beta_1^n|} |w_\alpha^{(n)}(b_1')|, \end{aligned}$$

where $d := b_1 - b_1'$. For any $s \in (b_1', b_1)$, without loss of generality, it is assumed that s is closer to the left endpoint $b_1$; than the right endpoint $b_1$; then we have $b_1 - s \ge d/2$. Thus

$$\begin{aligned} |\xi_\alpha^{(n)}(s)|^2 &= \frac{1}{b_1 - s} \int_{b_1}^{s} \left((b_1 - t)|\xi_\alpha^{(n)}(t)|^2\right)' dt \\ &= \frac{1}{b_1 - s} \int_{b_1}^{s} \Big[ - |\xi_\alpha^{(n)}(t)|^2 + 2(b_1 - t) \operatorname{Re}\left((\xi_\alpha^{(n)}(t))' \overline{\xi_\alpha^{(n)}(t)}\right) \Big] dt \\ &\le \frac{1}{b_1 - s} \int_{b_1'}^{b_1} |\xi_\alpha^{(n)}(t)|^2 dt + 2 \int_{b_1'}^{b_1} |\xi_\alpha^{(n)}(t)| |(\xi_\alpha^{(n)}(t))'| dt, \end{aligned}$$

which implies that

$$\|\xi_\alpha^{(n)}\|^2_{L^\infty([b_1', b_1])} \le \frac{2}{d} \|\xi_\alpha^{(n)}\|^2_{L^2([b_1', b_1])} + 2\|\xi_\alpha^{(n)}\|_{L^2([b_1', b_1])} \|(\xi_\alpha^{(n)})'\|_{L^2([b_1', b_1])}.$$

Therefore from the Young inequality,

$$\begin{aligned} &L \sum_{|n + n_\alpha| > N_1} |\beta_1^n|^5 |w_\alpha^{(n)}(b_1)|^2 \\ &\le L \sum_{|n + n_\alpha| > N_1} \left( \frac{|\beta_1^n|}{2} \|\xi_\alpha^{(n)}\|^2_{L^\infty([b_1', b_1])} + 2|\beta_1^n|^5 e^{-2d|\beta_1^n|} |w_\alpha^{(n)}(b_1')|^2 \right) \\ &\le CL \sum_{|n + n_\alpha| > N_1} \left( \left((d|\beta_1^n|)^{-1} + 1\right) |\beta_1^n|^2 \|\xi_\alpha^{(n)}\|^2_{L^2([b_1', b_1])} + \|(\xi_\alpha^{(n)})'\|^2_{L^2([b_1', b_1])} \right) \end{aligned}$$

$$+ 2L \max_{|n+n_\alpha|>N_1} \left(|\beta_1^n|^4 e^{-2d|\beta_1^n|}\right) \sum_{|n+n_\alpha|>N_1} |\beta_1^n||w_\alpha^{(n)}(b_1')|^2$$
$$:= \mathrm{I} + \mathrm{II}.$$

Note that

$$|\beta_1^n| \leq |\alpha_n + \alpha| \leq (1 + |\alpha_n + \alpha|^2)^{1/2}, \quad N_1 \leq C|\beta_1^n|, \quad \text{for } |n + n_\alpha| > N_1$$

and that

$$\|\xi\|^2_{H^1(\Omega)} = L \sum_{n\in\mathbb{Z}} \int_{b_2}^{b_1} \left((1 + |\alpha_n + \alpha|^2)|\xi_\alpha^{(n)}(x_3)|^2 + \left|\frac{d}{dx_3}\xi_\alpha^{(n)}(x_3)\right|^2\right) dx_3.$$

We have

$$\mathrm{I} \leq C\big((N_1 d)^{-1} + 1\big)\|\xi\|^2_{H^1(\Omega)}.$$

On the other hand, from Lemma 4.2, (4.4), and the fact that the function $t^4 e^{-t}$ is bounded on $(0, +\infty)$,

$$\mathrm{II} \leq C d^{-4} L \max_{|n+n_\alpha|>N_1} \left(|2d\beta_1^n|^4 e^{-2d|\beta_1^n|}\right) \|w\|^2_{H^{1/2}(\Gamma_1')} \leq C d^{-4} \|\xi\|^2_{L^2(\Omega)}.$$

Therefore,

$$L \sum_{|n+n_\alpha|>N_1} |\beta_1^n|^5 |w_\alpha^{(n)}(b_1)|^2 \leq C\big((N_1 d)^{-1} + 1 + d^{-4}\big)\|\xi\|^2_{H^1(\Omega)}. \tag{4.12}$$

Then the proof of Lemma 4.5 follows by combining (4.10) and (4.12) and setting

$$C_2 = C\big(1 + (N_1 d)^{-1/2} + d^{-2}\big). \tag{4.13}$$ ∎

**4.4. Proof of Theorem 3.1.** Next we prove Theorem 3.1. Let $N_j \geq \max(N_{j1}, N_{j2})$, $j = 1, 2$. First, from the definitions of $T^{(j,N_j)}$ and $\beta_j^n$ in (3.2) and (2.6),

$$\mathrm{Re} \int_{\Gamma_j} T^{(j,N_j)} \xi \overline{\xi} dx_1 = -L \sum_{|n+n_\alpha|\leq N_j} \mathrm{Im}(\beta_j^n)|\xi_\alpha^{(n)}|^2 \leq 0.$$

Therefore, from (4.5) and Lemma 4.4, we have

$$\|\xi\|^2_{H^1(\Omega)} \leq C_1 \left(\left(\sum_{T\in\mathcal{M}_h} \eta_T^2\right)^{1/2} + \sum_{j=1}^{2} e^{-|b_j - b_j'|\sqrt{(2\pi N_j/L)^2 - \mathrm{Re}\, k_j^2}}\right) \|\xi\|_{H^1(\Omega)}$$
$$+ C_3 \|\xi\|^2_{L^2(\Omega)}.$$

To estimate $\|\xi\|_{L^2(\Omega)}$, we use (4.6), Lemma 4.4, (4.4), and Lemma 4.5 to obtain

$$\|\xi\|^2_{L^2(\Omega)} \leq C_1 \left(\left(\sum_{T\in\mathcal{M}_h} \eta_T^2\right)^{1/2} + \sum_{j=1}^{2} e^{-|b_j - b_j'|\sqrt{(2\pi N_j/L)^2 - \mathrm{Re}\, k_j^2}}\right) C_0 \|\xi\|_{L^2(\Omega)}$$
$$+ C_2(N_1^{-2} + N_2^{-2})\|\xi\|^2_{H^1(\Omega)}.$$

By combining the above two estimates we have

$$\|\xi\|^2_{H^1(\Omega)} \le C_1(1+C_0C_3)\left(\left(\sum_{T\in\mathcal{M}_h}\eta_T^2\right)^{1/2}+\sum_{j=1}^{2}e^{-|b_j-b_j'|\sqrt{(2\pi N_j/L)^2-\operatorname{Re}k_j^2}}\right)\|\xi\|_{H^1(\Omega)} + C_2C_3(N_1^{-2}+N_2^{-2})\|\xi\|^2_{H^1(\Omega)}.$$

Choose integer $N_{j3}$ such that

$$C_2C_3N_{j3}^{-2} \le \frac{1}{4}, \quad j=1,2.$$

Then the proof of Theorem 3.1 follows by taking

$$N_{j0} = \max(N_{j1}, N_{j2}, N_{j3}), \quad \widetilde{C} = 2C_1(1+C_0C_3). \tag{4.14}$$

**4.5. TM polarization.** In this subsection, we present the parallel result for the grating problem (2.4) without providing detailed discussion.

The variational form in the TM polarization is as follows: Given incoming plane wave $u_I = e^{i\alpha x_1 - i\beta x_3}$, seek $u \in X(\Omega)$ such that

$$a_{TM}(u,\psi) = -k_1^{-2}\int_{\Gamma_1} 2i\beta u_I\overline{\psi}dx_1 \quad \text{for all } \psi \in X(\Omega), \tag{4.15}$$

where the sesquilinear form $a_{TM} : X(\Omega)\times X(\Omega) \to C$ is defined as

$$a_{TM}(\varphi,\psi) = \int_\Omega (k^{-2}(x)\nabla\varphi\cdot\nabla\overline{\psi} - \varphi\overline{\psi})dx - \sum_{j=1}^{2} k_j^{-2}\int_{\Gamma_j}(T^{(j)}\varphi)\overline{\psi}dx_1. \tag{4.16}$$

The truncated finite element formulation for (4.16) is as follows: Find $u_h^N \in V_h$ such that

$$a^N_{TM}(u_h^N,\psi_h) = -k_1^{-2}\int_{\Gamma_1} 2i\beta u_I\overline{\psi}_h dx_1 \quad \text{for all } \psi_h \in V_h, \tag{4.17}$$

where the sesquilinear form $a^N_{TM} : X(\Omega)\times X(\Omega) \to C$ is defined as

$$a^N_{TM}(\varphi,\psi) = \int_\Omega (k^{-2}(x)\nabla\varphi\cdot\nabla\overline{\psi} - \varphi\overline{\psi})dx - \sum_{j=1}^{2} k_j^{-2}\int_{\Gamma_j}(T^{(j,N_j)}\varphi)\overline{\psi}dx_1. \tag{4.18}$$

For any interior side $e \in \mathcal{B}_h$ which is the common side of $T_1$ and $T_2 \in \mathcal{M}_h$, we define the jump residual across $e$ as

$$J_e = -k^{-2}(x)(\nabla u_h^N|_{T_1}\cdot n_1 + \nabla u_h^N|_{T_2}\cdot n_2),$$

where $n_j$ is the unit outward normal vector to the boundary of $T_j$, $j=1,2$. If $e = \Gamma_{left}\bigcap \partial T$ for some element $T \in \mathcal{M}_h$ and $e'$ be a corresponding side on $\Gamma_{right}$ which also belongs to some element $T'$, then we define the jump residual as

$$J_e = \left(k^{-2}(x)\frac{\partial u_h^N}{\partial x_1}\right)\Big|_T - e^{-i\alpha L}\left(k^{-2}(x)\frac{\partial u_h^N}{\partial x_1}\right)\Big|_{T'},$$
$$J_{e'} = e^{i\alpha L}\left(k^{-2}(x)\frac{\partial u_h^N}{\partial x_1}\right)\Big|_T - \left(k^{-2}(x)\frac{\partial u_h^N}{\partial x_1}\right)\Big|_{T'}.$$

For any $e \in \mathcal{B}_h^1$ and $e' \in \mathcal{B}_h^2$, define the jump residual as follows:

$$J_e = 2k_1^{-2}\Big(T^{(1,N_1)}u_h^N(x_1,b_1) - \frac{\partial u_h^N}{\partial x_3}(x_1,b_1) - 2i\beta e^{i\alpha x_1 - i\beta b_1}\Big),$$
$$J_{e'} = 2k_2^{-2}\Big(T^{(2,N_2)}u_h^N(x_1,b_2) + \frac{\partial u_h^N}{\partial x_3}(x_1,b_2)\Big).$$

For any $T \in \mathcal{M}_h$, denote by $\eta_T$ the local error estimator, which is defined as follows:

$$\eta_T = h_T\|\mathcal{L}u_h^N\|_{L^2(T)} + \left(\frac{1}{2}\sum_{e\subset\partial T} h_e\|J_e\|_{L^2(e)}^2\right)^{\frac{1}{2}}, \tag{4.19}$$

where the element residual $\mathcal{L}u_h^N = \operatorname{div}(k^{-2}(x)\nabla u_h^N) + u_h^N$.

THEOREM 4.6. *Let $u$ and $u_h^N$ denote the solutions of* (4.15) *and* (4.17)*, respectively. Then there exist two integers $N_{j0}, j = 1, 2,$ independent of $h$ and satisfying $(2\pi N_{j0}/L)^2 > \operatorname{Re} k_j^2$, such that for $N_j \geq N_{j0}$ the following a posteriori error estimate holds:*

$$\|u - u_h^N\|_{H^1(\Omega)} \leq \widetilde{C}\left(\left(\sum_{T\in\mathcal{M}_h}\eta_T^2\right)^{1/2} + \sum_{j=1}^{2} e^{-|b_j - b_j'|\sqrt{(2\pi N_j/L)^2 - \operatorname{Re} k_j^2}}\right),$$

*where the constant $\widetilde{C}$ is independent of $h$, $N_1$, and $N_2$.*

## 5. Implementation and numerical examples.

In this section, we discuss the implementation of the adaptive finite element algorithm and present several numerical examples to demonstrate the competitive behavior of the proposed algorithm.

### 5.1. Adaptive algorithm.

Based on the a posteriori error estimate from Theorem 3.1 in the TE case and from Theorem 4.6 in the TM case, we use the PDE toolbox of MATLAB to implement the adaptive algorithm of the linear finite element formulation ($p = 1$). Theorem 3.1 shows us that the a posteriori error estimate consists of two parts: the finite element discretization error $\epsilon_h$ and the truncation error $\epsilon_{N_j}$ which depends on $N_j$, where

$$\epsilon_h = \left(\sum_{T\in\mathcal{M}_h}\eta_T^2\right)^{1/2}, \tag{5.1}$$

$$\epsilon_{N_j} = e^{-|b_j - b_j'|\sqrt{(2\pi N_j/L)^2 - \operatorname{Re} k_j^2}}, \quad j = 1, 2. \tag{5.2}$$

$\epsilon_h$ and $\epsilon_{N_j}$ should be changed respectively in the TM case according to Theorem 4.6. In our implementation, we can choose $b_j$, $b_j'$, and $N_j$ based on the restriction of $N_j$ and (5.2) such that the finite element discretization error is not contaminated by our truncation error, or more specifically, $\epsilon_{N_j}$ is required to be very small compared with $\epsilon_h$, say, $\epsilon_{N_j} \leq 10^{-8}$. Once $b_j$ and $N_j$ are fixed, we will use the adaptive strategy with Dörfler marking [24] to modify the grid according to the a posteriori error estimate (5.1). A good choice of $b_j$, $b_j'$, and $N_j$ in $\epsilon_{N_j}$ can greatly reduce the calculation amount. However, it is not clear to us how to determine the most optimal $N_j$, $b_j'$, and $b_j$ to improve the computational efficiency. For simplicity, in the following numerical experiments, $b_j'$ is chosen such that the grating structure lies exactly between $x_3 = b_j', j = 1, 2,$ and $b_j$ is determined through the relation that $|b_j - b_j'| = \min\{L/2, \lambda/2\}$,

Give $\kappa \in (0,1]$ and tolerance $TOL > 0$.
Choose $b_j$, $b'_j$ and $N_j$ such that $\epsilon_{N_j} \leq 10^{-8}$, $j = 1, 2$;
Generate an initial mesh $\mathcal{M}_h$ over $\Omega$ and compute error estimators;
While $\epsilon_h > TOL$ do
    choose $\widehat{\mathcal{M}}_h \subseteq \mathcal{M}_h$ according to the strategy $\eta_{\widehat{\mathcal{M}}_h} > \kappa \eta_h$
    refine all the elements in $\widehat{\mathcal{M}}_h$ and denote the obtained mesh still by $\mathcal{M}_h$,
    solve the discrete problem (3.3) or (4.17) on the new $\mathcal{M}_h$,
    compute the corresponding error estimators,
End while.

FIG. 3. *The adaptive algorithm.*

where $L$ is the period of the grating and $\lambda$ is the wavelength of the incident wave, and then $N_j$ is taken to be the smallest positive integer satisfying $\epsilon_{N_j} \leq 10^{-8}$. Recall that in section 3, we define, for any $T \in \mathcal{M}_h$, the local a posteriori error estimator as follows:

$$\eta_T = h_T \|\mathcal{L} u_h^N\|_{L^2(T)} + \left( \frac{1}{2} \sum_{e \subset \partial T} h_e \|J_e\|^2_{L^2(e)} \right)^{1/2}.$$

For any subset $\widehat{\mathcal{M}}_h$ of $\mathcal{M}_h$, denote by $\eta^2_{\widehat{\mathcal{M}}_h} = \sum_{T \in \widehat{\mathcal{M}}_h} \eta_T^2$ and let $\eta_h$ denote $\eta_{\mathcal{M}_h}$. In our algorithm, a factor of 0.15 is set on the error estimator as in the PDE toolbox of MATLAB. Now we briefly describe the adaptive algorithm in Figure 3.

**5.2. Numerical examples.** The following two examples are chosen to demonstrate the effectiveness of our algorithm. We also note that our adaptive algorithm with truncated DtN boundary condition is almost comparable in robustness to the adaptive PML algorithm [21, 9]. We set $\kappa = 0.5$ in the adaptive algorithm of Figure 3 and normalize the space variables so that $\mu = 1$.

*Example* 5.1. Consider the simplest grating structure, a straight line. Assume that a plane wave $u_{\mathrm{I}} = e^{i\alpha x_1 - i\beta x_3}$ is incident on the straight line $\{x_3 = 0\}$, which separates two homogeneous media whose dielectric coefficients are $\varepsilon_1$ and $\varepsilon_2$, respectively. The exact solution is known in [8]:

$$u = \begin{cases} u_{\mathrm{I}} + r e^{i\alpha x_1 + i\beta x_3} & \text{if } x_3 > 0, \\ t e^{i\alpha x_1 - i\widehat{\beta} x_3} & \text{if } x_3 < 0, \end{cases}$$

where $\widehat{\beta} = (k_2^2 - \alpha^2)^{1/2}$, $t = 2\beta/(\beta + \widehat{\beta})$, and $r = (\beta - \widehat{\beta})/(\beta + \widehat{\beta})$.

The parameters are chosen as $\varepsilon_1 = 1$, $\varepsilon_2 = (0.22 + 6.71i)^2$, $\theta = \pi/6$ and $\omega = \pi$. The domain is defined as $\Omega = [0, 2] \times [-1, 1]$. For our adaptive DtN algorithm, we choose $b_1 = 1$, $b_2 = -1$, $b'_1 = 0$, $b'_2 = 0$, $N_1 = 6$ and $N_2 = 4$.

Table 1 clearly shows the advantage of using adaptive mesh refinements. Moreover, it is shown that the a posteriori error estimate $\epsilon_h$ provides a rather good estimate for the priori error $e_h = \|\nabla(u - u_h^N)\|_{L^2(\Omega)}$. Figure 4 displays the curves of $\log e_h$ and $\log \epsilon_h$ versus $\log \mathrm{DoF}_h$ for our adaptive DtN method and the adaptive PML method, where $\mathrm{DoF}_h$ denotes the number of nodal points of the mesh $\mathcal{M}_h$ in domain $\Omega$ for our DtN method or in domain $D$ composed of $\Omega$ and the PML layers for the PML method. Figure 5 shows the error of the zero-order reflection efficiencies as a function of $\mathrm{DoF}_h$, respectively, for our adaptive DtN method and the adaptive PML method

Table 1

*Comparison of numerical results using adaptive (DtN) and uniform mesh refinements for Example* 5.1. *$DoF_h$ is the number of nodal points of mesh $\mathcal{M}_h$.*

| Adaptive mesh | | | | Uniform mesh | | | |
|---|---|---|---|---|---|---|---|
| $\mathrm{DoF}_h$ | $e_h$ | $\epsilon_h$ | $\epsilon_h/e_h$ | $\mathrm{DoF}_h$ | $e_h$ | $\epsilon_h$ | $\epsilon_h/e_h$ |
| 15 | 4.6817 | 3.85909 | 0.8243 | 15 | 4.6817 | 3.85909 | 0.8243 |
| 39 | 2.7703 | 2.79589 | 1.0092 | 49 | 2.6218 | 2.45519 | 0.9365 |
| 134 | 1.2586 | 1.31922 | 1.0481 | 177 | 1.3064 | 1.38364 | 1.0592 |
| 381 | 0.6480 | 0.724811 | 1.1186 | 673 | 0.6556 | 0.732994 | 1.1181 |
| 1505 | 0.3245 | 0.361331 | 1.1133 | 2625 | 0.3285 | 0.378147 | 1.1510 |
| 6257 | 0.1614 | 0.178102 | 1.1035 | 10369 | 0.1644 | 0.192346 | 1.1703 |
| 23500 | 0.0817 | 0.0901649 | 1.1040 | 41217 | 0.0822 | 0.0970503 | 1.1812 |

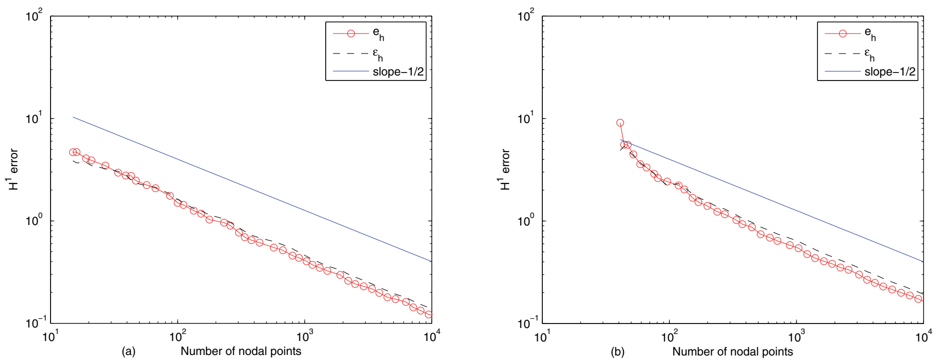


Fig. 4. *Quasi-optimality of the a priori and a posteriori error estimates for Example* 5.1. (a) *the adaptive DtN method;* (b) *the adaptive PML method.*

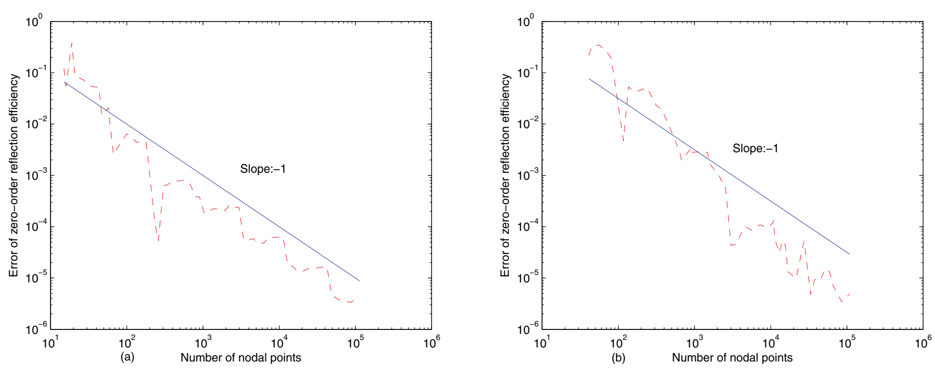


Fig. 5. *The error of the zero-order reflection efficiencies versus the number of nodal points for Example* 5.1. (a) *the adaptive DtN method;* (b) *the adaptive PML method.*

[21, 9], where the exact zero-order reflection efficiency is 0.9836391. It indicates that for the proposed two methods, the meshes and the associated numerical complexity are quasi-optimal: $\|\nabla(u-u_h^N)\|_{L^2(\Omega)} = \mathcal{O}(\mathrm{DoF}_h^{-1/2})$ is valid asymptotically, while the convergence rate of the zero-order reflection efficiencies is about $\mathcal{O}(\mathrm{DoF}_h^{-1})$.

Next we give further comparisons of the adaptive DtN method and the adaptive PML method in terms of accuracy and CPU time. For the adaptive PML method,

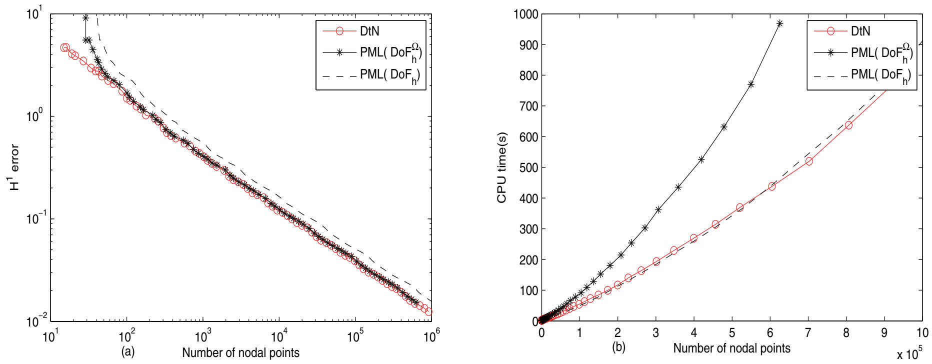


FIG. 6. (a) *The priori error versus the number of nodal points for Example* 5.1. (b) *CPU time versus the number of nodal points for Example* 5.1.

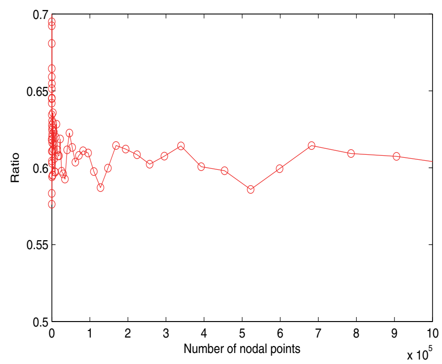


FIG. 7. *Ratio of* $DoF_h^\Omega/DoF_h$ *versus* $DoF_h$ *for Example* 5.1.

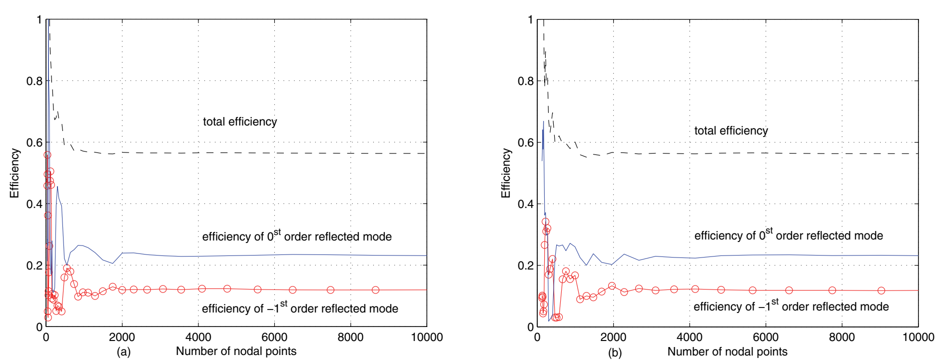


FIG. 8. *The zero-order reflection efficiency, the negative first-order reflection efficiency, and the total efficiency versus the number of nodal points for Example* 5.2. (a) *the adaptive DtN method;* (b) *the adaptive PML method.*

we denote by $\mathrm{DoF}_h$ the number of all nodal points in the mesh $\mathcal{M}_h$ and by $\mathrm{DoF}_h^\Omega$ the number of nodal points restricted to the computational domain $\Omega$. Clearly, the

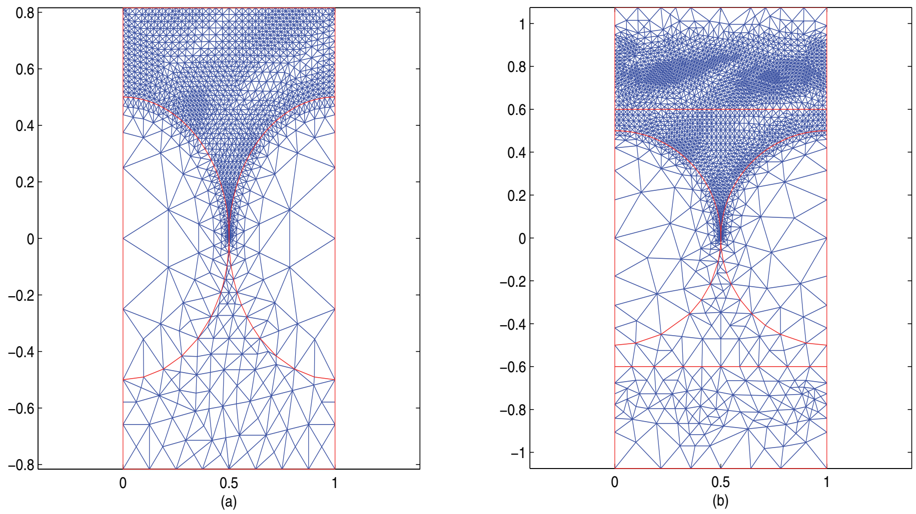


FIG. 9. *An adaptively refined mesh* (a) *with* 4024 *elements using the DtN method and* (b) *with* 6320 *elements using the PML method, for Example* 5.2.

accuracy of the adaptive PML solution is determined by $\mathrm{DoF}_h^\Omega$, the degrees of freedom in the computational domain $\Omega$. Figure 6(a) shows the curves of the priori error $e_h$ versus the number of nodal points for our adaptive DtN method and the adaptive PML method. Figure 6(b) shows the CPU time as a function of the number of nodal points for our adaptive DtN method and the adaptive PML method. From Figure 6, we observe that the error $e_h$ for the two methods is almost the same when $\mathrm{DoF}_h$ of our DtN method is equal to $\mathrm{DoF}_h^\Omega$ of the PML method, while the CPU time of the DtN method is a little shorter. The CPU time for the two methods is almost the same when the number of all nodal points $\mathrm{DoF}_h$ of both methods are equal, while accuracy of the DtN method is a little better. This inferiority of the adaptive PML method for Example 1 is due to a relatively large portion of degrees of Freedom being placed in the PML layers to approximate the solution there (see Figure 7). We remark that for higher-frequency problems, the ratio of degrees of freedom in the PML layers to those in the computational domain would be smaller (cf. [20]), while the number the truncation terms in our adaptive DtN method would be larger, and as a consequence the performances of the two methods should be similar. We would like to point out that most grating problems are essentially low-frequency problems (meaning that $k \times \mathrm{diam}(\Omega) = O(1)$).

This example clearly shows that our adaptive DtN method is effective and feasible, like the adaptive PML method.

*Example* 5.2. This example is a practical problem from [9] and is concerned with a cylindrical, metallic rod grating in TM polarization. The groove spacing is 1 $\mu$m, the radius of the rods is 0.5 $\mu$m, and their refractive index ($\sqrt{\varepsilon_2}$) is 1.3+7.6i. They are placed in a vacuum of $\varepsilon_1 = 1$ and illuminated under $\theta = \pi/6$ incidence angle by a 0.6328 $\mu$m-wavelength laser($\omega = 2\pi/0.6328$). Note that the parameters are taken as follows: $b_1 = 0.8164$, $b_2 = -0.8164$, $b_1' = 0.5$, $b_2' = -0.5$, $N_1 = 10$, and $N_2 = 10$.

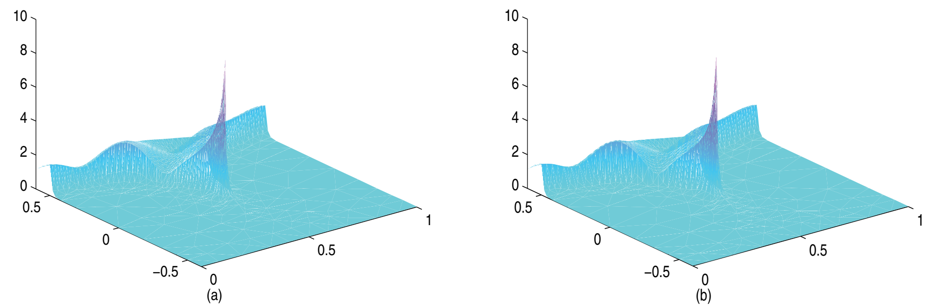


FIG. 10. *The surface plot of the amplitude of the associated numerical solution restricted in* $\Omega$ *on the mesh of Figure* 9 *for Example* 5.2*:* (a) *the adaptive DtN method;* (b)*: the adaptive PML method.*

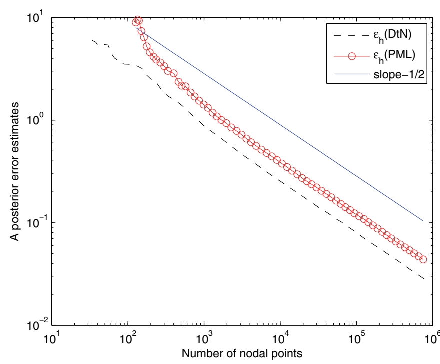


FIG. 11. *Quasi-optimality of the a posteriori error estimates for the adaptive DtN method and the adaptive PML method in Example* 5.2.

Figure 8 shows the zero-order reflection efficiency, the negative first-order reflection efficiency, and the total efficiency, as a function of the number of nodal points($\mathrm{DoF}_h$). It is clear that the efficiencies are convergent for both our adaptive DtN method and the adaptive PML method. The mesh plot and the amplitude of the associated numerical solution restricted in the domain $\Omega = [0, 1] \times [-0.6, 0.6]$ are shown for the two methods in Figures 9 and 10. It is easily seen that although there is a considerable difference in the meshes, the surface plots of the amplitude of the numerical solutions are pretty much the same for the two methods. It is also observed that both the adaptive DtN method and the adaptive PML method generate locally refined meshes in $\Omega$, which shows the ability of the two algorithms to capture the singularities of the problem. Figure 11 shows the curve of $\log \epsilon_h$ versus $\log \mathrm{DoF}_h$. It implies that decay of the a posteriori error estimates is $\mathcal{O}(\mathrm{DoF}_h^{-1/2})$ for the two algorithms.

Finally we conclude that our method is comparable to the adaptive PML method in the performance of approximation errors.

**6. Conclusion.** Based on the a posteriori error estimate, we presented an adaptive finite element method with DtN boundary condition for the diffraction grating problem. Numerical experiments included in this paper clearly demonstrate the competitive behavior of our proposed algorithm. Our point of view is that the adaptive DtN finite element method enriches the range of choices available for the numerical computation of wave propagation problems. The present work provides a viable alternative to the adaptive finite element method with PML for solving the same grating problem. We further hope the algorithm can be used to solve other scientific problems defined on an unbounded physical domain, especially when the PML techniques might not be applicable in those cases. Future work will be devoted to the extension of our analysis to the adaptive DtN finite element approximation of the two-dimensional diffraction grating problem governed by the Maxwell equations.

## REFERENCES


[1] T. Abboud, *Electromagnetic waves in periodic media*, in Proceedings of the 2nd International Conference on Mathematical and Numerical Aspects of Wave Propagation, Newark, DE, 1993, pp. 1–9.

[2] H. Ammari and G. Bao, *Maxwell's equations in periodic chiral structures*, Math. Nachr., 251 (2003), pp. 3–18.

[3] H. Ammari and J.-C. Nédélec, *Low-frequency electromagnetic scattering*, SIAM J. Math. Anal., 31 (2000), pp. 836–861.

[4] I. Babuška and A. Aziz, *Survey Lectures on Mathematical Foundations of the Finite Element Method, in the Mathematical Foundations of the Finite Element Method with Application to Partial Differential Equations*, A. Aziz, ed., Academic Press, New York, 1973.

[5] I. Babuška and W. C. Rheinboldt, *Error estimates for adaptive finite element computations*, SIAM J. Numer. Anal., 15 (1978), pp. 736–754.

[6] G. Bao, *Finite element approximation of time harmonic waves in periodic structures*, SIAM J. Numer. Anal., 32 (1995), pp. 1155–1169.

[7] G. Bao, *Numerical analysis of diffraction by periodic structures: TM polarization*, Numer. Math., 75 (1996), pp. 1–16.

[8] G. Bao, Y. Cao, and H. Yang, *Numerical solution of diffraction problems by a least-square finite element method*, Math. Methods Appl. Sci., 23 (2000), pp. 1073–1092.

[9] G. Bao, Z. Chen, and H. Wu, *Adaptive finite element method for diffraction gratings*, J. Opt. Soc. Amer. A, 22 (2005), pp. 1106–1114.

[10] G. Bao, L. Cowsar, and W. Masters, *Mathematical Modeling in Optical Science*, Frontiers in Appl. Math., SIAM, Philadelphia, 2001.

[11] G. Bao, D. C. Dobson, and J. A. Cox, *Mathematical studies in rigorous grating theory*, J. Opt. Soc. Amer. A, 12 (1995), pp. 1029–1042.

[12] G. Bao, P. Li, and H. Wu, *An adaptive edge element method with perfectly matched absorbing layers for wave scattering by biperiodic structures*, Math. Comp., 79 (2010), pp. 1–34.

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

[14] J.-P. Berenger, *A perfectly matched layer for the absorption of electromagnetic waves*, J. Comput. Phys., 114 (1994), pp. 185–200.

[15] P. Binev, W. Dahmen, and R. DeVore, *Adaptive finite element methods with convergence rates*, Numer. Math., 97 (2004), pp. 219–268.

[16] J. M. Cascon, C. Kreuzer, R. H. Nochetto, and K. G. Siebert, *Quasi-optimal convergence rate for an adaptive finite element method*, SIAM J. Numer. Anal., 46 (2008), pp. 2524–2550.

[17] J. Chen and Z. Chen, *An adaptive perfectly matched layer technique for 3-d time-harmonic electromagnetic scattering problems*, Math. Comp., 77 (2008), pp. 673–698.

[18] Z. Chen and S. Dai, *Adaptive Galerkin methods with error control for a dynamical Ginzburg–Landau model in superconductivity*, SIAM J. Numer. Anal., 38 (2001), pp. 1961–1985.

[19] Z. Chen and S. Dai, *On the efficiency of adaptive finite element methods for elliptic problems with discontinuous coefficients*, SIAM J. Sci. Comput., 24 (2002), pp. 443–462.

[20] Z. Chen and X. Liu, *An adaptive perfectly matched layer technique for time-harmonic scattering problems*, SIAM J. Numer. Anal., 43 (2005), pp. 645–671.

[21] 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 (2003), pp. 799–826.

[22] D. Dobson and A. Friedman, *The time-harmonic maxwell equations in a doubly periodic structure*, J. Math. Anal. Appl., 166 (1992), pp. 507–528.

[23] D. C. Dobson, *Optimal design of periodic antireflective structures for the Helmholtz equation*, European. J. Appl. Math, 4 (1993), pp. 321–340.

[24] W. Dörfler, *A convergent adaptive algorithm for Poisson's equation*, SIAM J. Numer. Anal., 33 (1996), pp. 1106–1124.

[25] N. H. Lord and A. J. Mulholland, *A dual weighted residual method applied to complex periodic gratings*, Proc. Roy. Soc. Edinburgh Sect. A, 469 (2013), 20130176.

[26] K. Mekchay and R. H. Nochetto, *Convergence of adaptive finite element methods for general second order linear elliptic PDEs*, SIAM J. Numer. Anal., 43 (2005), pp. 1803–1827.

[27] P. Monk, *A posteriori error indicators for Maxwell's equations*, J. Comput. Appl. Math., 100 (1998), pp. 173–190.

[28] P. Monk and E. Süli, *The adaptive computation of far-field patterns by a posteriori error estimation of linear functionals*, SIAM J. Numer. Anal., 36 (1998), pp. 251–274.

[29] P. Morin, R. H. Nochetto, and K. G. Siebert, *Data oscillation and convergence of adaptive FEM*, SIAM J. Numer. Anal., 38 (2000), pp. 466–488.

[30] P. Morin, R. H. Nochetto, and K. G. Siebert, *Convergence of adaptive finite element methods*, SIAM, Rev., 44 (2002), pp. 631–658.

[31] R. Petit, *Electromagnetic Theory of Gratings*, Topics in Current Physics 22, R. Petit, ed., Springer, Berlin, 1980.

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

[33] L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493.

[34] R. Stevenson, *Optimality of a standard adaptive finite element method*, Found. Comput. Math., 7 (2007), pp. 245–269.

[35] R. Verfürth, *A Review of A Posterior Error Estimation and Adaptive Mesh Refinement Techniques*, Teubner, Stuttgart, 1996.

[36] V. Yachin and K. Yasumoto, *Method of integral functionals for electromagnetic wave scattering from a double-periodic magnetodielectric layer*, J. Opt. Soc. Amer. A, 24 (2007), pp. 3606–3618.