SIAM J. NUMER. ANAL.
Vol. 59, No. 1, pp. 456–476

A RIGOROUS NUMERICAL ANALYSIS OF THE TRANSFORMED FIELD EXPANSION METHOD FOR DIFFRACTION BY PERIODIC, LAYERED STRUCTURES*

YOUNGJOON HONG† AND DAVID NICHOLLS‡

Abstract. Boundary perturbation methods have received considerable attention in recent years due to their ability to simulate solutions of differential equations of applied interest in a stable, robust, and highly accurate fashion. In this contribution we study the rigorous numerical analysis of a recently proposed high-order perturbation of surfaces method for scattering of electromagnetic waves by a doubly layered, periodic medium in transverse electric polarization. The algorithm in question is a transformed field expansion method which is discretized with a Fourier–Legendre-Galerkin, Taylor series approach. We prove not only results on existence and uniqueness of solutions but also theorems indicating that solutions of our scheme converge to these solutions with high-order spectral accuracy.

Key words. high-order perturbation of surfaces methods, high-order spectral methods, Helmholtz equation, diffraction gratings, layered media

AMS subject classifications. 65N35, 78A45, 78B22

DOI. 10.1137/20M131374X

1. Introduction.

We consider here the scattering of a time-harmonic electromagnetic plane wave by a periodically corrugated grating structure [32]. The scattering of linear waves involving periodic layered media plays a crucial role in a wide range of engineering and physics applications, e.g., materials science [17], nondestructive testing [38], sensing [19], geophysics [39], imaging [26], oceanography [9], and nanoplasmonics [34].

A number of computational methods have been developed for problems of scattering by periodic gratings. The most popular approaches to these problems are volumetric methods, such as finite differences and finite/spectral element methods [14, 3], but these methods are greatly disadvantaged with an unnecessarily large number of unknowns for piecewise homogeneous grating problems [27]. Interfacial methods based on integral equations (IEs) [13, 11, 24] are a natural alternative, but these also face several challenges. First, for periodic problems, the relevant Green function must be periodized, which greatly increases the computational cost. Additionally, these nonlocal IEs produce dense, nonsymmetric positive definite systems of linear equations which must be solved with each simulation.

A high-order perturbation of surfaces (HOPS) approach can avoid these concerns, such as the method of transformed field expansions (TFE) [28, 29], which we study here. These high-order algorithms were first developed by Bruno and Reitich for the two-dimensional scalar case [10] and later enhanced and stabilized by Nicholls and Reitich [28, 29] and Malcolm and Nicholls [25]. HOPS approaches are compelling, as

*Received by the editors January 17, 2020; accepted for publication (in revised form) November 16, 2020; published electronically February 18, 2021.
https://doi.org/10.1137/20M131374X
Funding: The work of the authors was supported by National Science Foundation grants DMS-1522548 and DMS-1813033.
†Department of Mathematics, Sungkyunkwan University, Suwon 440-746, Republic of Korea (hongyj@skku.edu); Department of Mathematics and Statistics, San Diego State University, San Diego, CA 92182 USA (yhong2@sdsu.edu).
‡Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago, Chicago, IL 60607 USA (davidn@uic.edu).

they maintain the advantageous properties of classical IE formulations (e.g., surface formulation and exact enforcement of far-field boundary conditions) while avoiding many of their shortcomings. For instance, since HOPS schemes utilize complex exponentials as basis functions in the lateral variable, the quasi-periodicity of solutions does not need to be explicitly enforced. In addition, due to the nature of the scheme, at every perturbation order, one need only invert a single, sparse operator corresponding to the flat-interface, order-zero approximation of the problem. The TFE method studied in this contribution was generalized by Nicholls and Shen to the case of irregular bounded obstacles in two [30] and three dimensions [16]. They later delivered a rigorous numerical analysis of the method [31], and we follow their strategy in this contribution. Subsequently, in [18, 21, 20, 22] the algorithms were extended to the case of periodic gratings separating multiply layered materials whose solutions are governed by either Helmholtz equations or the full Maxwell equations.

Of the immense literature (numbering several thousand papers) on the numerical simulation of this layered media problem (which we do not have the space to review here), we make special reference to the work of G. Bao and his group, as it is particularly relevant to our current approach. We encourage the interested reader to read the survey paper [6] and survey volume [4] for their efforts up to 2000. Beyond this, the original results on the weak formulation and finite element analysis of the problem which first appeared in [2, 3] has been extended to the least-squares framework [8] and the full vector Maxwell equations [5], including a periodic structure with perturbed interfaces [1], which we consider here. In addition, this group has done a great deal of work on the inverse problem of determining the geometrical features of the structure based on near-field imaging techniques. The paper [7] considers a problem particularly close to the one studied here, and the proofs appearing in our Appendix A use the same technology.

In this paper, we conduct a rigorous numerical analysis of the method developed by the authors [21, 20, 22] in the case of a doubly layered material with solutions satisfying a pair of Helmholtz equations coupled via the boundary conditions at the interface between the two. The TFE algorithm we derived is not only a stable and high-order numerical scheme, but it can also be used to directly establish the existence, uniqueness, and analyticity of solutions, as we presently demonstrate. For this purpose we establish a classical, but nontrivial, elliptic existence, uniqueness, and regularity theory by using the Green function and *a priori* estimates. The proof of our main result is based on analyticity estimates for the TFE expansions coupled to the convergence of the Fourier–Legendre–Galerkin method. Our developments illustrate the power and flexibility of the TFE approach for both numerical simulation and theoretical analysis.

2. Governing equations. To specify the problem and its geometry, we consider the two-dimensional Helmholtz problem which governs the scattering of electromagnetic waves in transverse electric polarization [32]

$$\Delta u + k_1^2 u = 0 \quad \text{in } z > g(x), \tag{2.1a}$$
$$\Delta v + k_2^2 v = 0 \quad \text{in } z < g(x), \tag{2.1b}$$
$$u - v = -u^{inc} \quad \text{at } z = g(x), \tag{2.1c}$$
$$\partial_N u - \partial_N v = -\partial_N u^{inc} \quad \text{at } z = g(x), \tag{2.1d}$$
$$\mathrm{OWC}[u] = 0 \quad \text{as } z \longrightarrow \infty, \tag{2.1e}$$
$$\mathrm{OWC}[v] = 0 \quad \text{as } z \longrightarrow -\infty, \tag{2.1f}$$
$$u(x+d, z) = e^{i\alpha d} u(x, z), \tag{2.1g}$$
$$v(x+d, z) = e^{i\alpha d} v(x, z), \tag{2.1h}$$

where $u^{inc} = e^{i\alpha x - i\gamma z}$, ∂_N is an upward pointing normal derivative and OWC connotes the outgoing wave condition, which we make precise presently.

2.1. Transparent boundary conditions. The usual procedure when implementing the TFE method is to truncate (if necessary) the unbounded problem domain to one of finite extent. For this, we introduce artificial boundaries above and below the structure and enforce transparent boundary conditions to equivalently solve (2.1). Introducing the planes $\{z = a > |g|_{L^\infty}\}$ and $\{z = b < -|g|_{L^\infty}\}$, we show that transparent boundary conditions can be enforced at these with Dirichlet–Neumann operators (DNOs) derived from the Rayleigh expansions [32]. These expansions are relevant, as they are the explicit solutions (obtained from separation of variables) of the problems on $\{z > a\}$ and $\{z < b\}$ on specification of Dirichlet data at the artificial boundaries, $\{z = a\}$ and $\{z = b\}$. More specifically, it is known [32] that

$$u = u(x,z) = \sum_{p=-\infty}^{\infty} \hat{\zeta}_p e^{i\alpha_p x} e^{i\gamma_{1,p}(z-a)}, \quad z > a,$$
$$v = v(x,z) = \sum_{p=-\infty}^{\infty} \hat{\psi}_p e^{i\alpha_p x} e^{i\gamma_{2,p}(b-z)}, \quad z < b,$$

where

$$\alpha_p := \alpha + \frac{2\pi}{d} p, \quad \gamma_{l,p} = \begin{cases} \sqrt{k_l^2 - \alpha_p^2}, & \alpha_p^2 \le k_l^2, \\ i\sqrt{\alpha_p^2 - k_l^2}, & \alpha_p^2 > k_l^2, \end{cases}$$

for $l = 1, 2$. We note that, upon evaluating at the artificial boundaries,

$$u(x,a) = \sum_{p=-\infty}^{\infty} \hat{\zeta}_p e^{i\alpha_p x} =: \zeta(x), \quad v(x,b) = \sum_{p=-\infty}^{\infty} \hat{\psi}_p e^{i\alpha_p x} =: \psi(x),$$

and from these we can compute the Neumann data at the artificial boundaries,

$$\partial_z u(x,a) = \sum_{p=-\infty}^{\infty} (i\gamma_{1,p}) \hat{\zeta}_p e^{i\alpha_p x}, \quad \partial_z v(x,b) = \sum_{p=-\infty}^{\infty} (-i\gamma_{2,p}) \hat{\psi}_p e^{i\alpha_p x}.$$

With these, we define the DNOs

$$T_1[\zeta] = T_1[u(x,a)] := \sum_{p=-\infty}^{\infty} (i\gamma_{1,p}) \hat{\zeta}_p e^{i\alpha_p x},$$
$$T_2[\psi] = T_2[v(x,b)] := \sum_{p=-\infty}^{\infty} (-i\gamma_{2,p}) \hat{\psi}_p e^{i\alpha_p x},$$

which are order-one Fourier multipliers. Using these, we can state (2.1) *equivalently* on the bounded domain $\{b < z < a\}$ as

$$\Delta u + k_1^2 u = 0 \quad \text{in } g(x) < z < a, \tag{2.2a}$$
$$\Delta v + k_2^2 v = 0 \quad \text{in } b < g(x) < z, \tag{2.2b}$$
$$u - v = -u^{inc} \quad \text{at } z = g(x), \tag{2.2c}$$

$$\partial_N u - \partial_N v = -\partial_N u^{inc} \qquad \text{at } z = g(x), \tag{2.2d}$$

$$\partial_z u - T_1[u] = 0 \qquad \text{at } z = a, \tag{2.2e}$$

$$\partial_z v - T_2[v] = 0 \qquad \text{at } z = b, \tag{2.2f}$$

$$u(x+d, z) = e^{i\alpha d} u(x,z), \tag{2.2g}$$

$$v(x+d, z) = e^{i\alpha d} v(x,z). \tag{2.2h}$$

3. TFE. We now recall the TFE method [28, 29], which begins with a domain flattening change of variables (also known as σ-coordinates [33] in the geophysical literature and the C-method [12] in the electromagnetics community). Subsequently, we make a boundary perturbation expansion which is solved recursively at each perturbation order.

3.1. The change of variables. We define the change of variables $x' = x$,

$$z_1 = a\left(\frac{z-g}{a-g}\right) \text{ for } g < z < a, \qquad z_2 = b\left(\frac{g-z}{g-b}\right) \text{ for } b < z < g,$$

and define

$$U_1(x', z_1) := u(x(x'), z(x', z_1, z_2)), \quad U_2(x', z_2) := v(x(x'), z(x', z_1, z_2)).$$

Using this change of variables, a long computation (see section 5) transforms (2.2) to the following system of equations:

$$\Delta_1 U_1 + k_1^2 U_1 = \frac{1}{G_1}(\partial_{x'} R_1^x + \partial_{z_1} R_1^z + R_1^0) =: R_1 \qquad \text{in } 0 < z_1 < a, \tag{3.1a}$$

$$\Delta_2 U_2 + k_2^2 U_2 = \frac{1}{G_2}(\partial_{x'} R_2^x + \partial_{z_2} R_2^z + R_2^0) =: R_2 \qquad \text{in } b < z_2 < 0, \tag{3.1b}$$

$$U_1 - U_2 = \xi_1 \qquad \text{at } z_1 = z_2 = 0, \tag{3.1c}$$

$$\partial_{z_1} U_1 - \partial_{z_2} U_2 = \xi_2 \qquad \text{at } z_1 = z_2 = 0, \tag{3.1d}$$

$$\partial_{z_1} U_1 - T_1[U_1] = -\frac{g}{a} T_1[U_1] =: J_1 \qquad \text{at } z_1 = a, \tag{3.1e}$$

$$\partial_{z_2} U_2 - T_2[U_2] = -\frac{g}{b} T_2[U_2] =: J_2 \qquad \text{at } z_2 = b, \tag{3.1f}$$

$$U_1(x'+d, z_1) = e^{i\alpha d} U_1(x', z_1), \tag{3.1g}$$

$$U_2(x'+d, z_2) = e^{i\alpha d} U_2(x', z_2), \tag{3.1h}$$

where the Laplacian operator Δ_l is defined by $\Delta_l = \partial_{x'}^2 + \partial_{z_l}^2$ for $l = 1, 2$. We refer the reader to section 5 for the specific formulas for R_l and ξ_l.

3.2. A high-order perturbation of surfaces method. We now introduce a boundary perturbation method to solve the transformed governing equations, (3.1). To begin, we assume that the deformation has the form

$$g(x') = \varepsilon f(x'), \quad f = \mathcal{O}(1),$$

and expand the fields in power series

$$\{U_1, U_2\} = \sum_{n=0}^{\infty} \{U_{1,n}, U_{2,n}\} \varepsilon^n.$$

As we shall soon see, not only does the interface need to be the graph of a function, $z_l = \varepsilon f(x')$, but also it must be sufficiently smooth; in this paper, we require $f \in C^{s+2}$ for $s \geq 0$. Inserting these expansions into (3.1) and equating at order $\mathcal{O}(\varepsilon^n)$ delivers

$$\Delta_1 U_{1,n} + k_1^2 U_{1,n} = R_{1,n} \quad \text{in } 0 < z_1 < a, \tag{3.2a}$$
$$\Delta_2 U_{2,n} + k_2^2 U_{2,n} = R_{2,n} \quad \text{in } b < z_2 < 0, \tag{3.2b}$$
$$U_{1,n} - U_{2,n} = \xi_{1,n} \quad \text{at } z_1 = z_2 = 0, \tag{3.2c}$$
$$\partial_{z_1} U_{1,n} - \partial_{z_2} U_{2,n} = \xi_{2,n} \quad \text{at } z_1 = z_2 = 0, \tag{3.2d}$$
$$\partial_{z_1} U_{1,n} - T_1[U_{1,n}] = -\frac{f}{a} T_1[U_{1,n-1}] =: J_{1,n} \quad \text{at } z_1 = a, \tag{3.2e}$$
$$\partial_{z_2} U_{2,n} - T_2[U_{2,n}] = -\frac{f}{b} T_2[U_{2,n-1}] =: J_{2,n} \quad \text{at } z_2 = b, \tag{3.2f}$$
$$U_{1,n}(x' + d, z_1) = e^{i\alpha d} U_{1,n}(x', z_1), \tag{3.2g}$$
$$U_{2,n}(x' + d, z_2) = e^{i\alpha d} U_{2,n}(x', z_2). \tag{3.2h}$$

Again, we refer the reader to the section 5 for the specific formulas for the right hand sides $R_{l,n}$ and $\xi_{l,n}$.

Considering the quasi-periodicity of solutions, we propose the generalized Fourier (Floquet) series expansions

$$U_{l,n}(x', z_l) = \sum_{p=-\infty}^{\infty} \widehat{U}_{l,n}^{(p)}(z_l) e^{i\alpha_p x'}, \quad R_{l,n}(x', z_l) = \sum_{p=-\infty}^{\infty} \widehat{R}_{l,n}^{(p)}(z_l) e^{i\alpha_p x'}, \tag{3.3a}$$
$$J_{l,n}(x') = \sum_{p=-\infty}^{\infty} \widehat{J}_{l,n}^{(p)} e^{i\alpha_p x'}, \quad \xi_{l,n}(x') = \sum_{p=-\infty}^{\infty} \widehat{\xi}_{l,n}^{(p)} e^{i\alpha_p x'}, \tag{3.3b}$$

for $l = 1, 2$. Inserting these expansions into (3.2), the governing equations are reduced to the one-dimensional boundary value problems

$$\partial_{z_1}^2 \widehat{U}_{1,n}^{(p)} + (k_1^2 - \alpha_p^2) \widehat{U}_{1,n}^{(p)} = \widehat{R}_{1,n}^{(p)} \quad \text{in } 0 < z_1 < a, \tag{3.4a}$$
$$\partial_{z_2}^2 \widehat{U}_{2,n}^{(p)} + (k_2^2 - \alpha_p^2) \widehat{U}_{2,n}^{(p)} = \widehat{R}_{2,n}^{(p)} \quad \text{in } b < z_2 < 0, \tag{3.4b}$$
$$\widehat{U}_{1,n}^{(p)} - \widehat{U}_{2,n}^{(p)} = \widehat{\xi}_{1,n}^{(p)} \quad \text{at } z_1 = z_2 = 0, \tag{3.4c}$$
$$\partial_{z_1} \widehat{U}_{1,n}^{(p)} - \partial_{z_2} \widehat{U}_{2,n}^{(p)} = \widehat{\xi}_{2,n}^{(p)} \quad \text{at } z_1 = z_2 = 0, \tag{3.4d}$$
$$\partial_{z_1} \widehat{U}_{1,n}^{(p)} - i\gamma_{1,p} \widehat{U}_{1,n}^{(p)} = -\frac{f}{a} (i\gamma_{1,p}) \widehat{U}_{1,n-1}^{(p)} =: \widehat{J}_{1,n}^{(p)} \quad \text{at } z_1 = a, \tag{3.4e}$$
$$\partial_{z_2} \widehat{U}_{2,n}^{(p)} + i\gamma_{2,p} \widehat{U}_{2,n}^{(p)} = -\frac{f}{b} (-i\gamma_{2,p}) \widehat{U}_{2,n-1}^{(p)} =: \widehat{J}_{2,n}^{(p)} \quad \text{at } z_2 = b. \tag{3.4f}$$

4. Function spaces. In order to use these TFE recursions in a direct proof of the existence, uniqueness, and analyticity of the solutions $\{u, v\}$ of (2.2), we must define our function spaces and state properties of these. To start, we recall, for the L^2 function $f = f(x')$, the classical Sobolev norm for any real $s \geq 0$ [23]:

$$\|f\|_{H^s}^2 := \sum_{p=-\infty}^{\infty} \langle p \rangle^{2s} \left| \hat{f}_p \right|^2, \quad \langle p \rangle^2 := 1 + |p|^2, \quad \hat{f}_p := \frac{1}{d} \int_0^d f(x') e^{-i\alpha_p x'} \, dx'.$$

For the L^2 function $w = w(x', z)$, we require the classical Sobolev norm for any integer $s \geq 0$ [15]:

$$\|w\|_{H^s}^2 := \sum_{k=0}^{s} \sum_{p=-\infty}^{\infty} \langle p \rangle^{2(s-k)} \int_{z_\ell}^{z_u} \left| \partial_z^k \hat{w}_p(z) \right|^2 \, dz.$$

With these norms, we define the function spaces, for real $s \geq 0$,

$$H^s([0,d]) := \left\{ f(x') \in L^2([0,d]) \mid \|f\|_{H^s} < \infty \right\}$$

and, for integer $s \geq 0$,

$$H^s([0,d] \times [z_\ell, z_u]) := \left\{ w(x', z) \in L^2([0,d] \times [z_\ell, z_u]) \mid \|w\|_{H^s} < \infty \right\}$$

and

$$\begin{aligned} H^s([0,d] \times [a,b]) := \{\{w_1, w_2\} \mid w_1 \in H^s([0,d] \times [0,a]), w_2 \in H^s([0,d] \times [b,0]), \\ w_1(0) = w_2(0)\}. \end{aligned}$$

Additionally, we will require their duals, H^{-s} [15].

We recall the following algebra property of Sobolev spaces (see, e.g., [28]), which allows us to estimate products of elements in these classes.

LEMMA 4.1. *Given any integer $s \geq 0$ and any $\sigma > 0$, there exists a constant $\kappa = \kappa(s, \sigma)$ such that if $f \in C^s([0,d])$ and $w \in H^s([0,d] \times [b,a])$, then*

$$\|fw\|_{H^s} \leq \kappa \, |f|_{C^s} \, \|w\|_{H^s} \, , \tag{4.1}$$

and if $\tilde{f} \in C^{s+1/2+\sigma}([0,d])$ and $\tilde{w} \in H^{s+1/2}([0,d])$, then

$$\left\| \tilde{f} \tilde{w} \right\|_{H^{s+1/2}} \leq \kappa \left| \tilde{f} \right|_{C^{s+1/2+\sigma}} \|\tilde{w}\|_{H^{s+1/2}} \, . \tag{4.2}$$

We also recall an elementary property of H^s.

LEMMA 4.2. *Given any integer $s \geq 0$, if $F \in H^s([0,d] \times [b,a])$, then $(a-z)F \in H^s([0,d] \times [b,a])$, and there exists a positive constant $Z_a = Z_a(s)$ such that*

$$\|(a-z)F\|_{H^s} \leq Z_a \, \|F\|_{H^s} \, .$$

As we shall see, the key tool for establishing our result is the following elliptic estimate, which allows us to show that unique solutions exist to the prototype problem above, (3.2), in an appropriate Sobolev space.

THEOREM 4.3. *Given any integer $s \geq 0$, if $\{F_1, F_2\} \in H^{s-1}([0,d] \times [b,a])$, $\xi \in H^{s+1/2}([0,d])$, and $\nu, K_1, K_2 \in H^{s-1/2}([0,d])$, then there exists a unique solution pair $\{u, v\} \in H^{s+1}([0,d] \times [b,a])$ of*

$$\Delta_1 u + k_1^2 u = F_1, \qquad 0 < z_1 < a, \tag{4.3a}$$

$$\Delta_2 v + k_2^2 v = F_2, \qquad b < z_1 < 0, \tag{4.3b}$$

$$u - v = \xi, \qquad z_1 = z_2 = 0, \tag{4.3c}$$

$$\partial_{z_1} u - \partial_{z_2} v = \nu, \qquad z_1 = z_2 = 0, \tag{4.3d}$$

$$\partial_{z_1} u - T_1[u] = K_1, \qquad z_1 = a, \tag{4.3e}$$

$$\partial_{z_2} v - T_2[v] = K_2, \qquad z_2 = b, \tag{4.3f}$$

$$u(x'+d,z) = e^{i\alpha d} u(x',z), \tag{4.3g}$$

$$v(x'+d,z) = e^{i\alpha d} v(x',z) \tag{4.3h}$$

such that, for a universal constant K_e,

$$\max\{\|u\|_{H^{s+1}}, \|v\|_{H^{s+1}}\} \le K_e \{\|F_1\|_{H^{s-1}} + \|F_2\|_{H^{s-1}} + \|\xi\|_{H^{s+1/2}} + \|\nu\|_{H^{s-1/2}} + \|K_1\|_{H^{s-1/2}} + \|K_2\|_{H^{s-1/2}}\}.$$

We give the proof in Appendix A.

5. Existence, uniqueness, and analyticity. To study the existence, uniqueness, and analyticity of solutions, we recall (3.2) and present precise expressions for the terms on the right hand sides. Recalling that $R_{l,n} = \partial_{x'} R^x_{l,n} + \partial_{z_l} R^z_{l,n} + R^0_{l,n}$, it can be shown that

$$\begin{aligned}
R^x_{1,n} &= \frac{2}{a} f \partial_{x'} U_{1,n-1} + \frac{a-z_1}{a}(\partial_{x'} f)\partial_{z_1} U_{1,n-1} - \frac{f^2}{a^2}\partial_{x'} U_{1,n-2} \\
&\quad - \frac{a-z_1}{a} f(\partial_{x'} f)\partial_{z_1} U_{1,n-2}, \\
R^z_{1,n} &= \frac{a-z_1}{a}(\partial_{x'} f)\partial_{x'} U_{1,n-1} - \frac{a-z_1}{a^2} f(\partial_{x'} f)\partial_{x'} U_{1,n-2} \\
&\quad - \frac{(a-z_1)^2}{a^2}(\partial_{x'} f)^2 \partial_{z_1} U_{1,n-2}, \\
R^0_{1,n} &= -\frac{1}{a}(\partial_{x'} f) U_{1,n-1} + (k_1^2)\frac{2f}{a} U_{1,n-1} + \frac{1}{a^2} f(\partial_{x'} f)\partial_{x'} U_{1,n-2} \\
&\quad + \frac{a-z_1}{a^2}(\partial_{x'} f)^2 \partial_{z_1} U_{1,n-2} - k_1^2 \frac{f^2}{a^2} U_{1,n-2},
\end{aligned}$$

similarly for $R_{2,n}$ and

$$\xi_{1,n} = (-1)^{n+1}\frac{(i\gamma f)^n}{n!} e^{i\alpha x}, \quad \xi_{2,n} = \frac{Q_{1,n} + Q_{2,n}}{ab},$$

where

$$\begin{aligned}
Q_{1,n} &= -iab\gamma \xi_{1,n} - iab\alpha(\partial_{x'} f)\xi_{1,n-1} - i\gamma(a-b) f \xi_{1,n-1} \\
&\quad - i\alpha(a-b) f(\partial_{x'} f)\xi_{1,n-2} + i\gamma f^2 \xi_{1,n-2} + i\alpha(\partial_{x'} f) f^2 \xi_{1,n-3}, \\
Q_{2,n} &= -af\partial_{z_1} U_{1,n-1} + ab(\partial_{x'} f)\partial_{x'} U_{1,n-1} + (a-b) f(\partial_{x'} f)\partial_{x'} U_{1,n-2} \\
&\quad - ab(\partial_{x'} f)^2 \partial_{z_1} U_{1,n-2} - (\partial_{x'} f) f^2 \partial_{x'} U_{1,n-3} - a(\partial_{x'} f)^2 f \partial_{z_1} U_{1,n-3} \\
&\quad - bf\partial_{z_2} U_{2,n-1} - ab(\partial_{x'} f)\partial_{x'} U_{2,n-1} - (a-b) f(\partial_{x'} f)\partial_{x'} U_{2,n-2} \\
&\quad + ab(\partial_{x'} f)^2 \partial_{z_2} U_{2,n-2} + (\partial_{x'} f) f^2 \partial_{x'} U_{2,n-3} - bf(\partial_{x'} f)^2 \partial_{z_2} U_{2,n-3}.
\end{aligned}$$

To begin our demonstration, we establish the analyticity of the Dirichlet data.

LEMMA 5.1. *Given any integer $s \ge 0$, if $f \in C^{s+2}([0,d])$, then*

$$\|\xi_{1,n}\|_{H^{s+3/2}} \le K_\xi B_\xi^n \tag{5.1}$$

for constants $K_\xi, B_\xi > 0$.

Proof. We note that $\xi_{1,n} = -i\gamma f \xi_{1,n-1}/n$ and use induction to prove this lemma. We begin at $n = 0$ and set

$$K_\xi := \|\xi_{1,0}\|_{H^{s+3/2}}.$$

We now assume (5.1) for all $n < \bar{n}$ and consider $\bar{n} > 1$, where we bound

$$\begin{aligned}\|\xi_{1,\bar{n}}\|_{H^{s+3/2}} &\le |\gamma|\, M\, |f|_{C^{s+3/2+\sigma}} \|\xi_{1,\bar{n}-1}\|_{H^{s+3/2}} \\ &\le |\gamma|\, M\, |f|_{C^{s+2}} K_\xi B_\xi^{\bar{n}-1}.\end{aligned}$$

By choosing $B_\xi > M|\gamma|\,|f|_{C^{s+2}}$, the lemma follows. ∎

We now provide the key inductive lemma which enables the proof of our result.

LEMMA 5.2. *Given any integer $s \ge 0$, if $f \in C^{s+2}([0,d])$ and*

$$\|U_{1,n}\|_{H^{s+2}} + \|U_{2,n}\|_{H^{s+2}} \le K B^n \quad \forall n < \bar{n}$$

for constants $K, B > 0$, then there exists a constant $\bar{C} > 0$ such that

$$\max\left\{\|R_{l,\bar{n}}\|_{H^s}, \|J_{l,\bar{n}}\|_{H^{s+1/2}}, \|\xi_{2,\bar{n}}\|_{H^{s+1/2}}\right\} \le K\bar{C}\left(B^{\bar{n}-1} + B^{\bar{n}-2} + B^{\bar{n}-3}\right)$$

for $l = 1, 2$.

Proof. For $l = 1, 2$, we recall that $R_{l,\bar{n}} = \partial_{x'} R^x_{l,\bar{n}} + \partial_{z_l} R^z_{l,\bar{n}} + R^0_{l,\bar{n}}$, so that one can deduce

$$\|R_{l,\bar{n}}\|_{H^s} \lesssim \left\|R^x_{l,\bar{n}}\right\|_{H^{s+1}} + \left\|R^z_{l,\bar{n}}\right\|_{H^{s+1}} + \left\|R^0_{l,\bar{n}}\right\|_{H^s},$$

where $\|A\| \lesssim \|B\|$ means that there exists a constant C, independent of all variables of importance, such that $\|A\| \le C\,\|B\|$. With the estimates

$$\begin{aligned}\left\|R^x_{l,\bar{n}}\right\|_{H^{s+1}} &\lesssim \|f\partial_{x'}U_{l,\bar{n}-1}\|_{H^{s+1}} + \|\partial_{x'}f\partial_{z_l}U_{l,\bar{n}-1}\|_{H^{s+1}} \\ &\quad + \left\|f^2\partial_{x'}U_{l,\bar{n}-2}\right\|_{H^{s+1}} + \|f(\partial_{x'}f)\partial_{z_l}U_{l,\bar{n}-2}\|_{H^{s+1}} \\ &\lesssim |f|_{C^{s+1}}\|U_{l,\bar{n}-1}\|_{H^{s+2}} + |f|_{C^{s+2}}\|U_{l,\bar{n}-1}\|_{H^{s+2}} \\ &\quad + |f|^2_{C^{s+1}}\|U_{l,\bar{n}-2}\|_{H^{s+2}} + |f|^2_{C^{s+2}}\|U_{l,\bar{n}-2}\|_{H^{s+2}} \\ &\lesssim 2\,|f|_{C^{s+2}} KB^{\bar{n}-1} + 2\,|f|^2_{C^{s+2}} KB^{\bar{n}-2}\end{aligned}$$

and

$$\begin{aligned}\left\|R^z_{l,\bar{n}}\right\|_{H^{s+1}} &\lesssim \|(\partial_{x'}f)\partial_{x'}U_{l,\bar{n}-1}\|_{H^{s+1}} + \|f(\partial_{x'}f)\partial_{x'}U_{l,\bar{n}-2}\|_{H^{s+1}} \\ &\quad + \left\|(\partial_{x'}f)^2\partial_{z_l}U_{l,\bar{n}-2}\right\|_{H^{s+1}} \\ &\lesssim |f|_{C^{s+2}} KB^{\bar{n}-1} + |f|^2_{C^{s+2}} KB^{\bar{n}-2}\end{aligned}$$

and

$$\begin{aligned}\left\|R^0_{l,\bar{n}}\right\|_{H^s} &\lesssim \|(\partial_{x'}f)U_{l,\bar{n}-1}\|_{H^s} + \|fU_{l,\bar{n}-1}\|_{H^s} + \|f(\partial_{x'}f)\partial_{x'}U_{l,\bar{n}-2}\|_{H^s} \\ &\quad + \left\|(\partial_{x'}f)^2\partial_{z_l}U_{l,\bar{n}-2}\right\|_{H^s} + \left\|f^2U_{l,\bar{n}-2}\right\|_{H^s} \\ &\lesssim 2\,|f|_{C^{s+2}} KB^{\bar{n}-1} + 3\,|f|^2_{C^{s+2}} KB^{\bar{n}-2},\end{aligned}$$

we find that

$$\|R_{l,\bar{n}}\|_{H^s} \lesssim K \left(|f|_{C^{s+2}} B^{\bar{n}-1} + |f|^2_{C^{s+2}} B^{\bar{n}-2} \right).$$

For $J_{l,\bar{n}}$, we can show that

$$\begin{aligned} \|J_{l,\bar{n}}\|_{H^{s+1/2}} &\lesssim \|f T_l[U_{l,\bar{n}-1}]\|_{H^{s+1/2}} \\ &\lesssim |f|_{C^{s+1/2+\sigma}} \|U_{l,\bar{n}-1}\|_{H^{s+3/2}} \\ &\lesssim |f|_{C^{s+1/2+\sigma}} K B^{\bar{n}-1}. \end{aligned}$$

Hence, we deduce that

$$\begin{aligned} \max\left\{ \|R_{l,\bar{n}}\|_{H^s}, \|J_{l,\bar{n}}\|_{H^{s+1/2}} \right\} &\le KC \left(|f|_{C^{s+2}} B^{\bar{n}-1} + |f|^2_{C^{s+2}} B^{\bar{n}-2} \right) \\ &\lesssim K \left(B^{\bar{n}-1} + B^{\bar{n}-2} \right). \end{aligned}$$

It remains to estimate $\xi_{2,\bar{n}}$, and for this, we use Lemma 5.1, which implies

$$\|\xi_{1,n}\|_{H^{s+1/2}} \le K_\xi B_\xi^n;$$

hence,

$$\|Q_{1,\bar{n}}\|_{H^{s+1/2}} \le KC \left(B_\xi^{n-1} + B_\xi^{n-2} + B_\xi^{n-3} \right).$$

In addition, we find, for $l = 1, 2$,

$$\begin{aligned} \|Q_{2,\bar{n}}\|_{H^{s+1/2}} &\lesssim \|\partial_{z_l} U_{l,\bar{n}-1}\|_{H^{s+1/2}} + \|\partial_{x'} U_{l,\bar{n}-1}\|_{H^{s+1/2}} + \|\partial_{x'} U_{l,\bar{n}-2}\|_{H^{s+1/2}} \\ &\quad + \|\partial_{z_l} U_{l,\bar{n}-2}\|_{H^{s+1/2}} + \|\partial_{x'} U_{l,\bar{n}-3}\|_{H^{s+1/2}} + \|\partial_{z_l} U_{l,\bar{n}-3}\|_{H^{s+1/2}} \\ &\lesssim \|U_{l,\bar{n}-1}\|_{H^{s+2}} + \|U_{l,\bar{n}-2}\|_{H^{s+2}} + \|U_{l,\bar{n}-3}\|_{H^{s+2}} \\ &\lesssim K \left(B^{\bar{n}-1} + B^{\bar{n}-2} + B^{\bar{n}-3} \right), \end{aligned}$$

and the lemma follows. ∎

We can now state and prove the main theorem of this section.

THEOREM 5.3. *Given any integer $s \ge 0$, if $f \in C^{s+2}([0,d])$ and $\xi_{1,n} \in H^{s+3/2}([0,d])$ such that*

$$\|\xi_{1,n}\|_{s+3/2} \le K_\xi B_\xi^n$$

for constants $K_\xi, B_\xi > 0$, then $U_{l,n} \in H^{s+2}([0,d] \times [b,a])$ for $l = 1, 2$ and

$$\|U_{1,n}\|_{H^{s+2}} + \|U_{2,n}\|_{H^{s+2}} \le K B^n \tag{5.2}$$

for some universal constant K.

Proof. We proceed by induction, and at order $n = 0$, Theorem 4.3 guarantees a unique solution such that

$$\|U_{1,0}\|_{H^{s+2}} + \|U_{2,0}\|_{H^{s+2}} \le K_e \|\xi_{1,0}\|_{H^{s+3/2}},$$

so we choose $K \ge K_e \|\xi_{1,0}\|_{H^{s+3/2}}$. We now assume (5.2) holds for all $n \le \bar{n}$, and from Theorem 4.3, we deduce that

$$\|U_{1,\bar{n}}\|_{H^{s+2}} + \|U_{2,\bar{n}}\|_{H^{s+2}} \leq C_1 \left(\|R_{1,\bar{n}}\|_{H^s} + \|R_{2,\bar{n}}\|_{H^s} + \|J_{1,\bar{n}}\|_{H^{s+1/2}} + \|J_{2,\bar{n}}\|_{H^{s+1/2}} + \|\xi_{2,\bar{n}}\|_{H^{s+1/2}} \right) + C_2 \|\xi_{1,\bar{n}}\|_{H^{s+3/2}}.$$

Appealing to Lemmas 5.1 and 5.2, we find that

$$\|U_{1,\bar{n}}\|_{H^{s+2}} + \|U_{2,\bar{n}}\|_{H^{s+2}} \leq 5C_1 K\bar{C} \left(B^{\bar{n}-1} + B^{\bar{n}-2} + B^{\bar{n}-3} \right) + C_2 K_\xi B_\xi^{\bar{n}}.$$

Now, on choosing $K > C_2 K_\xi$ and

$$B > \max \left\{ B_\xi, 5C_1\bar{C}, (5C_1\bar{C})^{1/2}, (5C_1\bar{C})^{1/3} \right\},$$

the theorem follows. ∎

Remark 5.4. We point out that this result is quite similar to Theorem 2.13 of [7], which establishes an analogous estimate in the single-layer setting.

6. Convergence analysis. We are now in a position to conduct a numerical analysis of our TFE approach. We recall the TFE recursions (3.2) and note that, in practice, we make use of the Floquet series representation, (3.3), and focus our attention on the reduced problem (3.4). We further specialize by splitting this into two: a homogeneous Helmholtz problem with inhomogeneous coupling ($\widehat{\xi}_{j,n}^{(p)} \not\equiv 0$) (see (B.1)) and an inhomogeneous Helmholtz problem with homogeneous coupling ($\widehat{\xi}_{j,n}^{(p)} \equiv 0$) (see (B.2)). Clearly, the solution of (3.4) is the sum of the solutions of these two problems, and, in practical numerical implementations, we need only solve the latter, as (B.1) can be solved explicitly via separation of variables, e.g., [18, 21, 20]. For this reason, we focus on (B.2), and, for simplicity, we suppress the index n. The weak form of this boundary value problem is

$$\text{Find } \widetilde{U}^{(p)} \in H^1(b,a) \text{ such that}$$
$$B(\widetilde{U}^{(p)}, \varphi) = R(\varphi) \quad \forall\, \varphi \in H^1(b,a), \tag{6.1}$$

where

$$\begin{aligned} B(\widetilde{U}^{(p)}, \varphi) &:= -i\gamma_{1,p} \widetilde{U}_1^{(p)}(a)\bar{\varphi}_1(a) - i\gamma_{2,p}\widetilde{U}_2^{(p)}(b)\bar{\varphi}_2(b) \\ &\quad + \int_b^a \partial_z \widetilde{U}^{(p)} \partial_z \bar{\varphi}\, dz - \gamma_p^2 \int_b^a \widetilde{U}^{(p)} \bar{\varphi}\, dz, \\ R(\varphi) &:= \widehat{J}_1^{(p)} \bar{\varphi}(a) - \widehat{J}_2^{(p)} \bar{\varphi}(b) + \int_b^a (-\widehat{R}^{(p)}) \bar{\varphi}\, dz. \end{aligned}$$

For our numerical analysis, we define the discrete function space

$$\begin{aligned} X_{M,p} = \operatorname{span} \Big\{ & u \in C(b,a) \mid u|_{(0,a)}, u|_{(b,0)} \in P_M, \\ & (\partial_z u - i\gamma_{1,p} u)(a) = \widehat{J}_1^{(p)}, (\partial_z u + i\gamma_{2,p} u)(b) = \widehat{J}_2^{(p)} \Big\}, \end{aligned}$$

where P_M is the space of all complex valued polynomials of degree less than or equal to M. The Legendre–Galerkin approximation of (6.1) is as follows:

$$\text{Find } \widetilde{U}^{(p),M} \in X_{M,p} \text{ such that}$$
$$B(\widetilde{U}^{(p),M}, \varphi_M) = R(\varphi_M) \quad \forall\, \varphi_M \in X_{M,p}. \tag{6.2}$$

To prove the main theorem of this section, the following interpolation result [35] is required for the projection ${}_0\Pi_M^1$ from $H^1(b,a)$ to P_M subject to the boundary conditions of the space $X_{M,p}$.

LEMMA 6.1. *There exists a mapping* ${}_0\Pi_M^1 : H^1(b,a) \to X_{M,p}$ *such that*

$$\big(\partial_z({}_0\Pi_M^1 V - V), \partial_z \varphi_M\big) = 0 \quad \forall\ \varphi_M \in X_{M,p}.$$

Moreover, for $1 \le l \le M+1$, *we have*

$$\left\|{}_0\Pi_M^1 V - V\right\|_{H^\mu} \lesssim \sqrt{\frac{(M-l+1)!}{M!}}(M+l)^{\mu-(l+1)/2}\left\|\partial_z^l V\right\|_{L^2},$$

where $\mu = 0, 1$.

Proof. We prove this lemma for $V_1 := V|_{(0,a)}$. By the straightforward change of variables $x = 2z/a - 1$, the domain of $V_1 \in H^1(0,a)$ can be transformed to the interval $(-1,1)$. Thus, we establish the result for a real valued function $v(x)$ on $\Lambda = (-1,1)$. Let $\Pi_M^{1,0}$ be the H_0^1-orthogonal projection operator onto $P_M \times P_M$, and, for any $v \in H^1(\Lambda)$, we define $v_*(x)$ by

$$v(x) = v_*(x) + \left(\frac{1+x}{2}\right)v(1) + \left(\frac{1-x}{2}\right)v(-1), \quad v_*(x) \in H_0^1(\Lambda).$$

Similarly, we define $\varphi_*(x)$ by

$$\varphi(x) = \varphi_*(x) + \left(\frac{1+x}{2}\right)v(1) + \left(\frac{1-x}{2}\right)v(-1), \quad \varphi_*(x) \in H_0^1(\Lambda),$$

for any $\varphi \in P_M$. Regarding

$${}_0\Pi_M^1 v(x) := \Pi_M^{1,0} v_*(x) + \left(\frac{1+x}{2}\right)v(1) + \left(\frac{1-x}{2}\right)v(-1),$$

we observe that

$$\begin{aligned}\big(\partial_x({}_0\Pi_M^1 v - v), \partial_x\varphi\big)_\Lambda &= \Big(\partial_x(\Pi_M^{1,0}v_* - v_*), \partial_x\varphi_*\Big)_\Lambda \\ &\quad + \left(\frac{v(1)}{2} - \frac{v(-1)}{2}\right)\int_{-1}^1 \partial_x(\Pi_M^{1,0}v_* - v_*)(x)\,dx \\ &= \Big(\partial_x(\Pi_M^{1,0}v_* - v_*), \partial_x\varphi_*\Big)_\Lambda = 0\end{aligned}$$

for $\varphi \in P_M$. By Theorem 3.39 in [35], we find

$$\begin{aligned}\left\|{}_0\Pi_M^1 v(x) - v\right\|_{H^\mu} &= \left\|\Pi_M^{1,0}v_* - v_*\right\|_{H^\mu} \\ &\lesssim \sqrt{\frac{(M-l+1)!}{M!}}(M+l)^{\mu-(l+1)/2}\left\|\partial_x^l v_*\right\|_{L^2}.\end{aligned}$$

For $l = 1$, by the Poincaré inequality, we derive that

$$\|\partial_x v_*\|_{L^2} \le \|\partial_x v\|_{L^2} + c\,(|v(1)| + |v(-1)|) \le c\,\|\partial_x v\|_{L^2}.$$

From this, the lemma follows. ∎

Now we are ready to prove the convergence theorem.

THEOREM 6.2. *Let $\widetilde{U}^{(p)}$ and $\widetilde{U}^{(p),M}$ be the solutions of* (6.1) *and* (6.2), *respectively. Then, for $1 \leq l \leq M+1$, we have*

$$\left\|\widetilde{U}^{(p)} - \widetilde{U}^{(p),M}\right\|_{H^1} + |\gamma_p| \left\|\widetilde{U}^{(p)} - \widetilde{U}^{(p),M}\right\|_{L^2} \leq (1+\gamma_p^2 M^{-1}) \sqrt{\frac{(M-l+1)!}{M!}} (M+l)^{(1-l)/2} \left\|\partial_z^l U^{(p)}\right\|_{L^2}.$$

Proof. Let

$$e_M := \widetilde{U}^{(p),M} - {}_0\Pi_M^1 \widetilde{U}^{(p)}, \quad \tilde{e}_M = \widetilde{U}^{(p)} - {}_0\Pi_M^1 \widetilde{U}^{(p)}.$$

For $\varphi_M \in X_{M,p}$, using (6.1) and (6.2), we find

$$B(\widetilde{U}^{(p)} - \widetilde{U}^{(p),M}, \varphi_M) = 0.$$

Using Lemma 6.1, we obtain

$$\begin{aligned} B(e_M, \varphi_M) &= B(\widetilde{U}^{(p),M} - \widetilde{U}^{(p)} + \widetilde{U}^{(p)} - {}_0\Pi_M^1 \widetilde{U}^{(p)}, \varphi_M) \\ &= B(\widetilde{U}^{(p)} - {}_0\Pi_M^1 \widetilde{U}^{(p)}, \varphi_M) \\ &= -\gamma_p^2 (\tilde{e}_M, \varphi_M) - i\gamma_{1,p} \tilde{e}_M(a) \bar{\varphi}_M(a) - i\gamma_{2,p} \tilde{e}_M(b) \bar{\varphi}_M(b). \end{aligned} \tag{6.3}$$

In view of (6.3), we rewrite (6.1) by replacing $\{U^{(p)}, \widehat{J}_1^{(p)}, \widehat{J}_2^{(p)}, \widehat{R}^{(p)}\}$ with

$$\{e_M, -i\gamma_{1,p}\tilde{e}_M(a), i\gamma_{2,p}\tilde{e}_M(b), \gamma_p^2 \tilde{e}_M\},$$

respectively. Then, by the regularity result (B.7) from Appendix B, we obtain that

$$\|e_M\|_{H^1}^2 + \gamma_p^2 \|e_M\|_{L^2}^2 \lesssim \gamma_p^4 \|\tilde{e}_M\|_{L^2}^2 + \gamma_{1,p}^2 |\tilde{e}_M(a)|^2 + \gamma_{2,p}^2 |\tilde{e}_M(b)|^2.$$

By the Gagliardo–Nirenberg interpolation inequality [36] and Lemma 6.1, we find

$$\begin{aligned} |\tilde{e}(\pm 1)| &\lesssim \|\tilde{e}_M\|_{L^2}^{1/2} \|\tilde{e}_M\|_{H^1}^{1/2} \\ &\lesssim \sqrt{\frac{(M-l+1)!}{M!}} (M+l)^{-l/2} \left\|\partial_z^l U^{(p)}\right\|_{L^2}. \end{aligned}$$

Using Lemma 6.1 again, we deduce that

$$\begin{aligned} &\left\|U^{(p)} - U^{(p),M}\right\|_{H^1} + |\gamma_p| \left\|U^{(p)} - U^{(p),M}\right\|_{L^2} \\ &\qquad \lesssim (\|e_M\|_{H^1} + |\gamma_p| \|e_M\|_{L^2} + \|\tilde{e}_M\|_{H^1} + |\gamma_p| \|\tilde{e}_M\|_{L^2}) \\ &\lesssim \left(1 + \gamma_p^2 M^{-1} + |\gamma_p| M^{-1/2}\right) \sqrt{\frac{(M-l+1)!}{M!}} (M+l)^{(1-l)/2} \left\|\partial_z^l U^{(p)}\right\|_{L^2}. \qquad \square \end{aligned}$$

We now reintroduce the index n and let

$$U_{l,n}^{(P),M}(x,z) := \sum_{p=-P}^{P} \widehat{U}_{l,n}^{(p),M}(z) e^{i\alpha_p x}, \quad l = 0, 1,$$

be the Fourier–Legendre approximation of the solution $U_{l,n}$ of (3.2). Using the same argument as in Theorem 3.3 in [31], we can prove the following estimate.

THEOREM 6.3. *For any integer $r \geq 1$, if $U \in H^r$, then*

$$\left\|\nabla(U_{l,n} - U_{l,n}^{(P),M})\right\|_{L^2} + k_l \left\|U_{l,n} - U_{l,n}^{(P),M}\right\|_{L^2}$$
$$\lesssim \left(P^{1-r} + (1 + k_l^2 M^{-1})\sqrt{\frac{(M-r+1)!}{M!}}(M+r)^{(1-r)/2}\right) \|U_{l,n}\|_{H^r}.$$

Finally, if we choose

$$U_{l,N}^{(P),M}(x,z) := \sum_{n=0}^{N} U_{l,n}^{(P),M}(x,z)\varepsilon^n$$

as our approximation to the solution U_m of (3.2), then, using Theorem 6.3 and Theorem 2.1 of [31], we have the final result.

THEOREM 6.4. *For any integer $r \geq 2$, if $f \in C^r([0,d])$, $\xi_1 \in H^{r-1/2}([0,d])$, and $\xi_2 \in H^{r-3/2}([0,d])$, then we have, for $l = 1, 2$,*

$$\left\|\nabla(U_l - U_{l,N}^{(P),M})\right\|_{L^2} + k_l \left\|U_l - U_{l,N}^{(P),M}\right\|_{L^2} \lesssim (B\varepsilon)^{N+1}$$
$$+ \left(P^{1-r} + (1 + k_l^2 M^{-1})\sqrt{\frac{(M-r+1)!}{M!}}(M+r)^{(1-r)/2}\right)$$
$$\times \left(\|\xi_1\|_{H^{r-1/2}} + \|\xi_2\|_{H^{r-3/2}}\right).$$

Remark 6.5. A similar result appears in Theorem 3.6 of [7]. The difference is that our new theorem concerns convergence of the *fields* as discretization parameters are refined, while [7] estimate errors in the interface reconstruction.

7. Conclusions. In this paper, we have provided a rigorous numerical analysis of a HOPS algorithm for electromagnetic scattering. Introducing DNOs at artificial boundaries placed above the top and below the bottom of the structure, we equivalently reformulated the governing Helmholtz equations for the doubly layered medium on a bounded domain. Using a suitable change of variables, the governing equations on a separable geometry with flat interfaces were derived. Introducing boundary perturbations, we described the scattered field in a Taylor series; more precisely, we derived a sequence of linear boundary value problems to be solved at each perturbation order resulting in the TFE algorithm. Our approach to establishing the convergence and accuracy of the TFE methodology is to combine analyticity theorems with results on Legendre–Galerkin methods. Our developments clearly point toward several extensions of great importance. In particular, our approach must be generalized to the three-dimensional vector wave equations of electromagnetics and linear elastodynamics. These extensions are not straightforward, as more complicated boundary conditions between layers are required. Hence, the algorithmic differences will be significant, and we will describe them in a future publication.

Appendix A. Proof of the elliptic estimate: Theorem 4.3. To begin our proof of Theorem 4.3, we state two classic results [29] on solutions of Helmholtz problems on each of the two layers separately.

THEOREM A.1. *Given any integer $s \geq 0$, if $F_1 \in H^{s-1}([0,d] \times [0,a])$, $U \in H^{s+1/2}([0,d])$, and $K_1 \in H^{s-1/2}([0,d])$, then there exists a unique solution $u \in$*

$H^{s+1}([0,d] \times [0,a])$ *of*

$$\Delta_1 u + k_1^2 u = F_1, \qquad 0 < z_1 < a, \tag{A.1a}$$

$$u = U, \qquad z_1 = 0, \tag{A.1b}$$

$$\partial_{z_1} u - T_1[u] = K_1, \qquad z_1 = a, \tag{A.1c}$$

such that

$$\|u\|_{H^{s+1}} \le C_u \left\{ \|F_1\|_{H^{s-1}} + \|U\|_{H^{s+1/2}} + \|K_1\|_{H^{s-1/2}} \right\}.$$

In addition, if $\tilde{U} = [-\partial_{z_1} u]_{z_1=0}$ *and we define the DNO*

$$\mathcal{G} : (U, K_1, F_1) \to \tilde{U}, \quad \mathcal{G}[U, K_1, F_1] = G^{(0)}[U] + G^{(a)}[K_1] + G^{([0,a])}[F_1],$$

then

$$\begin{aligned}
\left\| G^{(0)}[U] \right\|_{H^{s-1/2}} &\le C_{G^{(0)}} \|U\|_{H^{s+1/2}}, \\
\left\| G^{(a)}[K_1] \right\|_{H^{s-1/2}} &\le C_{G^{(a)}} \|K_1\|_{H^{s-1/2}}, \\
\left\| G^{([0,a])}[F_1] \right\|_{H^{s-1/2}} &\le C_{G^{([0,a])}} \|F_1\|_{H^{s-1}}.
\end{aligned}$$

Proof. For clarity of presentation, we drop the "1" subscript on all variables. Due to the quasi-periodic boundary conditions, we posit expansions

$$\{u, F\}(x,z) = \sum_{p=-\infty}^{\infty} \{\hat{u}_p, \hat{F}_p\}(z) e^{i\alpha_p x}, \quad \{U, K\}(x) = \sum_{p=-\infty}^{\infty} \{\hat{U}_p, \hat{K}_p\} e^{i\alpha_p x},$$

and (A.1) delivers the two-point boundary value problem

$$\begin{aligned}
&\partial_z^2 \hat{u}_p + \gamma_p^2 \hat{u}_p = \hat{F}_p, && 0 < z < a, \\
&\hat{u}_p(0) = \hat{U}_p, \\
&\partial_z \hat{u}_p(a) - (i\gamma_p)\hat{u}_p(a) = \hat{K}_p,
\end{aligned}$$

where

$$\gamma_p = \begin{cases} \gamma_p' := \sqrt{k^2 - \alpha_p^2}, & \alpha_p^2 < k^2, \\ 0, & \alpha_p^2 = k^2, \\ i\gamma_p'' := i\sqrt{\alpha_p^2 - k^2}, & \alpha_p^2 > k^2, \end{cases}$$
$$\gamma_p', \gamma_p'' \in \mathbf{R}, \quad \gamma_p', \gamma_p'' > 0.$$

It is not difficult to show that the unique solution of this problem is given by

$$\hat{u}_p(z) = \hat{U}_p \Phi_0(z;p) + \hat{K}_p e^{i\gamma_p a} \Phi_a(z;p) - I_0[\hat{F}_p](z) - I_a[\hat{F}_p](z),$$

where

$$\Phi_0(z;p) = e^{i\gamma_p z} := \begin{cases} e^{i\gamma_p' z}, & \alpha_p^2 < k^2, \\ 1, & \alpha_p^2 = k^2, \\ e^{-\gamma_p'' z}, & \alpha_p^2 > k^2, \end{cases}$$

and

$$\Phi_a(z;p) = \frac{\sinh(\gamma_p z)}{\gamma_p} := \begin{cases} \frac{\sin(\gamma_p' z)}{\gamma_p'}, & \alpha_p^2 < k^2, \\ z, & \alpha_p^2 = k^2, \\ \frac{\sinh(\gamma_p'' z)}{\gamma_p''}, & \alpha_p^2 > k^2, \end{cases}$$

and

$$I_0[\hat{F}_p](z) := \int_0^z \Phi_0(z;p)\Phi_a(s;p)\hat{F}_p(s)\,ds,$$
$$I_a[\hat{F}_p](z) := \int_z^a \Phi_0(s;p)\Phi_a(z;p)\hat{F}_p(s)\,ds.$$

It is straightforward to compute that

$$\partial_z I_0[\hat{F}_p](z) = \Phi_0(z;p)\Phi_a(z;p)\hat{F}_p(z) + \int_0^z (\partial_z\Phi_0(z;p))\Phi_a(s;p)\hat{F}_p(s)\,ds,$$
$$\partial_z I_a[\hat{F}_p](z) = -\Phi_0(z;p)\Phi_a(z;p)\hat{F}_p(z) + \int_z^a \Phi_0(s;p)(\partial_z\Phi_a(z;p))\hat{F}_p(s)\,ds.$$

Noting the cancellation in the *sum* of the terms $\partial_z I_0$ and $\partial_z I_a$, we realize

$$\partial_z \hat{u}_p(z) = \hat{U}_p \partial_z \Phi_0(z;p) + \hat{K}_p e^{i\gamma_p a}\partial_z\Phi_a(z;p) - \tilde{I}_0[\hat{F}_p](z) - \tilde{I}_a[\hat{F}_p](z),$$

where

$$\tilde{I}_0[\hat{F}_p](z) := \int_0^z (\partial_z\Phi_0(z;p))\Phi_a(s;p)\hat{F}_p(s)\,ds,$$
$$\tilde{I}_a[\hat{F}_p](z) := \int_z^a \Phi_0(s;p)(\partial_z\Phi_a(z;p))\hat{F}_p(s)\,ds.$$

If we evaluate this at $z = 0$, we find

$$\begin{aligned} -\partial_z \hat{u}_p(0) &= -\hat{U}_p\partial_z\Phi_0(0;p) - \hat{K}_p e^{i\gamma_p a}\partial_z\Phi_a(0;p) + \tilde{I}_0[\hat{F}_p](0) + \tilde{I}_a[\hat{F}_p](0) \\ &= -\hat{U}_p(i\gamma_p) - \hat{K}_p e^{i\gamma_p a} + \int_0^a e^{i\gamma_p s}\cosh(\gamma_p z)\hat{F}_p(s)\,ds. \end{aligned}$$

With these, it is easy to see that

$$G^{(0)}[U] = -\sum_{p=-\infty}^{\infty} (\partial_z\Phi_0)(0;p)\hat{U}_p e^{i\alpha_p x} = \sum_{p=-\infty}^{\infty} (-i\gamma_p)\hat{U}_p e^{i\alpha_p x}$$

and

$$G^{(a)}[K] = -\sum_{p=-\infty}^{\infty} e^{i\gamma_p a}(\partial_z\Phi_a)(0;p)\hat{K}_p e^{i\alpha_p x} = \sum_{p=-\infty}^{\infty} (-e^{i\gamma_p a})\hat{K}_p e^{i\alpha_p x}$$

and

$$G^{([0,a])}[F] = \sum_{p=-\infty}^{\infty} \int_0^a \left(e^{i\gamma_p s}\cosh(\gamma_p z)\hat{F}_p(s)\,ds \right) e^{i\alpha_p x}.$$

Regarding the estimates, these follow from the asymptotic estimates of $\|\Phi_0\|_{L^2(dz)}$, $\|\Phi_a\|_{L^2(dz)}$, $\|I_0[F]\|_{L^2(dz)}$, $\|I_a[F]\|_{L^2(dz)}$, $\|\tilde{I}_0[F]\|_{L^2(dz)}$, and $\|\tilde{I}_a[F]\|_{L^2(dz)}$. ∎

The analogue in the lower layer is the following result. It is established in an almost identical fashion as Theorem A.1.

Theorem A.2. *Given any integer $s \geq 0$, if $F_2 \in H^{s-1}([0,d] \times [b,0])$, $V \in H^{s+1/2}([0,d])$, and $K_2 \in H^{s-1/2}([0,d])$, then there exists a unique solution $v \in H^{s+1}([0,d] \times [b,0])$ of*

(A.2a) $$\Delta_2 v + k_2^2 v = F_2, \qquad b < z_2 < 0,$$

(A.2b) $$v = V, \qquad z_2 = 0,$$

(A.2c) $$\partial_{z_2} v - T_2[v] = K_2, \qquad z_2 = b,$$

such that

$$\|v\|_{H^{s+1}} \leq C_v \left\{ \|F_2\|_{H^{s-1}} + \|V\|_{H^{s+1/2}} + \|K_2\|_{H^{s-1/2}} \right\}.$$

In addition, if $\tilde{V} = [\partial_{z_2} v]_{z_2=0}$ and we define the DNO

$$\mathcal{J} : (V, K_2, F_2) \to \tilde{V}, \quad \mathcal{J}[V, K_2, F_2] = J^{(0)}[V] + J^{(b)}[K_2] + J^{([b,0])}[F_2],$$

then

$$\begin{aligned} \left\| J^{(0)}[V] \right\|_{H^{s-1/2}} &\leq C_{J^{(0)}} \|V\|_{H^{s+1/2}}, \\ \left\| J^{(b)}[K_2] \right\|_{H^{s-1/2}} &\leq C_{J^{(b)}} \|K_2\|_{H^{s-1/2}}, \\ \left\| J^{([b,0])}[F_2] \right\|_{H^{s-1/2}} &\leq C_{J^{([b,0])}} \|F_2\|_{H^{s-1}}. \end{aligned}$$

In addition, we require the following result on the boundary conditions which couple u and v at the interface $z_1 = z_2 = 0$.

Theorem A.3. *Given any integer $s \geq 0$, if $Q \in H^{s+1/2}([0,d])$ and $R \in H^{s-1/2}([0,d])$, then there exists a unique solution pair $U, V \in H^{s+1/2}([0,d])$ of*

(A.3a) $$U - V = Q,$$

(A.3b) $$G^{(0)}[U] + J^{(0)}[V] = R$$

such that

$$\max \left\{ \|U\|_{H^{s+1/2}}, \|V\|_{H^{s1/2}} \right\} \leq C_0 \left\{ \|Q\|_{H^{s+1/2}} + \|R\|_{H^{s-1/2}} \right\}.$$

Proof. The result follows simply from the well-known expressions for the flat-interface DNOs

$$\begin{aligned} G^{(0)}[U] &= G^{(0)} \left[\sum_{p=-\infty}^{\infty} \hat{U}_p e^{i\alpha_p x} \right] = \sum_{p=-\infty}^{\infty} (-i\gamma_{1,p}) \hat{U}_p e^{i\alpha_p x}, \\ J^{(0)}[U] &= J^{(0)} \left[\sum_{p=-\infty}^{\infty} \hat{V}_p e^{i\alpha_p x} \right] = \sum_{p=-\infty}^{\infty} (-i\gamma_{2,p}) \hat{V}_p e^{i\alpha_p x}, \end{aligned}$$

so that the governing equations become

$$\begin{pmatrix} 1 & -1 \\ (-i\gamma_{1,p}) & (-i\gamma_{2,p}) \end{pmatrix} \begin{pmatrix} \hat{U}_p \\ \hat{V}_p \end{pmatrix} = \begin{pmatrix} \hat{Q}_p \\ \hat{R}_p \end{pmatrix} \quad \forall\ p \in \mathbf{Z}.$$

These are readily solved,

$$\begin{pmatrix} \hat{U}_p \\ \hat{V}_p \end{pmatrix} = \frac{1}{(i\gamma_{1,p}) + (i\gamma_{2,p})} \begin{pmatrix} (-i\gamma_{2,p}) & 1 \\ (i\gamma_{1,p}) & 1 \end{pmatrix} \begin{pmatrix} \hat{Q}_p \\ \hat{R}_p \end{pmatrix},$$

and the $\{\hat{U}_p, \hat{V}_p\}$ have the right decay to verify the conclusions of the theorem. ∎

We can now proceed to our principal result, Theorem 4.3.

Proof. [Theorem 4.3] We begin by rewriting (4.3) as

$$\begin{aligned}
&\text{(A.4a)} && \Delta_1 u + k_1^2 u = F_1, && 0 < z_1 < a, \\
&\text{(A.4b)} && u = U, && z_1 = 0, \\
&\text{(A.4c)} && \partial_{z_1} u - T_1[u] = K_1, && z_1 = a, \\
&\text{(A.4d)} && \Delta_2 v + k_2^2 v = F_2, && b < z_2 < 0, \\
&\text{(A.4e)} && v = V, && z_2 = 0, \\
&\text{(A.4f)} && \partial_{z_2} v - T_2[v] = K_2, && z_2 = b, \\
&\text{(A.4g)} && U - V = \xi, && z_1 = z_2 = 0, \\
&\text{(A.4h)} && \tilde{U} + \tilde{V} = -\nu, && z_1 = z_2 = 0.
\end{aligned}$$

From Theorem A.1, we see that, provided that $F_1 \in H^{s-1}$, $K_1 \in H^{s-1/2}$, and $U \in H^{s+1/2}$, (A.4a)–(A.4c) delivers a unique solution $u \in H^{s+1}$ as desired. The functions F_1 and K_1 in the correct spaces are provided, so we merely need show that U is in $H^{s+1/2}$. In a similar fashion, Theorem A.2 guarantees that if $F_2 \in H^{s-1}$, $K_2 \in H^{s-1/2}$, and $V \in H^{s+1/2}$, then (A.4d)–(A.4f) provides a unique solution $v \in H^{s+1}$. As before, the functions F_2 and K_2 in the correct spaces are provided, so we are left to show that V is in $H^{s+1/2}$.

Thus, all that remains is to consider (A.4g)–(A.4h), which we write in terms of DNOs as

$$\begin{aligned}
&U - V = \xi, \\
&\left(G^{(0)}[U] + G^{(a)}[K_1] + G^{([0,a])}[F_1]\right) + \left(J^{(0)}[V] + J^{(b)}[K_2] + J^{([b,0])}[F_2]\right) = -\nu
\end{aligned}$$

or

$$\begin{aligned}
&U - V = \xi, \\
&G^{(0)}[U] + J^{(0)}[V] = -\nu - G^{(a)}[K_1] - G^{([0,a])}[F_1] - J^{(b)}[K_2] - J^{([b,0])}[F_2].
\end{aligned}$$

Theorem A.3 delivers the required solutions $U, V \in H^{1/2}$ provided that

$$\begin{aligned}
&Q = \xi \in H^{s+1/2}, \\
&R = -\nu - G^{(a)}[K_1] - G^{([0,a])}[F_1] - J^{(b)}[K_2] - J^{([b,0])}[F_2] \in H^{s-1/2},
\end{aligned}$$

both of which are true from (i) our hypotheses on ξ, ν, K_1, K_2, F_1, and F_2 and (ii) the mapping properties of $G^{(a)}$, $G^{([0,a])}$, $J^{(b)}$, and $J^{([b,0])}$ established in Theorems A.1 and A.2. ∎

Appendix B. Regularity of solutions of the weak formulation. We now produce an elliptic regularity theory for solutions of the boundary value problem (3.4). (For the sake of simplicity, we drop the indices $\{p, n\}$.) Noting that $\gamma_{l,p}^2 = k_l^2 - \alpha_p^2$, we split (3.4) into two BVPs: one with inhomogeneous coupling (which, due to the homogeneous Helmholtz equations, we can solve explicitly with Fourier analysis),

$$\partial_{z_1}^2 \check{U}_1 + \gamma_1^2 \check{U}_1 = 0, \quad 0 < z_1 < a, \tag{B.1a}$$
$$\partial_{z_2}^2 \check{U}_2 + \gamma_2^2 \check{U}_2 = 0, \quad b < z_2 < 0, \tag{B.1b}$$
$$\check{U}_1 - \check{U}_2 = \widehat{\xi}_1, \quad z_1 = z_2 = 0, \tag{B.1c}$$
$$\partial_{z_1} \check{U}_1 - \partial_{z_2} \check{U}_2 = \widehat{\xi}_2, \quad z_1 = z_2 = 0, \tag{B.1d}$$
$$\partial_{z_1} \check{U}_1 - i\gamma_1 \check{U}_1 = 0, \quad z_1 = a, \tag{B.1e}$$
$$\partial_{z_2} \check{U}_2 + i\gamma_2 \check{U}_2 = 0, \quad z_2 = b, \tag{B.1f}$$

and one with homogeneous coupling (but inhomogeneous Helmholtz equations),

$$\partial_{z_1}^2 \widetilde{U}_1 + \gamma_1^2 \widetilde{U}_1 = \widehat{R}_1, \quad 0 < z_1 < a, \tag{B.2a}$$
$$\partial_{z_2}^2 \widetilde{U}_2 + \gamma_2^2 \widetilde{U}_2 = \widehat{R}_2, \quad b < z_2 < 0, \tag{B.2b}$$
$$\widetilde{U}_1 - \widetilde{U}_2 = 0, \quad z_1 = z_2 = 0, \tag{B.2c}$$
$$\partial_{z_1} \widetilde{U}_1 - \partial_{z_2} \widetilde{U}_2 = 0, \quad z_1 = z_2 = 0, \tag{B.2d}$$
$$\partial_{z_1} \widetilde{U}_1 - i\gamma_1 \widetilde{U}_1 = \widehat{J}_1, \quad z_1 = a, \tag{B.2e}$$
$$\partial_{z_2} \widetilde{U}_2 + i\gamma_2 \widetilde{U}_2 = \widehat{J}_2, \quad z_2 = b. \tag{B.2f}$$

To study the regularity of solutions of (B.2), we find the variational formulation as in [36, 37],

$$\begin{aligned} \int_b^a \partial_z \widetilde{U} \partial_z \bar{\varphi} - \gamma^2 \int_b^a \widetilde{U} \bar{\varphi} - i\gamma_1 \widetilde{U}_1(a) \bar{\varphi}_1(a) + i\gamma_2 \widetilde{U}_2(b) \bar{\varphi}_2(b) \\ = \widehat{J}_1 \bar{\varphi}_1(a) - \widehat{J}_2 \bar{\varphi}(b) + \int_b^a (-\widehat{R}) \bar{\varphi}, \end{aligned} \tag{B.3}$$

take $\varphi = \widetilde{U}$, and consider the imaginary and real parts, respectively. For the imaginary part, we find

$$-\gamma_1 \left| \widetilde{U}_1(a) \right|^2 - \gamma_2 \left| \widetilde{U}_2(b) \right|^2 = \operatorname{Im} \left\{ (-\widehat{R}, \widetilde{U}) \right\} + \operatorname{Im} \left\{ \widehat{J}_1 \widetilde{U}_1(a) \right\} - \operatorname{Im} \left\{ \widehat{J}_2 \widetilde{U}_2(b) \right\}.$$

With this, we estimate

$$\begin{aligned} \gamma_1 \left| \widetilde{U}_1(a) \right|^2 + \gamma_2 \left| \widetilde{U}_2(b) \right|^2 \le \frac{\kappa_M \delta_1}{2} \left\| \widetilde{U} \right\|_{L^2}^2 + \frac{1}{2\delta_1 \kappa_M} \left\| \widehat{R} \right\|_{L^2}^2 \\ + \frac{\gamma_1}{2} \left| \widetilde{U}_1(a) \right|^2 + \frac{\gamma_2}{2} \left| \widetilde{U}_2(b) \right|^2 + \frac{1}{2\gamma_1} \left| \widehat{J}_1 \right|^2 + \frac{1}{2\gamma_2} \left| \widehat{J}_2 \right|^2, \end{aligned} \tag{B.4}$$

where $\kappa_M := \max(|\gamma_1|, |\gamma_2|)$ and $\delta_1 > 0$ will be chosen later. For the real part, we deduce that

$$\left\| \partial_z \widetilde{U} \right\|_{L^2}^2 - \gamma^2 \left\| \widetilde{U} \right\|_{L^2}^2 = \operatorname{Re} \left\{ (-\widehat{R}, \widetilde{U}) \right\} + \operatorname{Re} \left\{ \widehat{J}_1 \widetilde{U}_1(a) \right\} - \operatorname{Re} \left\{ \widehat{J}_2 \widetilde{U}_2(b) \right\},$$

and this implies

$$\begin{aligned}\left\|\partial_z\widetilde{U}\right\|_{L^2}^2 \leq \kappa_M^2\left\|\widetilde{U}\right\|_{L^2}^2 + \delta_2\kappa_M^2\left|\widetilde{U}_1(a)\right|^2 + \delta_2\kappa_M^2\left|\widetilde{U}_2(b)\right|^2 \\ + \frac{1}{4\delta_2\kappa_M^2}\left|\widehat{J}_1\right|^2 + \frac{1}{4\delta_2\kappa_M^2}\left|\widehat{J}_2\right|^2 + \frac{\delta_3\kappa_M^2}{2}\left\|\widetilde{U}\right\|_{L^2}^2 + \frac{1}{2\delta_3\kappa_M^2}\left\|\widehat{R}\right\|_{L^2}^2,\end{aligned} \tag{B.5}$$

where $\delta_2, \delta_3 > 0$ will also be chosen later. Using (B.4), we deduce that

$$\kappa_m\left(\left|\widetilde{U}_1(a)\right|^2 + \left|\widetilde{U}_2(b)\right|^2\right) \leq \kappa_M\delta_1\left\|\widetilde{U}\right\|_{L^2}^2 + \frac{1}{\delta_1\kappa_M}\left\|\widehat{R}\right\|_{L^2}^2 + \frac{1}{\gamma_1}\left|\widehat{J}_1\right|^2 + \frac{1}{\gamma_2}\left|\widehat{J}_2\right|^2,$$

where $\kappa_m := \min(\gamma_1, \gamma_2)$, and this implies

$$\left|\widetilde{U}_1(a)\right|^2 + \left|\widetilde{U}_2(b)\right|^2 \leq \tau\delta_1\left\|\widetilde{U}\right\|_{L^2}^2 + \frac{1}{\delta_1\kappa_M\kappa_m}\left\|\widehat{R}\right\|_{L^2}^2 + \frac{1}{\gamma_1\kappa_m}\left|\widehat{J}_1\right|^2 + \frac{1}{\gamma_2\kappa_m}\left|\widehat{J}_2\right|^2, \tag{B.6}$$

where $\tau = \kappa_M/\kappa_m$. Using (B.6) and (B.5), we derive that

$$\begin{aligned}\left\|\partial_z\widetilde{U}\right\|_{L^2} \leq \left(\kappa_M^2 + \delta_2\kappa_M^2\tau\delta_1 + \frac{\delta_3\kappa_M^2}{2}\right)\left\|\widetilde{U}\right\|_{L^2}^2 + \left(\frac{\delta_2}{\delta_1}\tau^2 + \frac{1}{2\delta_3\kappa_M^2}\right)\left\|\widehat{R}\right\|_{L^2}^2 \\ + \left(\delta_2\tau^2 + \frac{1}{r\delta_2\kappa_M^2}\right)\left(\left|\widehat{J}_1\right|^2 + \left|\widehat{J}_2\right|^2\right).\end{aligned}$$

Setting $\delta_2 = \delta_3/(2\delta_1\tau)$, we obtain

$$\begin{aligned}\left\|\partial_z\widetilde{U}\right\|_{L^2}^2 \leq \left(\kappa_M^2 + \delta_3\kappa_M^2\right)\left\|\widetilde{U}\right\|_{L^2}^2 + \left(\frac{\delta_3\tau}{2\delta_1^2} + \frac{1}{2\delta_3\kappa_M^2}\right)\left\|\widehat{R}\right\|_{L^2}^2 \\ + \left(\delta_2\tau^2 + \frac{1}{r\delta_2\kappa_M^2}\right)\left(\left|\widehat{J}_1\right|^2 + \left|\widehat{J}_2\right|^2\right).\end{aligned}$$

Regarding (B.3) again, we now take the test function

$$\varphi = 2z\partial_z\widetilde{U} = \begin{cases} 2z\partial_z\widetilde{U}_1, & z \in (0, a) =: I_1, \\ 2z\partial_z\widetilde{U}_2, & z \in (b, 0) =: I_2. \end{cases}$$

Then the weak form (B.3) becomes

$$\begin{aligned}&\left\|\partial_z\widetilde{U}_1\right\|_{L^2(I_1)}^2 + \left\|\partial_z\widetilde{U}_2\right\|_{L^2(I_2)}^2 + a\left|\partial_z\widetilde{U}_1(a)\right|^2 - b\left|\widetilde{\partial}_z U_2(b)\right|^2 \\ &\quad + \gamma_1^2\left\|\widetilde{U}_1\right\|_{L^2(I_1)}^2 + \gamma_2^2\left\|\widetilde{U}_2\right\|_{L^2(I_2)}^2 \\ &\leq \kappa_M^2 M\left(\left|\widetilde{U}_1(a)\right|^2 + \left|\widetilde{U}_2(b)\right|^2\right) + \frac{1}{2}\left\|\partial_z\widetilde{U}\right\|_{L^2}^2 + 8M^2\left\|\widehat{R}\right\|_{L^2}^2 \\ &\quad + \left(\frac{m}{2} + \frac{8}{m}4M^2\kappa_M^2\right)\left(\left|\partial_z\widetilde{U}_1(a)\right|^2 + \left|\partial_z\widetilde{U}_2(b)\right|^2\right) + \frac{32M^2}{m}\left(\left|\widehat{J}_1\right|^2 + \left|\widehat{J}_2\right|^2\right),\end{aligned}$$

where $M := \max(|a|, |b|)$ and $m := \min(|a|, |b|)$. Hence, we deduce that

$$\frac{1}{2}\left\|\partial_z \widetilde{U}\right\|_{L^2}^2 + \frac{m}{2}\left(\left|\partial_z \widetilde{U}_1(a)\right|^2 + \left|\partial_z \widetilde{U}_2(b)\right|^2\right) + \kappa_m^2 \left\|\widetilde{U}\right\|_{L^2}^2$$
$$\leq \left(\kappa_M^2 M + \frac{32M^2\kappa_M^2}{m}\right)\left(\tau\delta_1 \left\|\widetilde{U}\right\|_{L^2}^2 + \frac{1}{\delta_1\kappa_M\kappa_m}\left\|\widehat{R}\right\|_{L^2}^2 + \frac{1}{\kappa_m^2}\left(\left|\widehat{J}_1\right|^2 + \left|\widehat{J}_2\right|^2\right)\right)$$
$$+ 8M^2\left\|\widehat{R}\right\|_{L^2}^2 + \frac{32M^2}{m}\left(\left|\widehat{J}_1\right|^2 + \left|\widehat{J}_2\right|^2\right),$$

and this implies

$$\frac{1}{2}\left\|\partial_z \widetilde{U}\right\|_{L^2}^2 + \frac{m}{2}(\left|\partial_z \widetilde{U}_1(a)\right|^2 + \left|\partial_z \widetilde{U}_2(b)\right|^2)$$
$$+ \left(\kappa_m^2 - \left(\kappa_M^2 M + \frac{32M^2\kappa_M^2}{m}\right)\tau\delta_1\right)\left\|\widetilde{U}\right\|_{L^2}^2$$
$$\leq C\left(\left\|\widehat{R}\right\|_{L^2}^2 + \left|\widehat{J}_1\right|^2 + \left|\widehat{J}_2\right|^2\right).$$

By choosing $\delta_1 < (1/(2\tau^3))(M + 32M^2/m)^{-1}$, we derive our required estimate

$$\left\|\partial_z \widetilde{U}\right\|_{L^2}^2 + \kappa_m \left\|\widetilde{U}\right\|_{L^2}^2 \leq C\left(\left\|\widehat{R}\right\|_{L^2}^2 + \left|\widehat{J}_1\right|^2 + \left|\widehat{J}_2\right|^2\right). \tag{B.7}$$

REFERENCES

[1] H. Ammari and G. Bao, *Maxwell's equations in a perturbed periodic structure*, Model. Comput. Opt. Electromagn., 16 (2002), pp. 99–112.

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

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

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

[5] G. Bao and D. C. Dobson, *On the scattering by a biperiodic structure*, Proc. Amer. Math. Soc., 128 (2000), pp. 2715–2723.

[6] 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.

[7] G. Bao and P. Li, *Convergence analysis in near-field imaging*, Inverse Problems, 30 (2014), p. 085008.

[8] G. Bao and H. Yang, *A least-squares finite element analysis for diffraction problems*, SIAM J. Numer. Anal., 37 (2000), pp. 665–682.

[9] L. M. Brekhovskikh and Y. P. Lysanov, *Fundamentals of Ocean Acoustics*, Springer-Verlag, Berlin, 1982.

[10] O. Bruno and F. Reitich, *Numerical solution of diffraction problems: A method of variation of boundaries*, J. Opt. Soc. Am. A, 10 (1993), pp. 1168–1175.

[11] O. P. Bruno, S. P. Shipman, C. Turc, and S. Venakides, *Superalgebraically convergent smoothly windowed lattice sums for doubly periodic Green functions in three-dimensional space*, Proc. A., 472 (2016), pp. 20160255, 19.

[12] J. Chandezon, D. Maystre, and G. Raoult, *A new theoretical method for diffraction gratings and its numerical application*, J. Opt., 11 (1980), pp. 235–241.

[13] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, 2nd ed., Springer-Verlag, Berlin, 1998.

[14] D. C. Dobson, *A variational method for electromagnetic diffraction in biperiodic structures*, RAIRO Modél. Math. Anal. Numér., 28 (1994), pp. 419–439.

[15] L. C. Evans, *Partial Differential Equations*, 2nd ed., American Mathematical Society, Providence, RI, 2010.

[16] Q. Fang, D. P. Nicholls, and J. Shen, *A stable, high-order method for three-dimensional bounded-obstacle scattering*, J. Comput. Phys., 224 (2007), pp. 1145–1169.
[17] C. Godrèche, ed., *Solids Far from Equilibrium*, Cambridge University Press, Cambridge, 1992.
[18] Y. He, D. P. Nicholls, and J. Shen, *An efficient and stable spectral method for electromagnetic scattering from a layered periodic structure*, Journal of Computational Physics, 231 (2012), pp. 3007–3022.
[19] J. Homola, *Surface plasmon resonance sensors for detection of chemical and biological species*, Chem. Rev., 108 (2008), pp. 462–493.
[20] Y. Hong and D. P. Nicholls, *A high-order perturbation of surfaces method for scattering of linear waves by periodic multiply layered gratings in two and three dimensions*, J. Comput. Phys., 345 (2017), pp. 162–188.
[21] Y. Hong and D. P. Nicholls, *A stable high-order perturbation of surfaces method for numerical simulation of diffraction problems in triply layered media*, Journal of Computational Physics, 330 (2017), pp. 1043–1068.
[22] Y. Hong and D. P. Nicholls, *A high-order perturbation of surfaces method for vector electromagnetic scattering by doubly layered periodic crossed gratings*, J. Comput. Phys., 372 (2018), pp. 748–772.
[23] R. Kress, *Linear Integral Equations*, 3rd ed., Springer-Verlag, Berlin, 2014.
[24] J. Lai, M. Kobayashi, and A. Barnett, *A fast and robust solver for the scattering from a layered periodic structure containing multi-particle inclusions*, J. Comput. Phys., 298 (2015), pp. 194–208.
[25] A. Malcolm and D. P. Nicholls, *A field expansions method for scattering by periodic multilayered media*, J. Acoust. Soc. Amer., 129 (2011), pp. 1783–1793.
[26] F. Natterer and F. Wübbeling, *Mathematical Methods in Image Reconstruction*, SIAM Monographs on Mathematical Modeling and Computation, SIAM, Philadelphia, 2001.
[27] D. P. Nicholls, S.-H. Oh, T. W. Johnson, and F. Reitich, *Launching surface plasmon waves via vanishingly small periodic gratings*, J. Opt. Soc. Amer., A, 33 (2016), pp. 276–285.
[28] D. P. Nicholls and F. Reitich, *A new approach to analyticity of Dirichlet-Neumann operators*, Proc. Roy. Soc. Edinburgh Sect. A, 131 (2001), pp. 1411–1433.
[29] D. P. Nicholls and F. Reitich, *Shape deformations in rough surface scattering: Improved algorithms*, J. Opt. Soc. Amer. A, 21 (2004), pp. 606–621.
[30] D. P. Nicholls and J. Shen, *A stable, high-order method for two-dimensional bounded-obstacle scattering*, SIAM J. Sci. Comput., 28 (2006), pp. 1398–1419.
[31] D. P. Nicholls and J. Shen, *A rigorous numerical analysis of the transformed field expansion method*, SIAM J. Numer. Anal., 47 (2009), pp. 2708–2734.
[32] R. Petit, ed., *Electromagnetic Theory of Gratings*, Springer-Verlag, Berlin, 1980.
[33] N. A. Phillips, *A coordinate system having some special advantages for numerical forecasting*, J. Atmos. Sci., 14 (1957), pp. 184–185.
[34] H. Raether, *Surface Plasmons on Smooth and Rough Surfaces and on Gratings*, Springer-Verlag, Berlin, 1988.
[35] J. Shen, T. Tang, and L.-L. Wang, *Spectral Methods*, Springer Series in Computational Mathematics 41, Springer-Verlag, Berlin, 2011.
[36] J. Shen and L.-L. Wang, *Spectral approximation of the Helmholtz equation with high wave numbers*, SIAM J. Numer. Anal., 43 (2005), pp. 623–644.
[37] J. Shen and L.-L. Wang, *Analysis of a spectral-Galerkin approximation to the Helmholtz equation in exterior domains*, SIAM J. Numer. Anal., 45 (2007), pp. 1954–1978.
[38] P. J. Shull, *Nondestructive Evaluation: Theory, Techniques, and Applications*, Marcel Dekker, New York, 2002.
[39] J. Virieux and S. Operto, *An overview of full-waveform inversion in exploration geophysics*, Geophysics, 74 (2009), pp. WCC1–WCC26.