Journal of Computational Physics 345 (2017) 162–188

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

A high-order perturbation of surfaces method for scattering of linear waves by periodic multiply layered gratings in two and three dimensions

Youngjoon Hong, David P. Nicholls *

Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago, Chicago, IL 60607, USA

A R T I C L E I N F O

Article history:
Received 2 February 2017
Received in revised form 8 May 2017
Accepted 9 May 2017
Available online 12 May 2017

Keywords:
High-order spectral methods
Linear wave scattering
Periodic multiply layered media
High-order perturbation of surfaces methods

A B S T R A C T

The capability to rapidly and robustly simulate the scattering of linear waves by periodic, multiply layered media in two and three dimensions is crucial in many engineering applications. In this regard, we present a High-Order Perturbation of Surfaces method for linear wave scattering in a multiply layered periodic medium to find an accurate numerical solution of the governing Helmholtz equations. For this we truncate the bi-infinite computational domain to a finite one with artificial boundaries, above and below the structure, and enforce transparent boundary conditions there via Dirichlet–Neumann Operators. This is followed by a Transformed Field Expansion resulting in a Fourier collocation, Legendre–Galerkin, Taylor series method for solving the problem in a transformed set of coordinates. Assorted numerical simulations display the spectral convergence of the proposed algorithm.

1. Introduction

The scattering of linear waves by periodic, multiply layered media in two and three dimensions arises in many engineering and physics applications. Examples exist in materials science [1], geophysics [2,3], imaging [4], oceanography [5], and nanoplasmonics [6–8]. It is clear that the capability to rapidly and robustly simulate such interactions with high accuracy is of fundamental importance to practical applications. The most popular approaches to these problems are volumetric in nature and include the Finite Difference Method [9,10], the Finite Element Method [11,12], the Discontinuous Galerkin Method [13], the Spectral Element Method [14], and Spectral Methods [15–17]. However, these methods are greatly disadvantaged with an unnecessarily large number of unknowns for layered media problems (see, e.g., the discussion in [18,19]). Interfacial methods based on Integral Equations (IEs) [20–24] are a natural alternative but these also face several challenges. First, for the parameterized problems we consider here (characterized by the interface height/slope ε) a new simulation must be run for every configuration in the family of interest. Additionally, for every invocation there is a dense, non-symmetric positive definite system of linear equations generated by the IEs which must be inverted.

High-Order Perturbation of Surfaces (HOPS) methods can avoid these concerns. These highly accurate algorithms, based upon the low-order theories of Rayleigh [25] and Rice [26], were first developed by Bruno and Reitich [27] and later enhanced and stabilized by Nicholls and Reitich [28], and Nicholls and Malcolm [29]. HOPS approaches are compelling as they maintain the advantageous properties of classical IE formulations (e.g., surface formulation, exact enforcement of

* Corresponding author.
E-mail addresses: hongy@uic.edu (Y. Hong), davidn@uic.edu (D.P. Nicholls).

http://dx.doi.org/10.1016/j.jcp.2017.05.017
0021-9991/

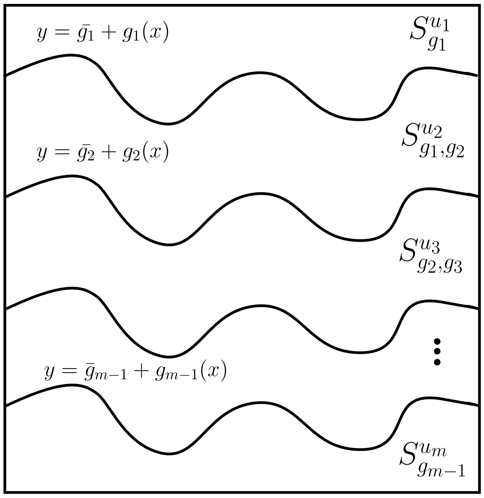

Fig. 1. A depiction of a multiply layered grating structure.

far-field boundary conditions, and quasiperiodic boundary conditions) while avoiding their shortcomings. For instance, in the setting of parameterized interfaces, scattering returns from an entire family of such configurations can be obtained with negligible additional cost beyond a single simulation. In addition, the perturbative nature of the algorithm allows one to invert, at every perturbation order, the *same* sparse (after Fast Fourier Transformation) operator corresponding to the flat-interface, order-zero approximation of the problem.

The method we present here is a generalization of the work of Nicholls and Shen on irregular, bounded obstacles in two [30] and three dimensions [31]. (Nicholls and Shen then gave a rigorous numerical analysis of a wide class of HOPS schemes in [32].) Subsequent to this He, Nicholls, and Shen [33] devised a non-trivial extension to the case of periodic gratings separating two materials of different dielectric constants. In [34] the authors made the further extension to the case of three layers. This work addressed the additional complication of waves propagating both up and down in a vertically bounded layer in between. The novelties of the current contribution include the highly non-trivial extension to three dimensions and an arbitrary number of layers. In addition, we not only demonstrate its applicability to *large* deformations with the use of Padé summation [35], but we also state and briefly prove the existence, uniqueness, and analyticity of solutions to our governing equations. A careful numerical analysis of the convergence of our scheme to these solutions we save for a forthcoming publication.

When considering real-world applications such as seismic imaging, underwater acoustics, and the modeling of plasmonic nanostructures, the multiply layered nature of the structure is essential. In the current contribution, we study linear waves interacting with periodic gratings separating several layers of materials with different dielectric constants; see, e.g., Fig. 1. The first step of our method is to equivalently reformulate the governing equations on a *bounded* domain with artificial boundaries and transparent boundary conditions (implemented via Dirichlet–Neumann Operators) above and below the interfaces of the structure. With a nonlinear change of variables, the computational domain can be transformed to one with flat interfaces between material layers, at the cost of nonlinear terms in the governing equations. Using boundary perturbations, we express the scattered field in a Taylor series and write a sequence of linear problems to be solved at each perturbation order for higher order corrections; this is the essence of the Transformed Field Expansions (TFE) algorithm. Due to inhomogeneities in the governing Helmholtz equations arising from the change of variables, a vertical discretization is required which we implement with a spectrally accurate Legendre–Galerkin method [36,37]. This Legendre–Galerkin algorithm is slightly unorthodox in that the standard basis is supplemented with additional connecting basis functions across the layer boundaries [34].

The article is organized as follows: In Section 2 the governing equations for linear waves interacting with a periodic multiply layered structure are given. The TFE method in this setting is described in Section 3, together with a discussion of the Legendre–Galerkin scheme we implemented for the vertical discretization in Sections 4 and 5. An assortment of numerical experiments are presented in Section 6 including simulations of smooth and rough interfaces, of small and large size, in two and three dimensions.

2. Governing equations

In this section, we describe the governing equations of linear waves scattered by a multiply layered medium. The geometry is depicted in Fig. 1. Dielectrics occupy each of the m-many domains: The upper material fills the region

$$S^{u_1}_{g_1} := \{y > \bar{g}_1 + g_1(x)\},$$

the middle layers occupy

$$S^{u_j}_{g_{j-1},g_j} := \{\bar{g}_j + g_j < y < \bar{g}_{j-1} + g_{j-1}\},$$

for $2 \le j \le m-1$, and the lower dielectric fills

$$S^{u_m}_{g_{m-1}} := \{y < \bar{g}_{m-1} + g_{m-1}(x)\},$$

where in all of these domains $x \in \mathbf{R}^2$. The grating interfaces are described by the functions $\{g_j\}$, $1 \le j \le m-1$, which we assume to be $d = (d_1, d_2)$-periodic

$$g_j(x+d) = g_j(x), \quad 1 \le j \le m-1,$$

and the $\bar{g}_j$ are constants. We define the scattered fields

$$v_1 = v_1(x,y), \quad v_j = v_j(x,y), \quad v_m = v_m(x,y),$$

in $S^{u_1}_{g_1}$, $S^{u_j}_{g_{j-1},g_j}$, and $S^{u_m}_{g_{m-1}}$, respectively for $2 \le j \le m-1$. The incident radiation in the upper layer is specified by $v^{\text{inc}} = \exp(i\alpha \cdot x - i\beta y)$.

We seek outgoing, $\alpha = (\alpha_1, \alpha_2)$-quasiperiodic solutions of the following system of Helmholtz equations

$$\Delta v_1 + k_1^2 v_1 = 0, \qquad \text{in } S^{u_1}_{g_1}, \tag{1a}$$
$$\Delta v_j + k_j^2 v_j = 0, \qquad \text{in } S^{u_j}_{g_{j-1},g_j}, \quad 2 \le j \le m-1, \tag{1b}$$
$$\Delta v_m + k_m^2 v_m = 0, \qquad \text{in } S^{u_m}_{g_{m-1}}, \tag{1c}$$
$$v_1 - v_2 = -v^{\text{inc}}, \qquad \text{at } y = \bar{g}_1 + g_1(x), \tag{1d}$$
$$\partial_{N_{g_1}} v_1 - \tau_1^2 \partial_{N_{g_1}} v_2 = -\partial_{N_{g_1}} v^{\text{inc}}, \qquad \text{at } y = \bar{g}_1 + g_1(x), \tag{1e}$$
$$v_j - v_{j+1} = 0, \qquad \text{at } y = \bar{g}_j + g_j(x), \quad 2 \le j \le m-1, \tag{1f}$$
$$\partial_{N_{g_j}} v_j - \tau_j^2 \partial_{N_{g_{j+1}}} v_2 = 0, \qquad \text{at } y = \bar{g}_j + g_j(x), \quad 2 \le j \le m-1, \tag{1g}$$
$$\text{OWC}[v_1] = 0, \qquad y \to \infty, \tag{1h}$$
$$\text{OWC}[v_m] = 0, \qquad y \to -\infty, \tag{1i}$$

where k_j is the wavenumber in layer j, N_{g_j} is an upward pointing normal vector, and

$$\tau_j^2 = \begin{cases} 1, & \text{for Transverse Electric (TE) polarization,} \\ (k_j/k_{j+1})^2, & \text{for Transverse Magnetic (TM) polarization,} \end{cases}$$

for $1 \le j \le m-1$. Here, “OWC” stands for the outgoing (upward/downward propagating) wave condition which we make precise presently.

2.1. Transparent boundary conditions

The usual procedure when implementing the TFE method is to truncate the bi-infinite domain into one of finite extent. For this we introduce artificial boundaries above and below the structure, and enforce transparent boundary conditions to equivalently solve (1). Introducing the planes $y = a > \bar{g}_1 + |g_1|_\infty$ and $y = b < \bar{g}_{m-1} + |g_{m-1}|_\infty$, we define the domains

$$S^a := \{y > a\}, \quad S^b := \{y < b\},$$
$$S^{a,u_1}_{g_1} := \{\bar{g}_1 + g_1(x) < y < a\}, \quad S^{u_m,b}_{g_{m-1}} := \{b < y < \bar{g}_{m-1} + g_{m-1}(x)\};$$

see Fig. 2 for more details. Transparent boundary conditions can be enforced with Dirichlet–Neumann Operators (DNOs) derived from the Rayleigh expansions. These are relevant as they solve the problem on S^a and S^b explicitly upon specification of Dirichlet data at the artificial boundaries, $\{y = a\}$ and $\{y = b\}$. More specifically, it is known [38] that

$$v_1^+(x,y) = \sum_{p=-\infty}^{\infty} \hat{\zeta}_p e^{i\alpha_p \cdot x + i\beta_p^1 (y-a)}, \quad y > a,$$

and

$$v_m^-(x,y) = \sum_{p=-\infty}^{\infty} \hat{\psi}_p e^{i\alpha_p \cdot x + i\beta_p^m (b-y)}, \quad y < b,$$

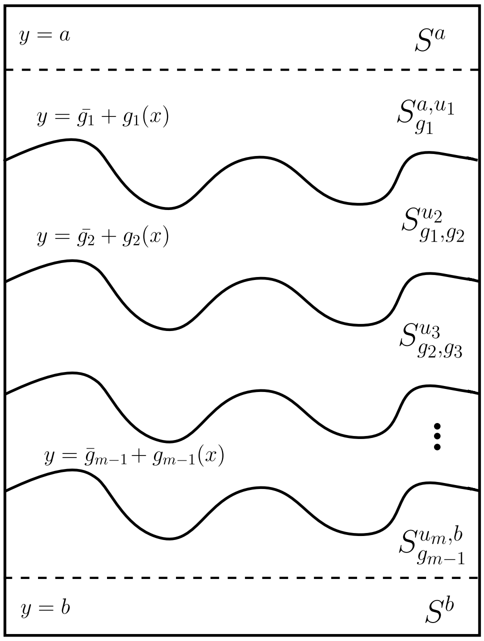

Fig. 2. A depiction of the multiply layered grating structure with artificial boundaries at $y = a$ and $y = b$.

where $p = (p_1, p_2)$, our summation notation connotes a double sum, $\alpha_{s,p} = \alpha_s + (2\pi/d_s)p_s$ for $s \in \{1, 2\}$, and $\alpha_p = (\alpha_{1,p_1}, \alpha_{2,p_2})$. For $j \in \{1, m\}$

$$\beta_p^j := \begin{cases} \sqrt{k_j^2 - |\alpha_p|^2}, & p \in \mathcal{U}^j, \\ i\sqrt{|\alpha_p|^2 - k_j^2}, & p \notin \mathcal{U}^j, \end{cases}$$

and the set of propagating modes is

$$\mathcal{U}^j := \left\{ p \mid |\alpha_p|^2 < k_j^2 \right\}.$$

These solutions satisfy the Dirichlet conditions

$$v_1^+(x,a) = \sum_{p=-\infty}^{\infty} \hat{\zeta}_p e^{i\alpha_p \cdot x} = \zeta(x), \quad v_m^-(x,b) = \sum_{p=-\infty}^{\infty} \hat{\psi}_p e^{i\alpha_p \cdot x} = \psi(x).$$

From these we can compute the Neumann data at the artificial boundaries,

$$\partial_y v_1^+(x,a) = \sum_{p=-\infty}^{\infty} (i\beta_p^1)\hat{\zeta}_p e^{i\alpha_p \cdot x}, \quad \partial_y v_m^-(x,b) = \sum_{p=-\infty}^{\infty} (-i\beta_p^m)\hat{\psi}_p e^{i\alpha_p \cdot x},$$

and thus we define the DNOs

$$T_1[\zeta] := \sum_{p=-\infty}^{\infty} (i\beta_p^1)\hat{\zeta}_p e^{i\alpha_p \cdot x}, \quad T_m[\psi] := \sum_{p=-\infty}^{\infty} (-i\beta_p^m)\hat{\psi}_p e^{i\alpha_p \cdot x},$$

which are order-one Fourier multipliers.

Using the DNOs at the artificial boundaries we write (1) *equivalently* on the bounded domain $\{b < y < a\}$,

$$\Delta v_1 + k_1^2 v_1 = 0, \qquad \text{in } S_{g_1}^{a,u_1}, \tag{2a}$$

$$\Delta v_j + k_j^2 v_j = 0, \qquad \text{in } S_{g_{j-1},g_j}^{u_j}, \quad 2 \le j \le m-1, \tag{2b}$$

$$\Delta v_m + k_m^2 v_m = 0, \qquad \text{in } S_{g_{m-1}}^{u_m,b}, \tag{2c}$$

$$v_1 - v_2 = -v^{\text{inc}}, \qquad \text{at } y = \bar{g}_1 + g_1(x), \tag{2d}$$

$$\begin{aligned}
&\partial_{N_{g_1}} v_1 - \tau_1^2 \partial_{N_{g_1}} v_2 = -\partial_{N_{g_1}} v^{\mathrm{inc}}, && \text{at } y = \bar{g}_1 + g_1(x), && \text{(2e)}\\
&v_j - v_{j+1} = 0, && \text{at } y = \bar{g}_j + g_j(x), \quad 2 \le j \le m-1, && \text{(2f)}\\
&\partial_{N_{g_j}} v_j - \tau_j^2 \partial_{N_{g_{j+1}}} v_j = 0, && \text{at } y = \bar{g}_j + g_j(x), \quad 2 \le j \le m-1, && \text{(2g)}\\
&\partial_y v_1 - T_1[v_1] = 0, && \text{at } y = a, && \text{(2h)}\\
&\partial_y v_m - T_m[v_m] = 0, && \text{at } y = b, && \text{(2i)}
\end{aligned}$$

which is our governing problem.

3. Transformed field expansions

We now recall the TFE method. This algorithm begins with a domain flattening change of variables (also known as σ-coordinates [39] in the geophysical literature and the C-method [40] 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 (nonlinear) change of variables

$$\begin{aligned}
&x' = x,\\
&y_1 = a\left(\frac{y - (\bar{g}_1 + g_1)}{a - (\bar{g}_1 + g_1)}\right) + \bar{g}_1\left(\frac{a - y}{a - (\bar{g}_1 + g_1)}\right),\\
&\qquad \bar{g}_1 + g_1 < y < a,\\
&y_m = b\left(\frac{y - (\bar{g}_{m-1} + g_{m-1})}{a - (\bar{g}_{m-1} + g_{m-1})}\right) + \bar{g}_{m-1}\left(\frac{b - y}{b - (\bar{g}_{m-1} + g_{m-1})}\right),\\
&\qquad b < y < \bar{g}_{m-1} + g_{m-1},\\
&y_j = \bar{g}_j\left(\frac{y - (\bar{g}_{j-1} + g_{j-1})}{(\bar{g}_j + g_j) - (\bar{g}_{j-1} + g_{j-1})}\right) + \bar{g}_{j-1}\left(\frac{(\bar{g}_j + g_j) - y}{(\bar{g}_j + g_j) - (\bar{g}_{j-1} + g_{j-1})}\right),\\
&\qquad \bar{g}_j + g_j < y < \bar{g}_{j-1} + g_{j-1},
\end{aligned}$$

and define

$$u_j = u_j(x', y) = v_j(x(x'), y(x', y_1, \dots, y_m)), \quad 1 \le j \le m.$$

By setting $\bar{g}_m := b$ and $\bar{g}_0 := a$, we find

$$b = \bar{g}_m \le y_m \le \bar{g}_{m-1} \le y_{m-1} \le \bar{g}_{m-2} \le \dots \le \bar{g}_1 \le y_1 \le \bar{g}_0 = a.$$

Using this change of variables, a long computation (see Appendix A) transforms (2) to the following system of equations

$$\begin{aligned}
&\Delta_{x,y_j} u_j + k_j^2 u_j = R_j, && \bar{g}_j < y_j < \bar{g}_{j-1}, \quad 1 \le j \le m, && \text{(3a)}\\
&u_1 - u_2 = -e^{i\alpha\cdot x} e^{-i\beta(\bar{g}_1 + g_1(x))}, && y_1 = y_2 = \bar{g}_1, && \text{(3b)}\\
&u_j - u_{j+1} = 0, && y_j = y_{j+1} = \bar{g}_j, \quad 2 \le j \le m-1, && \text{(3c)}\\
&\partial_{y_1} u_1 - \tau_1^2 \partial_{y_2} u_2 = J_1, && y_1 = y_2 = \bar{g}_1, && \text{(3d)}\\
&\partial_{y_j} u_j - \tau_j^2 \partial_{y_{j+1}} u_{j+1} = J_j, && y_j = y_{j+1} = \bar{g}_j, \quad 2 \le j \le m-1, && \text{(3e)}\\
&\partial_{y_1} u_1 - T_1[u_1] = -\frac{g_1}{a - \bar{g}_1} T_1[u_1], && y_1 = a, && \text{(3f)}\\
&\partial_{y_m} u_m - T_m[u_m] = -\frac{g_{m-1}}{b - \bar{g}_{m-1}} T_m[u_m], && y_m = b, && \text{(3g)}
\end{aligned}$$

where we have dropped the primes for notational convenience. We refer the reader to Appendix A for the specific formulas for the right hand sides R_j and J_j, (A.2) and (A.4), respectively.

3.2. A high-order perturbation of surfaces method

Upon consideration of deformations of the form

$$g_j(x) = \varepsilon f_j(x),$$

we introduce a HOPS method to study the transformed equations in (3). Making expansions of the fields

$$u_j = \sum_{n=0}^{\infty} u_{j,n}(x,y)\varepsilon^n, \quad 1 \le j \le m, \tag{4}$$

we find that (3) demands at each order of ε

$$\Delta_{x,y_j} u_{j,n} + k_j^2 u_{j,n} = R_{j,n}, \qquad \bar{g}_j < y_j < \bar{g}_{j-1}, \quad 1 \le j \le m, \tag{5a}$$

$$u_{1,n} - u_{2,n} = -e^{i\alpha\cdot x} e^{-i\beta\bar{g}_1} \frac{(-i\beta f_1)^n}{n!}, \qquad y_1 = y_2 = \bar{g}_1, \tag{5b}$$

$$u_{j,n} - u_{j+1,n} = 0, \qquad y_j = y_{j+1} = \bar{g}_j, \quad 2 \le j \le m-1, \tag{5c}$$

$$\partial_{y_1} u_{1,n} - \tau_1^2 \partial_{y_2} u_{2,n} = J_{1,n}, \qquad y_1 = y_2 = \bar{g}_1, \tag{5d}$$

$$\partial_{y_j} u_{j,n} - \tau_j^2 \partial_{y_{j+1}} u_{j+1,n} = J_{j,n}, \qquad y_j = y_{j+1} = \bar{g}_j, \quad 2 \le j \le m-1, \tag{5e}$$

$$\partial_{y_1} u_{1,n} - T_1[u_{1,n}] = -\frac{f_1}{a - \bar{g}_1} T_1[u_{1,n-1}], \qquad y_1 = a, \tag{5f}$$

$$\partial_{y_m} u_{m,n} - T_m[u_{m,n}] = -\frac{f_{m-1}}{b - \bar{g}_{m-1}} T_m[u_{m,n-1}], \qquad y_m = b. \tag{5g}$$

Again, the reader should see Appendix A for the particular forms for the $R_{j,n}$ and $J_{j,n}$, (A.5) and (A.6), respectively.

Considering the quasiperiodicity of solutions, we use the generalized Fourier (Floquet) series expansions

$$u_{j,n}(x,y) = \sum_{p=-\infty}^{\infty} u_{j,n}^p(y) e^{i\alpha_p\cdot x}, \qquad 1 \le j \le m, \tag{6a}$$

$$R_{j,n}(x,y) = \sum_{p=-\infty}^{\infty} R_{j,n}^p(y) e^{i\alpha_p\cdot x}, \qquad 1 \le j \le m, \tag{6b}$$

$$J_{j,n}(x) = \sum_{p=-\infty}^{\infty} J_{j,n}^p e^{i\alpha_p\cdot x}, \qquad 1 \le j \le m-1. \tag{6c}$$

Inserting these expansions into (5), we obtain, using $(\beta_p^j)^2 = k_j^2 - |\alpha_p|^2$,

$$\partial_y^2 u_{j,n}^p(y) + (\beta_p^j)^2 u_{j,n}^p(y) = R_{j,n}^p(y), \qquad \bar{g}_j < y < \bar{g}_{j-1}, \quad 1 \le j \le m, \tag{7a}$$

$$u_{1,n}^p(\bar{g}_1) - u_{2,n}^p(\bar{g}_1) = -e^{-i\beta\bar{g}_1} \frac{(-i\beta f_1)^n}{n!} =: \zeta_n^p, \tag{7b}$$

$$u_{j,n}^p(\bar{g}_j) - u_{j+1,n}^p(\bar{g}_j) = 0, \qquad 2 \le j \le m-1, \tag{7c}$$

$$\partial_{y_1} u_{1,n}^p(\bar{g}_1) - \tau_1^2 \partial_{y_2} u_{2,n}^p(\bar{g}_1) = J_{1,n}^p, \tag{7d}$$

$$\partial_{y_j} u_{j,n}^p(\bar{g}_j) - \tau_j^2 \partial_{y_{j+1}} u_{j+1,n}^p(\bar{g}_j) = J_{j,n}^p, \qquad 2 \le j \le m-1, \tag{7e}$$

$$\partial_{y_1} u_{1,n}^p - i\beta_p^1 u_{1,n}^p = B_{1,n}^p, \qquad y_1 = a, \tag{7f}$$

$$\partial_{y_m} u_{m,n}^p + i\beta_p^m u_{m,n}^p = B_{m,n}^p, \qquad y_m = b, \tag{7g}$$

and

$$B_{1,n}^p := -\frac{f_1}{a - \bar{g}_1} (i\beta_p^1)[u_{1,n-1}^p], \tag{7h}$$

$$B_{m,n}^p := -\frac{f_{m-1}}{b - \bar{g}_{m-1}} (-i\beta_p^m)[u_{m,n-1}^p]. \tag{7i}$$

3.3. Existence, uniqueness, and analyticity of solutions

With these recursions we can now present a brief sketch of the proof that the scattered field u depends analytically upon the deformation parameter ε. Before beginning we must define an "admissible configuration" where we can expect the existence of a unique solution. For this we appeal to the classical results of Keller [41], later extended in the "Integrated Solution Method" of Zhang [42,43] (see also [44]).

In order to establish existence and uniqueness of solutions to (5) we express solutions in the form (6) which delivers (7). We rewrite these in Keller's notation

$$\mathbf{u}'(y) = \mathbf{M}(y)\mathbf{u}(y) + \mathbf{g}(y), \qquad a < y < b,$$
$$\sum_{j=0}^{m} \mathbf{B}_j \mathbf{u}(\bar{g}_j) = \mathbf{b},$$

where

$$\mathbf{u} = \begin{pmatrix} u_{1,n}^p \\ \partial_y u_{1,n}^p \\ u_{2,n}^p \\ \partial_y u_{2,n}^p \\ \vdots \\ u_{m,n}^p \\ \partial_y u_{m,n}^p \end{pmatrix}, \quad \mathbf{g} = \begin{pmatrix} 0 \\ R_{1,n}^p \\ 0 \\ R_{2,n}^p \\ \vdots \\ 0 \\ R_{m,n}^p \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} B_{1,n}^p \\ \zeta_n^p \\ J_{1,n}^p \\ 0 \\ J_{2,n}^p \\ \vdots \\ 0 \\ J_{m,n}^p \\ B_{m,n}^p \end{pmatrix},$$

and

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_1 & 0 & 0 & \cdots & 0 \\ 0 & \mathbf{M}_2 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & 0 & \mathbf{M}_{m-1} & 0 \\ 0 & \cdots & 0 & 0 & \mathbf{M}_m \end{pmatrix}, \quad \mathbf{M}_j = \begin{pmatrix} 0 & 1 \\ -(\beta_p^j)^2 & 0 \end{pmatrix},$$

and

$$\mathbf{B}_0 = \begin{pmatrix} -(i\beta_p^1) & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix},$$

and

$$\mathbf{B}_j = \begin{pmatrix} 0 & 0 & & & \cdots & & & & 0 & 0 \\ \vdots & & & & & & & & & \vdots \\ 0 & 0 & & & \cdots & & & & 0 & 0 \\ 0 & \cdots & 0 & 1 & 0 & -1 & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 1 & 0 & -\tau_j^2 & 0 & \cdots & 0 \\ 0 & 0 & & & \cdots & & & & 0 & 0 \\ \vdots & & & & & & & & & \vdots \\ 0 & 0 & & & \cdots & & & & 0 & 0 \end{pmatrix},$$

and

$$\mathbf{B}_{m+1} = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & (i\beta_p^m) & 1 \end{pmatrix}.$$

Let $\mathbf{\Phi}(y)$ be the fundamental matrix solution of the system

$$\mathbf{\Phi}'(y) = \mathbf{M}(y)\mathbf{\Phi}(y), \quad \mathbf{\Phi}(0) = I,$$

where I is the $(2m) \times (2m)$ identity matrix. Keller shows [41] that the two-point value problem above has a unique solution if and only if

$$Q := \sum_{j=0}^{m} \mathbf{B}_j \mathbf{\Phi}(\bar{g}_j, \bar{g}_0) \tag{8}$$

is non-singular. It is not hard to show that

$$e^{\mathbf{M}y} = \mathbf{\Phi}(y) = \begin{pmatrix} \mathbf{\Phi}_1 & 0 & 0 & \vdots & 0 \\ 0 & \mathbf{\Phi}_2 & 0 & \vdots & 0 \\ \vdots & & & & \vdots \\ 0 & \cdots & 0 & \mathbf{\Phi}_{m-1} & 0 \\ 0 & \cdots & 0 & 0 & \mathbf{\Phi}_m \end{pmatrix}, \quad \mathbf{\Phi}_j(y) = e^{\mathbf{M}_j y},$$

and

$$\mathbf{\Phi}_j(y) = \begin{cases} \begin{pmatrix} 1 & y \\ 0 & 1 \end{pmatrix}, & \beta_p^j = 0, \\ \begin{pmatrix} \cos(\beta_p^j y) & \sin(\beta_p^j y)/\beta_p^j \\ -\beta_p^j \sin(\beta_p^j y) & \cos(\beta_p^j y) \end{pmatrix}, & p \in \mathcal{U}^j, \\ \begin{pmatrix} \cosh(Im[\beta_p^j]y) & \sinh(Im[\beta_p^j]y)/Im[\beta_p^j] \\ Im[\beta_p^j]\sinh(Im[\beta_p^j]y) & \cosh(Im[\beta_p^j]y) \end{pmatrix}, & p \notin \mathcal{U}^j. \end{cases}$$

With this notation we can now define a τ-allowable configuration.

Definition 1. A *configuration*, $\mathcal{C}$, of the structure is the $(2m+4)$-tuple

$$\mathcal{C} := (d_1, d_2, \alpha_1, \alpha_2, k_1, \ldots, k_m, \beta^1, \ldots, \beta^m),$$

to which we can assign a matrix $Q = Q(\mathcal{C})$, cf. (8). For any $\tau > 0$, the set of *τ-allowable configurations* is defined by

$$\mathcal{L}_\tau := \{\mathcal{C} \mid |\det(Q(\mathcal{C}))| > \tau\}. \tag{9}$$

Theorem 2. *Given an integer $s \geq 0$ and any $\sigma > 0$, if $f_j \in C^{s+3/2+\sigma}(d)$ and the configuration is τ-allowable (cf., (9)) then there exists a unique solution (4) of (5) satisfying*

$$\left\| \partial_y u_n(x,0) \right\|_{H^s} \leq K B^n, \tag{10a}$$

$$\sum_{p=-\infty}^{\infty} |p|^{2s+1} \left\| \partial_y \hat{u}_{j,n}(p,0) \right\|_{L^2(dy)}^2 \leq K^2 B^{2n}, \tag{10b}$$

$$\sum_{p=-\infty}^{\infty} |p|^{2s+3} \left\| \hat{u}_{j,n}(p,0) \right\|_{L^2(dy)}^2 \leq K^2 B^{2n}, \tag{10c}$$

for any $B > C(s) \max_{1 \leq j \leq m} \left| f_j \right|_{C^{s+3/2+\sigma}}$ where K is a universal constant and C only depends on s.

Proof. In short, the proof follows that given for Theorem 5.1 in [45] which consists of an elliptic estimate (Lemmas 5.2 & 5.3) followed by a recursive lemma (Lemma 5.4). The former is built upon the classical elliptic theory found in Ladyzhenskaya and Ural'tseva [46] consisting of two assertions: Existence and uniqueness of solutions in appropriate Sobolev classes. Uniqueness is guaranteed by our demand that the configuration be τ allowable. Existence of solutions with appropriate estimates can be shown in much the same was as in [45]. The recursive estimate shows that, given that (10) is true for all $n < N$, then the right-hand-sides of (5) can be bounded in the following way, e.g.,

$$\left\| R_{j,n} \right\|_{H^s} \leq C K \left| f_j \right|_{C^{s+3/2+\sigma}} B^{N-1},$$

for some $C > 0$. Given these two results we can show, e.g., that

$$\left\| \partial_y u_N(x,0) \right\|_{H^s} \leq K_e \left\| R_{j,N} \right\|_{H^s} \leq K_e C K \left| f_j \right|_{C^{s+3/2+\sigma}} B^{N-1} \leq K B^N,$$

where K_e is the constant of the elliptic estimate, provided that B is chosen sufficiently large (greater than a multiple of the maximum of the $\left| f_j \right|_{C^{s+3/2+\sigma}}$). □

4. Weak formulation

In order to find approximate solutions of (7) numerically, we construct a weak formulation of the problem by decomposing the solution into two parts

$$u^p_{j,n} = \tilde{u}^p_{j,n} + \check{u}^p_{j,n}, \quad 1 \leq j \leq m. \tag{11}$$

We choose the first term $\tilde{u}^p_{j,n}$ to solve (7) with $R^p_{j,n} \equiv 0$, and the second term $\check{u}^p_{j,n}$ to solve (7) with homogeneous boundary data, $B^p_{s,n} \equiv \zeta^p_n \equiv J^p_{j,n} \equiv 0$. In other words, dropping the indices $\{n, p\}$ for the sake of simplicity, we rewrite the homogeneous equations for $\tilde{u}^p_{j,n}$

$$\partial_y^2 \tilde{u}_j(y) + (\beta_p^j)^2 \tilde{u}_j(y) = 0 \qquad \bar{g}_j < y < \bar{g}_{j-1}, \quad 1 \leq j \leq m, \tag{12a}$$
$$\tilde{u}_1(\bar{g}_1) - \tilde{u}_2(\bar{g}_1) = \zeta, \tag{12b}$$
$$\tilde{u}_j(\bar{g}_j) - \tilde{u}_{j+1}(\bar{g}_j) = 0, \qquad 2 \leq j \leq m-1, \tag{12c}$$
$$\partial_{y_1} \tilde{u}_1(\bar{g}_1) - \tau_1^2 \partial_{y_2} \tilde{u}_2(\bar{g}_1) = J_1, \tag{12d}$$
$$\partial_{y_j} \tilde{u}_j(\bar{g}_j) - \tau_j^2 \partial_{y_{j+1}} \tilde{u}_{j+1}(\bar{g}_j) = J_j, \qquad 2 \leq j \leq m-1, \tag{12e}$$
$$\partial_{y_1} \tilde{u}_1 - i\beta^1 \tilde{u}_1 = B_1, \qquad y_1 = a, \tag{12f}$$
$$\partial_{y_m} \tilde{u}_m + i\beta^m \tilde{u}_m = B_m, \qquad y_m = b. \tag{12g}$$

A general solution of these equations, (12), is easily found

$$\tilde{u}_j = M_j e^{i\beta^j y} + N_j e^{-i\beta^j y}, \quad 1 \leq j \leq m.$$

Using the boundary conditions in (12), the coefficients of M_j and N_j can be explicitly computed. It remains to investigate the equations for $\check{u}_j$ which can be solved by a High-Order Spectral method [37]

$$\partial_y^2 \check{u}_j(y) + (\beta_p^j)^2 \check{u}_j(y) = R_j(y) \qquad \bar{g}_j < y < \bar{g}_{j-1}, \quad 1 \leq j \leq m, \tag{13a}$$
$$\check{u}_1(\bar{g}_1) - \check{u}_2(\bar{g}_1) = 0, \tag{13b}$$
$$\check{u}_j(\bar{g}_j) - \check{u}_{j+1}(\bar{g}_j) = 0, \qquad 2 \leq j \leq m-1, \tag{13c}$$
$$\partial_{y_1} \check{u}_1(\bar{g}_1) - \tau_1^2 \partial_{y_2} \check{u}_2(\bar{g}_1) = 0, \tag{13d}$$
$$\partial_{y_j} \check{u}_j(\bar{g}_j) - \tau_j^2 \partial_{y_{j+1}} \check{u}_{j+1}(\bar{g}_j) = 0, \qquad 2 \leq j \leq m-1, \tag{13e}$$
$$\partial_{y_1} \check{u}_1 - i\beta^1 \check{u}_1 = 0, \qquad y_1 = a, \tag{13f}$$
$$\partial_{y_m} \check{u}_m + i\beta^m \check{u}_m = 0, \qquad y_m = b. \tag{13g}$$

The weak formulation of (13) is: Find $U \in H^1([b,a])$ such that

$$\begin{aligned}
&(\tilde{k}^2 U, \varphi) - (\partial_y U, \partial_y \varphi) + \sum_{j=1}^{m-1} (1 - \tau_j^2) \partial_y \check{u}_{j+1}(\bar{g}_j) \bar{\varphi}(\bar{g}_j) \\
&\quad = (\mathcal{R}, \varphi) - i\beta^1 \check{u}_1(a) \bar{\varphi}(a) - i\beta^m \check{u}_m(b) \bar{\varphi}(b), \quad \forall \varphi \in H^1([b,a]),
\end{aligned}$$

where

$$\begin{aligned}
&U := \check{u}_j, && y \in I_j, \\
&\varphi := \varphi_j, && y \in I_j, \\
&\mathcal{R} := R_j, && y \in I_j, \\
&\tilde{k}^2 := (\beta_p^j)^2, && y \in I_j, \\
&I_j := \{\bar{g}_{j-1} \leq y \leq \bar{g}_j\},
\end{aligned}$$

for $1 \le j \le m$. Here the inner product is defined on the interval (a, b) by

$$(u, v) := \int_a^b u\bar{v}\, dx,$$

where $\bar{v}$ is the complex conjugate of v.

To construct a Legendre–Galerkin method as in [33,34], we define the finite dimensional function space $X_{N_y} \subset H^1([b, a])$ by

$$X_{N_y} := \{\varphi_i \in P_{N_y}(I_j) \mid \partial_y \varphi_1(a) - i\beta^1 \varphi_1(a) = 0, \partial_y \varphi_m(b) + i\beta^m \varphi_m(b) = 0\},$$

where P_{N_y} is the space of polynomials of degree less than N_y. Then, the Legendre–Galerkin formulation is to find $U_{N_y} \in X_{N_y}$ such that

$$\begin{aligned}
&(\tilde{k}^2 U_{N_y}, \varphi_{N_y}) - (\partial_y U_{N_y}, \partial_y \varphi_{N_y}) + \sum_{j=1}^{m-1} (1 - \tau_j^2) \left[\partial_y (\check{u}_{j+1})_{N_y} \bar{\varphi}_{N_y}\right]_{y=\bar{g}_j} \\
&\quad = (\mathcal{I}_{N_y} \mathcal{R}_{N_y}, \varphi_{N_y}) - \left[i\beta^1 (\check{u}_1)_{N_y} \bar{\varphi}_{N_y}\right]_{y=a} \\
&\qquad - \left[i\beta^m (\check{u}_m)_{N_y} \bar{\varphi}_{N_y}\right]_{y=b}, \quad \forall \varphi_{N_y} \in X_{N_y},
\end{aligned}$$

where $\mathcal{I}_{N_y}$ is the projection operator onto P_{N_y}. By integration by parts on each subdomain I_j, the equivalent variational formulation is derived: Find $U_{N_y} \in X_{N_y}$ such that

$$\begin{aligned}
&(\tilde{k}^2 U_{N_y}, \varphi_{N_y}) + (\partial_y^2 U_{N_y}, \varphi_{N_y}) - \sum_{j=1}^{m-1} \tau_j^2 \left[\partial_y (\check{u}_{j+1})_{N_y} \bar{\varphi}_{N_y}\right]_{y=\bar{g}_j} \\
&\quad + \sum_{j=1}^{m-1} \left[\partial_y (\check{u}_j)_{N_y} \varphi_{N_y}\right]_{y=\bar{g}_j} = (\mathcal{I}_{N_y} \mathcal{R}_{N_y}, \varphi_{N_y}), \quad \forall \varphi_{N_y} \in X_{N_y}.
\end{aligned}$$

5. A Legendre–Galerkin numerical method in enriched spaces

Following the spectral Legendre–Galerkin approach [36,37], we consider basis functions as combinations of Legendre polynomials $L_j(y)$. For $y \in I_l$ and $1 \le l \le m$, we define

$$\phi_{l,j} := (1+i) L_j\left(\frac{2(y - \overline{g}_{l-1})}{\overline{g}_{l-1} - \overline{g}_l} + 1\right) + a_{l,j} L_{j+1}\left(\frac{2(y - \overline{g}_{l-1})}{\overline{g}_{l-1} - \overline{g}_l} + 1\right) + b_{l,j} L_{j+2}\left(\frac{2(y - \overline{g}_{l-1})}{\overline{g}_{l-1} - \overline{g}_l} + 1\right),$$

such that

$$\begin{aligned}
&\partial_y \phi_{1,j}(\overline{g}_0) - i\beta^1 \phi_{1,j}(\overline{g}_0) = 0, && \phi_{1,j}(\overline{g}_1) = 0, \\
&\phi_{s,j}(\overline{g}_{s-1}) = 0, && \phi_{s,j}(\overline{g}_s) = 0, \quad 2 \le s \le m-1, \\
&\partial_y \phi_{m,j}(\overline{g}_m) + i\beta^m \phi_{m,j}(\overline{g}_m) = 0, && \phi_{m,j}(\overline{g}_{m-1}) = 0,
\end{aligned}$$

where $j = 1, \ldots, N_y - 2$. Note that these Legendre–Galerkin basis functions vanish at the transition layers at $y = \overline{g}_s$ where $1 \le s \le m-1$. For this reason, we introduce additional basis functions which have the nonzero value of $(1+i)$ on $y = \overline{g}_s$ (see Fig. 3):

$$\eta_1(y) = \begin{cases} c_1(y - \overline{g}_1) + (1+i), & \overline{g}_1 \le y \le \overline{g}_0, \\ d_1(y - \overline{g}_1) + (1+i), & \overline{g}_2 \le y \le \overline{g}_1, \\ 0, & \text{otherwise}, \end{cases}$$

and

$$\eta_s(y) = \begin{cases} c_s(y - \overline{g}_s) + (1+i), & \overline{g}_s \le y \le \overline{g}_{s-1}, \\ d_s(y - \overline{g}_s) + (1+i), & \overline{g}_{s+1} \le y \le \overline{g}_s, \\ 0, & \text{otherwise}, \end{cases}$$

$2 \le s \le m-2$, and

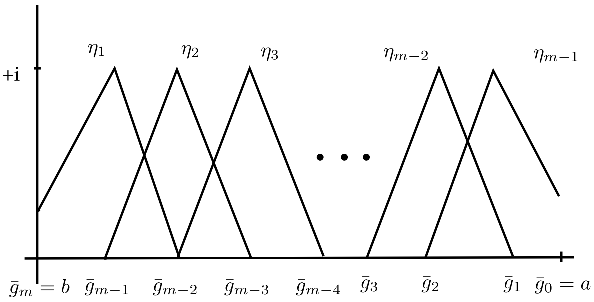

Fig. 3. A depiction of the additional basis functions η_s.

$$\eta_{m-1}(y) = \begin{cases} c_{m-1}(y-\overline{h}) + (1+i), & \overline{g}_{m-1} \le y \le \overline{g}_{m-2}, \\ d_{m-1}(y-\overline{h}) + (1+i), & \overline{g}_m \le y \le \overline{g}_{m-1}, \\ 0, & \text{otherwise.} \end{cases}$$

Enforcing the conditions

$$\begin{aligned} &(\partial_y \eta_1 - i\beta^1 \eta_1)(\overline{g}_0) = 0, \\ &\eta_1(\overline{g}_2) = 0, \\ &\eta_s(\overline{g}_{s-1}) = 0, && 2 \le s \le m-2, \\ &\eta_s(\overline{g}_{s+1}) = 0, && 2 \le s \le m-2, \\ &(\partial_y \eta_{m-1} + i\beta^m \eta_{m-1})(\overline{g}_m) = 0, \\ &\eta_{m-1}(\overline{g}_{m-2}) = 0, \end{aligned}$$

we find

$$\begin{aligned} &c_1 = \frac{i\beta^u}{(1+i) - i\beta^1(\overline{g}_0 - \overline{g}_1)}, \\ &d_1 = \frac{1+i}{\overline{g}_1 - \overline{g}_2}, \\ &c_s = \frac{1+i}{\overline{g}_s - \overline{g}_{s+1}}, && 2 \le s \le m-2, \\ &d_s = \frac{1+i}{\overline{g}_s - \overline{g}_{s-1}}, && 2 \le s \le m-2, \\ &c_{m-1} = \frac{1+i}{\overline{g}_{m-1} - \overline{g}_{m-2}}, \\ &d_{m-1} = -\frac{i\beta^m}{(1+i) + i\beta^m(\overline{g}_m - \overline{g}_{m-1})}. \end{aligned}$$

With these we construct the basis functions defined on all $\{\bar{g}_m < y < \bar{g}_0\}$

$$\tilde{\phi}_j(y) = \begin{cases} \phi_{1,j}(y), & \bar{g}_1 < y < \bar{g}_0, \\ 0, & \bar{g}_m < y < \bar{g}_1, \end{cases} \quad j = 0, \dots, N_y - 2,$$

and

$$\tilde{\phi}_{s(N_y-1)+j}(y) = \begin{cases} 0, & \bar{g}_{s-1} < y < \bar{g}_0, \\ \phi_{s,j}(y), & \bar{g}_s < y < \bar{g}_{s-1}, \\ 0, & \bar{g}_m < y < \bar{g}_s, \end{cases} \quad j = 0, \dots, N_y - 2,$$

for $2 \le s \le m-2$, and

$$\tilde{\phi}_{(m-1)(N_y-1)+j}(y) = \begin{cases} 0, & \bar{g}_{m-1} < y < \bar{g}_{m-2}, \\ \phi_{3,j}(y), & \bar{g}_m < y < \bar{g}_{m-1}, \end{cases} \quad j = 0, \dots, N_y - 2,$$

and finally,

$$\tilde{\phi}_{m(N_y-1)+(q-1)}(y) = \eta_q(y), \quad \bar{g}_{q+1} < y < \bar{g}_{q-1}, \quad 1 \le q \le m-1.$$

Setting $\bar{N} = m(N_y-1) - 1 + (m-1)$, our numerical approximation reads

$$u_{N_y}(y) := \sum_{j=0}^{\bar{N}} \hat{u}_j \tilde{\phi}_j(y),$$

and we seek

$$\mathbf{u}_s = (\hat{u}_{(s-1)(N_y-1)}, \hat{u}_{(s-1)(N_y-1)+1}, \ldots, \hat{u}_{(s-1)(N_y-1)+N_y-2})^T, \quad 1 \le s \le m,$$

where we are given

$$\begin{aligned}
&\mathbf{f} = (\hat{f}_0, \ldots, \hat{f}_{m(N_y-1)-1})^T, \\
&\hat{f}_j := (\mathcal{I}_N f, \tilde{\phi}_j), \qquad\qquad j = 0, \ldots, \bar{N}.
\end{aligned}$$

We also define the matrices

$$(A_s)_{lj} = (\partial_y^2 \tilde{\phi}_{(s-1)(N_y-1)+j}, \tilde{\phi}_{(s-1)(N_y-1)+l})_{I_s} + \tilde{k}_s^2 (\tilde{\phi}_{(s-1)(N_y-1)+j}, \tilde{\phi}_{(s-1)(N_y-1)+l})_{I_s},$$

where $0 \le l, j \le N_y - 2$ and $1 \le s \le m$. We define the column vectors

$$\begin{aligned}
a_{12} &= (\partial_y^2 \tilde{\phi}_{m(N_y-1)} + k_1^2 \tilde{\phi}_{m(N_y-1)}, \tilde{\phi}_j)_{I_1}, \\
c_{12}^q &= (\partial_y^2 \tilde{\phi}_{m(N_y-1)+(q-1)} + k_{q+1}^2 \tilde{\phi}_{m(N_y-1)+(q-1)}, \tilde{\phi}_{q(N_y-1)+j})_{I_{q+1}}, \\
d_{12}^q &= (\partial_y^2 \tilde{\phi}_{m(N_y-1)+q} + k_{q+1}^2 \tilde{\phi}_{m(N_y-1)+q}, \tilde{\phi}_{q(N_y-1)+j})_{I_{q+1}}, \\
b_{12} &= (\partial_y^2 \tilde{\phi}_{\bar{N}} + k_m^2 \tilde{\phi}_{\bar{N}}, \tilde{\phi}_{(m-1)(N_y-1)+j})_{I_m},
\end{aligned}$$

and row vectors

$$\begin{aligned}
a_{21} &= (\partial_y^2 \tilde{\phi}_j + k_1^2 \tilde{\phi}_j, \tilde{\phi}_{m(N_y-1)})_{I_1} + \partial_y \tilde{\phi}_j(\bar{g}_1), \\
c_{21}^q &= (\partial_y^2 \tilde{\phi}_{q(N_y-1)+j} + k_{q+1}^2 \tilde{\phi}_{q(N_y-1)+j}, \tilde{\phi}_{m(N_y-1)+(q-1)})_{I_{q+1}} \\
&\quad - \tau_q^2 \partial_y \tilde{\phi}_{q(N_y-1)+j}(\bar{g}_q), \\
d_{21}^q &= (\partial_y^2 \tilde{\phi}_{q(N_y-1)+j} + k_{q+1}^2 \tilde{\phi}_{q(N_y-1)+j}, \tilde{\phi}_{m(N_y-1)+q})_{I_{q+1}} + \partial_y \tilde{\phi}_{q(N_y-1)+j}(\bar{g}_{q+1}), \\
b_{21} &= (\partial_y^2 \tilde{\phi}_{(m-1)(N_y-1)+j} + k_m^2 \tilde{\phi}_{(m-1)(N_y-1)+j}, \tilde{\phi}_{\bar{N}})_{I_m} \\
&\quad - \tau_{m-1}^2 \partial_y \tilde{\phi}_{(m-1)(N_y-1)+j}(\bar{g}_{m-1}),
\end{aligned}$$

for $0 \le j \le N_y - 2$ and $1 \le q \le m-2$. Moreover, we set

$$\begin{aligned}
a_{22}^q &= (\partial_y^2 \tilde{\phi}_{m(N_y-1)+(q-1)} + k^2 \tilde{\phi}_{m(N_y-1)+(q-1)}, \tilde{\phi}_{m(N_y-1)+(q-1)}) \\
&\quad + \partial_y \tilde{\phi}_{m(N_y-1)+(q-1)}(\bar{g}_q^+) - \tau_q^2 \partial_y \tilde{\phi}_{m(N_y-1)+(q-1)}(\bar{g}_q^-), \\
a_{23}^s &= (\partial_y^2 \tilde{\phi}_{m(N_y-1)+s} + k^2 \tilde{\phi}_{m(N_y-1)+s}, \tilde{\phi}_{m(N_y-1)+(s-1)}) \\
&\quad + [\partial_y \tilde{\phi}_{m(N_y-1)+s}(\bar{g}_s^+) - \tau_s^2 \partial_y \tilde{\phi}_{m(N_y-1)+s}(\bar{g}_s^-)] \tilde{\phi}_{m(N_y-1)+(s-1)}(\bar{g}_s), \\
a_{32}^s &= (\partial_y^2 \tilde{\phi}_{m(N_y-1)+(s-1)} + k^2 \tilde{\phi}_{m(N_y-1)+(s-1)}, \tilde{\phi}_{m(N_y-1)+s}) \\
&\quad + [\partial_y \tilde{\phi}_{m(N_y-1)+(s-1)}(\bar{g}_{s+1}^+) \\
&\quad - \tau_{s+1}^2 \partial_y \tilde{\phi}_{m(N_y-1)+(s-1)}(\bar{g}_{s+1}^-)] \tilde{\phi}_{m(N_y-1)+s}(\bar{g}_{s+1}),
\end{aligned}$$

where $1 \le q \le m-1$ and $1 \le s \le m-2$. Here, $\partial_y \tilde{\phi}_n(\bar{g}_l^-)$ and $\partial_y \tilde{\phi}_n(\bar{g}_l^+)$ stand for the left and right derivatives at $\bar{g}_l$, respectively. The Legendre–Galerkin scheme demands the $m(N_y-1) + (m-1) = mN_y - 1$ equations:

$$\mathbf{M}\mathbf{u} = \mathbf{f},$$

where $\mathbf{M}$ is a block matrix

$$\mathbf{M} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

The block matrix A is defined as

$$A = \begin{pmatrix} A_1 & & & & 0 \\ & A_2 & & & \\ & & \ddots & & \\ & & & A_{m-1} & \\ 0 & & & & A_m \end{pmatrix},$$

and the bidiagonal matrices B and C are defined as

$$B = \begin{pmatrix} a_{12} & 0 & \dots & & 0 \\ c_{12}^1 & d_{12}^1 & & & \\ 0 & c_{12}^1 & d_{12}^1 & & \\ & & \ddots & & \\ & & & c_{12}^{m-2} & d_{12}^{m-2} \\ 0 & \dots & & 0 & b_{12} \end{pmatrix},$$

and

$$C = \begin{pmatrix} (a_{21})^T & (c_{21}^1)^T & 0 & \dots & & 0 \\ 0 & (d_{12}^1)^T & (c_{21}^2)^T & & & \\ & & (d_{12}^2)^T & & & \\ & & & \ddots & (c_{21}^{m-2})^T & \\ 0 & \dots & & & (d_{21}^{m-2})^T & (b_{21})^T \end{pmatrix}.$$

Finally the tridiagonal matrix D is given by

$$D = \begin{pmatrix} a_{22}^1 & a_{23}^1 & & & 0 \\ a_{32}^1 & a_{22}^2 & a_{23}^2 & & \\ 0 & a_{32}^2 & a_{22}^3 & a_{23}^3 & \\ & & & \ddots & \\ 0 & & a_{32}^{m-3} & a_{32}^{m-2} & a_{22}^{m-2} \\ 0 & & & a_{32}^{m-2} & a_{22}^{m-1} \end{pmatrix}.$$

6. Numerical simulations

We now conduct a variety of numerical experiments with an implementation of the algorithm described above. We begin by demonstrating the accuracy of the Legendre–Galerkin solver for the boundary value problem (13) at the heart of our numerical method versus an exact solution. We follow this by presenting the performance of our full scattering solver using “energy defect” as an indicator of convergence [38].

6.1. Simulations of a boundary value problem

To begin our numerical simulations, we investigate the numerical approximation of the reduced problem (13) which is at the core of our full solver. Utilizing the algorithm proposed in Section 5 we look for numerical convergence of the following reduced problem

$$\partial_y^2 u_j(y) + k_j^2 u(y) = F_j(y), \qquad \bar{g}_j < y < \bar{g}_{j-1}, \quad 1 \le j \le m, \tag{14a}$$

$$u_j(\bar{g}_j) - u_{j+1}(\bar{g}_j) = 0, \qquad 1 \le j \le m-1, \tag{14b}$$

$$\partial_y u_j(\bar{g}_j) - \tau_j^2 \partial_y u_{j+1}(\bar{g}_j) = 0, \qquad 1 \le j \le m-1, \tag{14c}$$

$$\partial_y u_1(a) - i\beta^1 u_1(a) = 0, \tag{14d}$$

$$\partial_y u_m(b) + i\beta^m u_m(b) = 0. \tag{14e}$$

As a test of convergence we consider the following functions and parameters which collectively define an exact solution in TE polarization ($\tau_j = 1$)

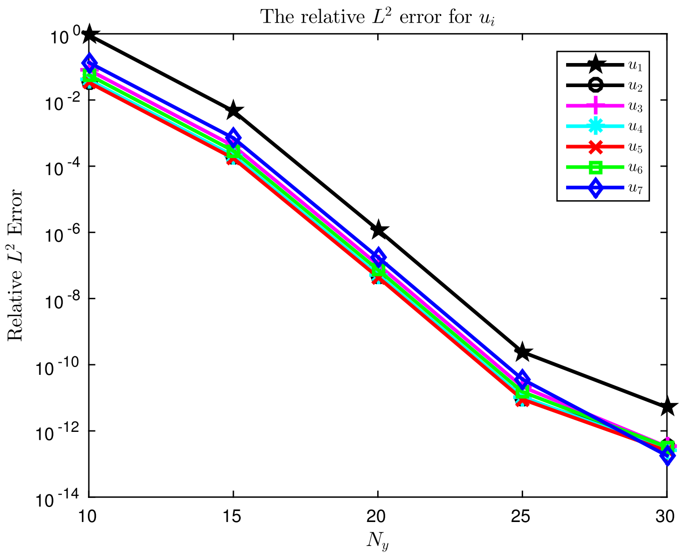

Fig. 4. Relative L^2 error, (17), of our Legendre–Galerkin approximation of (14) in configuration (15) versus number of basis functions N_y.

$$
\begin{aligned}
&u_j = \sin(\pi y)(a-y)^2(b-y)^2, \quad \bar{g}_j < y < \bar{g}_{j-1},\\
&a = 5, \quad b = -5, \quad \bar{g}_j = a - j, \quad \tau_j = 1, \quad 1 \le j \le 6,\\
&k_l = 1.5 - 0.1l, \quad \beta_1 = 1 + i, \quad \beta_7 = 1 - i, \quad 1 \le l \le 7.
\end{aligned} \tag{15}
$$

As a second example, we consider a TM simulation ($\tau_j \neq 1$)

$$
\begin{aligned}
&u_j = \prod_{q=1}^{7-j} \tau_q (a-y)^2(b-y)^2 \prod_{s=1}^{6}(y-\bar{g}_s), \quad \bar{g}_j < y < \bar{g}_{j-1},\\
&a = 5, \quad b = -10, \quad \bar{g}_j = a - 2j, \quad \tau_j = 1 + 2j, \quad 1 \le j \le 6,\\
&k_l = 12 - l, \quad \beta_1 = 1 + i, \quad \beta_7 = 2 - 3i, \quad 1 \le l \le 7.
\end{aligned} \tag{16}
$$

In configurations (15) and (16) the functions u_j satisfy the boundary conditions in (14), and one can easily find the corresponding terms F_j which render them exact solutions by direct calculation. To test numerical convergence, we define the relative L^2 error

$$
\frac{\|u_{\text{ex}} - u_{N_y}\|_{L^2}}{\|u_{\text{ex}}\|_{L^2}}, \tag{17}
$$

where u_{ex} is the exact solution and u_{N_y} is the numerical approximation. Figs. 4 and 5 display the spectral rate of convergence which our algorithm typically delivers in this simplified setting.

6.2. Simulations of a multiply layered medium: smooth interfaces

We now perform numerical experiments of a multiply layered medium whose scattering returns are governed by (2). Unlike the simplified problems in Section 6.1, exact solutions are unavailable so we utilize the widely accepted error measurement of energy defect [38,27,47–49]. In more detail, consider the Rayleigh expansions in the upper and lower layers

$$
u_1(x,y) = \sum_{p=-\infty}^{\infty} \hat{u}_1^p e^{i\beta_p^1 y} e^{i\alpha_p x}, \quad u_m(x,y) = \sum_{p=-\infty}^{\infty} \hat{u}_m^p e^{-i\beta_p^m y} e^{i\alpha_p x}.
$$

With these the efficiencies are defined as

$$
\begin{aligned}
e_p^1 &:= \frac{\beta_p^1}{\beta} \left|\hat{u}_1^p\right|^2, \quad p \in \mathcal{U}^1,\\
e_p^m &:= \frac{\beta_p^m}{\beta} \left|\hat{u}_m^p\right|^2, \quad p \in \mathcal{U}^m.
\end{aligned}
$$

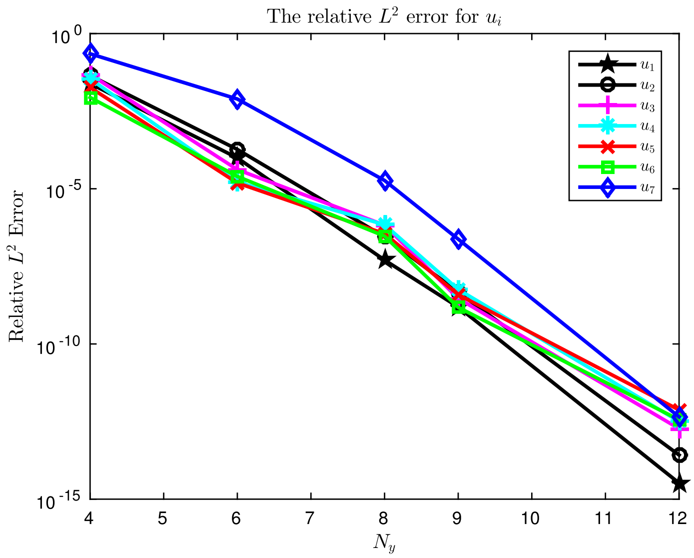

Fig. 5. Relative L^2 error, (17), of our Legendre–Galerkin approximation of (14) in configuration (16) versus number of basis functions N_y.

The efficiencies measure the energy at wavenumber p propagated away from the grating interface. If all materials in the structure are lossless ($k_l \in \mathbf{R}$) energy is conserved so

$$\text{(TE mode):} \qquad \sum_{p\in\mathcal{U}^1} e_p^1 + \sum_{p\in\mathcal{U}^m} e_p^m = 1,$$

$$\text{(TM mode):} \qquad \sum_{p\in\mathcal{U}^1} e_p^1 + \frac{k_1^2}{k_m^2}\sum_{p\in\mathcal{U}^m} e_p^m = 1.$$

Hence, we define the "energy defect" for TE and TM modes

$$\text{(TE mode):} \qquad \delta_{TE} = 1 - \sum_{p\in\mathcal{U}^1} e_p^1 - \sum_{p\in\mathcal{U}^m} e_p^m,$$

$$\text{(TM mode):} \qquad \delta_{TM} = 1 - \sum_{p\in\mathcal{U}^1} e_p^1 - \frac{k_1^2}{k_m^2}\sum_{p\in\mathcal{U}^m} e_p^m,$$

which will be zero for an exact solution.

We pursue a sequence of simulations to show the spectral convergence of the proposed Legendre–Galerkin method. For a set of smooth interface tests, we use the following data:

$$a = 5, \quad b = -8, \quad \alpha = 0.1, \quad d = 2\pi. \tag{18}$$

In our numerical experiments we use the following parameters to display the performance of our methods

$$N = \text{Perturbation order}, \quad N_x = \text{Number of basis functions in } x,$$
$$N_y = \text{Number of basis functions in } y, \quad \varepsilon = \text{Perturbation parameter},$$
$$k_j = \text{wave number in layer } j.$$

In Fig. 6 we display, in both TE and TM polarizations, the energy defect versus number of perturbation orders retained for the parameter choices

$$N_x = N_y = 40, \quad \bar{g}_j = a - j, \quad 1 \le j \le 6,$$
$$g_j(x) = \begin{cases} \varepsilon \sin(x), & \text{for } j \text{ is odd}, \\ \varepsilon \sin(x)/2, & \text{for } j \text{ is even}, \end{cases}$$
$$k_l = l + 0.5, \quad 1 \le l \le 7. \tag{19}$$

The figure shows the spectral convergence of the energy defect as the perturbation order N is refined. We also find that the energy defect decays rapidly to machine precision for small choices of ε.

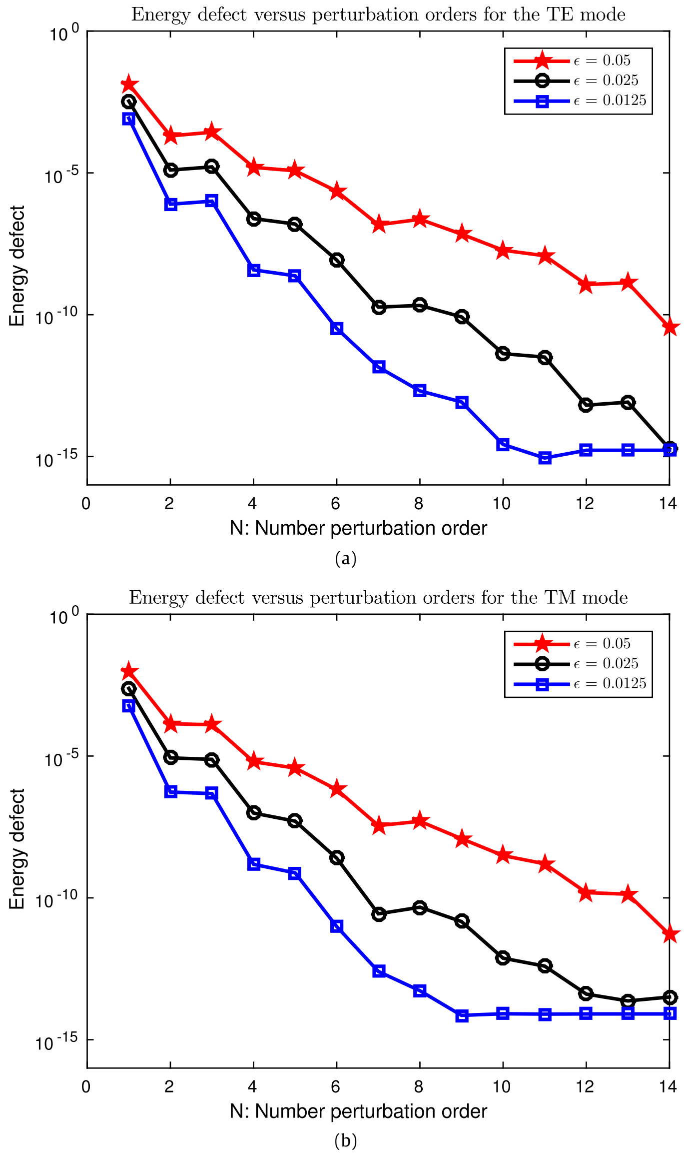

Fig. 6. (a) Energy defect versus perturbation order, N, for smooth interface configuration (18) and (19): TE mode with 7 layers. (b) Energy defect versus perturbation order, N, for smooth interface configuration (18) and (19): TM mode with 7 layers.

In Fig. 7 we repeat the same experiment (18) and (19) for fixed value of $\varepsilon = 0.025$ while varying the spatial discretization parameters N_x and N_y with 7 layers

$$\begin{aligned} &\bar{g}_j = a - j, \quad 1 \le j \le 6, \quad \varepsilon = 0.025, \\ &g_j(x) = \begin{cases} \varepsilon \sin(x), & \text{for } j \text{ is odd}, \\ \varepsilon \sin(x)/2, & \text{for } j \text{ is even}, \end{cases} \\ &k_l = l + 0.5, \quad 1 \le l \le 7. \end{aligned} \tag{20}$$

This clearly shows the spectral convergence of the energy defect versus the number of Fourier and Legendre coefficients.

We then focus on the behavior of our method with different numbers of layers. For this, we choose the following parameters:

$$\begin{aligned} &N_x = N_y = 40, \quad N = 15, \quad \varepsilon = 0.025, \\ &g_j(x) = \begin{cases} \varepsilon \sin(x), & \text{for } j \text{ is odd}, \\ \varepsilon \sin(x)/2, & \text{for } j \text{ is even}, \end{cases} \end{aligned}$$

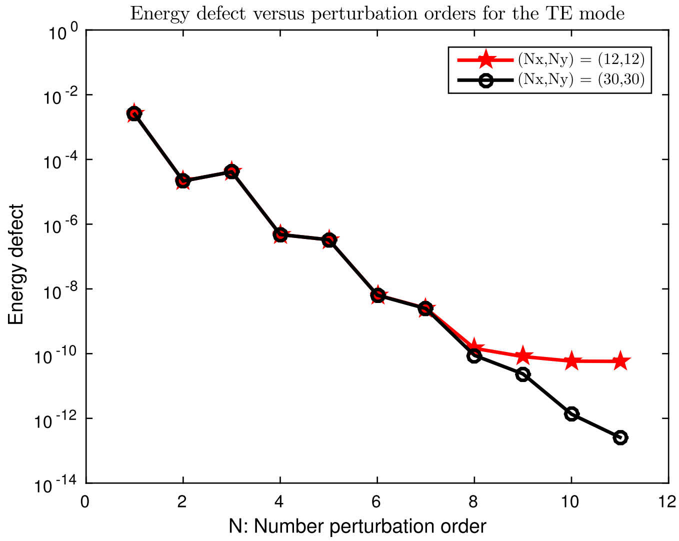

Fig. 7. Energy defect versus perturbation order, N, for smooth interface configuration (18) and (20): TE mode with 7 layers.

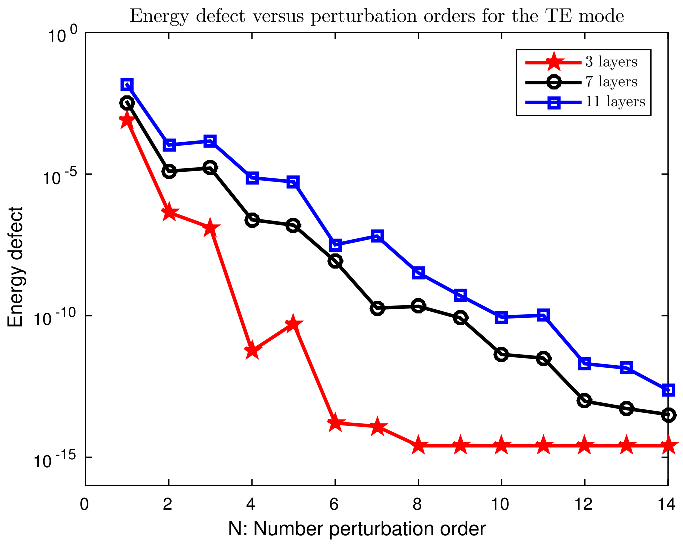

Fig. 8. Energy defect versus perturbation order, N, for smooth interface configuration (18) and (21): TE mode with different numbers of layers.

$$\bar{g}_j = a - j, \quad 1 \le j \le m-1, \quad k_l = l + 0.25, \quad 1 \le l \le m, \tag{21}$$

where m is the number of layers. Fig. 8 displays the energy defect versus number of perturbation orders. As anticipated, we observe rapid convergence to machine precision for a small number of layers, however, we still observe spectral convergence with (relatively) large m.

If the wavenumbers are not large, as in (19), then we observe rapid convergence to machine precision. However, as we increase to larger values of k, our convergence deteriorates due to the severe under-resolution of our parameter choices. To address this concern, in Fig. 9, we revisit these computations with the following parameters

$$\begin{aligned}
&N = 15, \quad \varepsilon = 0.025, \\
&g_j(x) = \begin{cases} \varepsilon \sin(x), & \text{for } j \text{ is odd}, \\ \varepsilon \sin(x)/2, & \text{for } j \text{ is even}, \end{cases} \\
&(k_1, k_2, k_3, k_4, k_5) = (1.5, 2.5, 40, 41, 1), \\
&\bar{g}_j = a - j, \quad 1 \le j \le 4.
\end{aligned} \tag{22}$$

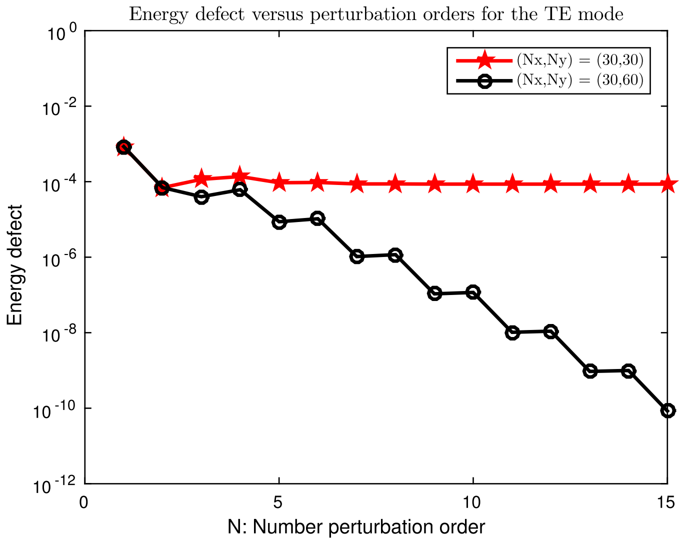

Fig. 9. Energy defect versus perturbation order, N, for smooth interface configuration (18) and (22) with high wavenumbers: TE mode with 5 layers.

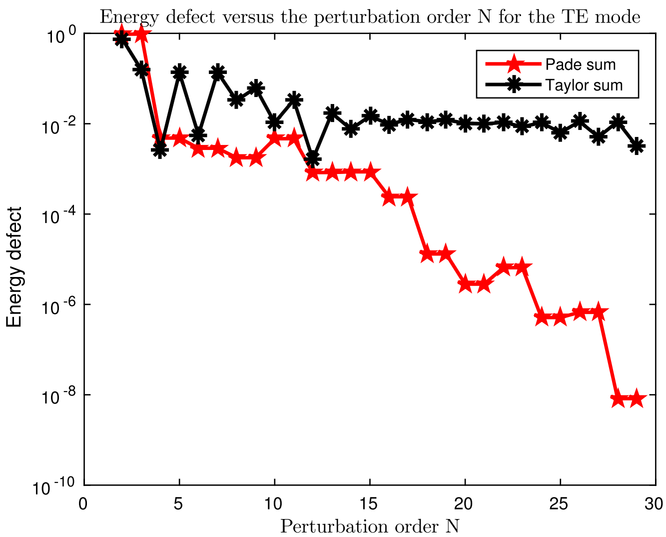

Fig. 10. Energy defect versus perturbation order, N, for smooth interface configuration (18) and (23) with large deformation: TE mode with 5 layers.

As we see, if we refine our spatial discretization then we recover the striking convergence we saw for the smallest values of k. To be more specific, as we display in Fig. 9, if we choose $N_x = 30$ and $N_y = 60$, we can realize results with excellent accuracy for large wave numbers.

To close, we examine the possibility of using our new algorithm for deformations of large size. To study this we revisit the calculations specified in (19), with the parameter choices

$$\begin{aligned}
&N = 30, \quad \varepsilon = 1.0, \\
&\bar{g}_0 = a = 6, \quad \bar{g}_1 + g_1 = 4 + \varepsilon \sin(x), \quad \bar{g}_2 + g_2 = 2 + \varepsilon \sin(x)/2, \\
&\bar{g}_3 + g_3 = 0 + \varepsilon \sin(x), \quad \bar{g}_4 = b = -2, \\
&(k_1, k_2, k_3, k_4) = (1.0, 1.25, 1.5, 1.75), \quad (N_x, N_y) = (60, 60).
\end{aligned} \tag{23}$$

In Fig. 10 we display numerical simulations with $\varepsilon = 1$ on these four layers. As in [50,51] simple Taylor summation in perturbation order N does not work well. However, if Padé approximation [35] is utilized then excellent results can be realized showing that, indeed, *large* deformations can be readily simulated.

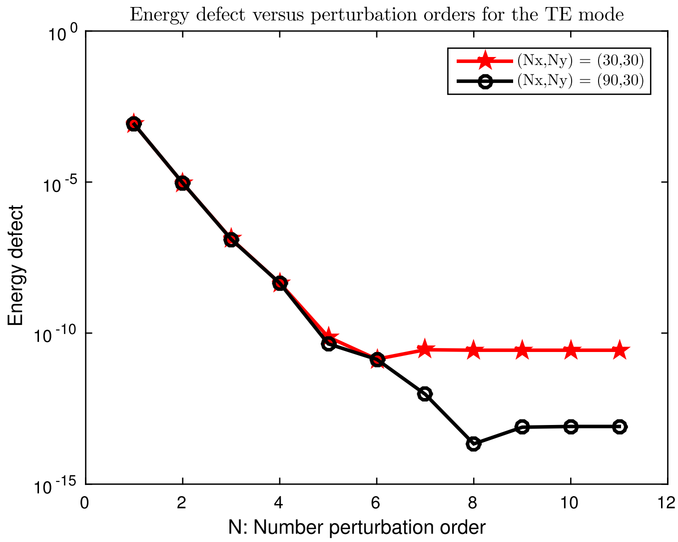

Fig. 11. Energy defect versus perturbation order, N, for rough interface configuration (18) and (26): TE mode with 6 layers.

6.3. Simulations of a multiply layered medium: rough interfaces

To continue our investigation we conduct a numerical simulation with rough interfaces

$$f_r(x) = \left(2 \times 10^{-4}\right) x^4 (2\pi - x)^4 - c_0,$$
$$f_L(x) = \begin{cases} -\frac{2}{\pi} x + 1, & 0 \le x \le \pi, \\ \frac{2}{\pi} x - 3, & \pi \le x \le 2\pi, \end{cases}$$

where c_0 is chosen so that f_r has zero mean, which possess C^4 (but not C^5) and Lipschitz regularity, respectively [28,51]. For our numerical experiments we use the Fourier series representations

$$f_r(x) = \sum_{k=1}^{\infty} \frac{96(2k^2\pi^2 - 21)}{125k^8} \cos(kx), \tag{24a}$$

$$f_L(x) = \sum_{k=1}^{\infty} \frac{8}{\pi^2(2k-1)^2} \cos((2k-1)x), \tag{24b}$$

which we truncate after wavenumber $P = 40$,

$$f_{r,P}(x) = \sum_{k=1}^{P} \frac{96(2k^2\pi^2 - 21)}{125k^8} \cos(kx), \tag{25a}$$

$$f_{L,P}(x) = \sum_{k=1}^{P/2} \frac{8}{\pi^2(2k-1)^2} \cos((2k-1)x). \tag{25b}$$

For these simulations we choose the following parameters with 6 layers:

$$\begin{aligned} & a = 4, \quad b = -2, \quad \bar{g}_j = a - j, \quad 1 \le j \le 5, \\ & k_l = l + 0.5, \quad 1 \le l \le 6, \\ & g_1(x) = g_5(x) = \varepsilon \sin x, \quad g_2(x) = g_4(x) = \varepsilon f_{r,40}, \quad g_3(x) = \varepsilon f_{L,40}, \\ & \alpha = 0.5, \quad d = 2\pi, \quad \varepsilon = 0.025. \end{aligned} \tag{26}$$

In Fig. 11 we display results of this experiment with rough interfaces. Evidently, our new method is applicable to configurations with finite and even Lipschitz smoothness, provided with sufficient resolution. In Fig. 12 we present the real and imaginary parts of the scattered solution for the rough interface configuration (26). The layer interfaces are depicted in dashed red lines.

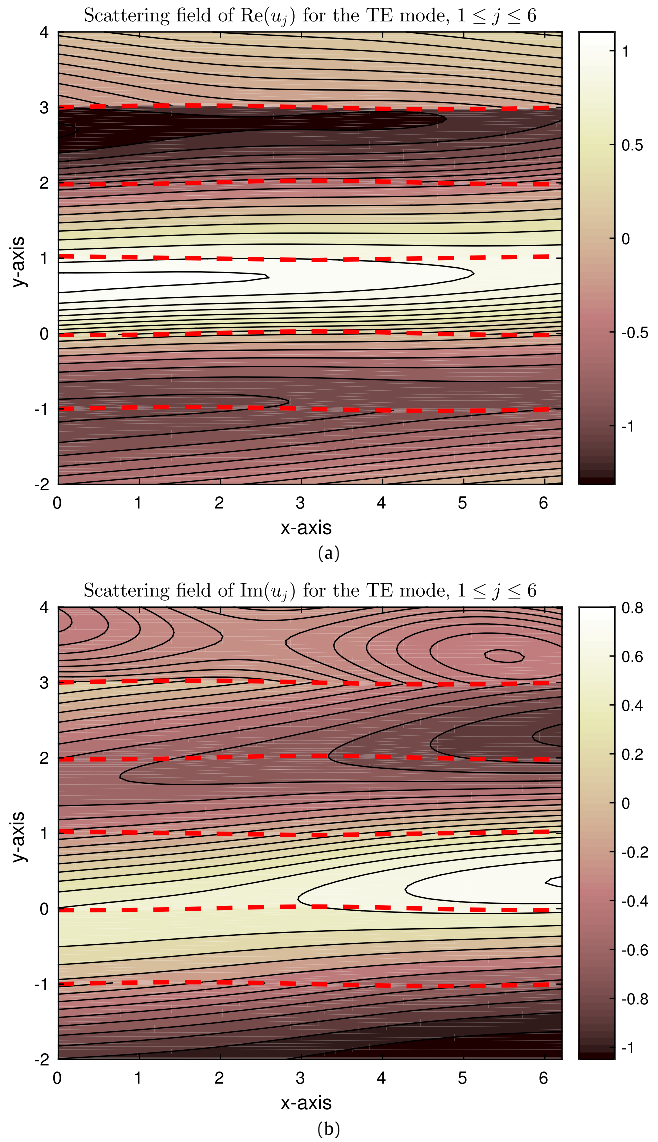

Fig. 12. (a) Plot of the real part of the scattered field for rough interface (26): TE mode with 6 layers. Layer interfaces depicted in dashed red lines. (b) Plot of the imaginary part of the scattered field for rough interface (26): TE mode with 6 layers. Layer interfaces depicted in dashed red lines. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

6.4. Simulations of a triply layered medium in three dimensions

To close our study, we consider three dimensional scattering of scalar waves in a triply layered periodic medium. This bridges to our forthcoming article which studies the vector wave scattering problems of electromagnetics (in three dimensions) and linear elastodynamics.

In Fig. 13 we display the energy defect versus number of perturbation orders retained for the following parameter choices with three layers

$$\begin{aligned}
&N_{x_1} = N_{x_2} = N_y = 20, \quad a = 6, \quad \bar{g}_1 = 2, \quad \bar{g}_2 = 0, \quad b = -4, \\
&k_1 = 1.5, \quad k_2 = 2.5, \quad k_3 = 3.5, \quad d_1 = d_2 = 2\pi, \quad \alpha_1 = 0.1, \quad \alpha_2 = 0.3, \\
&g_1(x) = \varepsilon \cos(x_1)\sin(x_2), \quad g_2(x) = \varepsilon \sin(x_1)\cos(x_2).
\end{aligned} \tag{27}$$

These experiments are conducted for TE and TM modes with various values of ε. The figure shows the spectral convergence in the energy defect as the perturbation order N is refined as in two-dimensional problem. We also find that the energy defect decays rapidly to machine precision for small choices of ε.

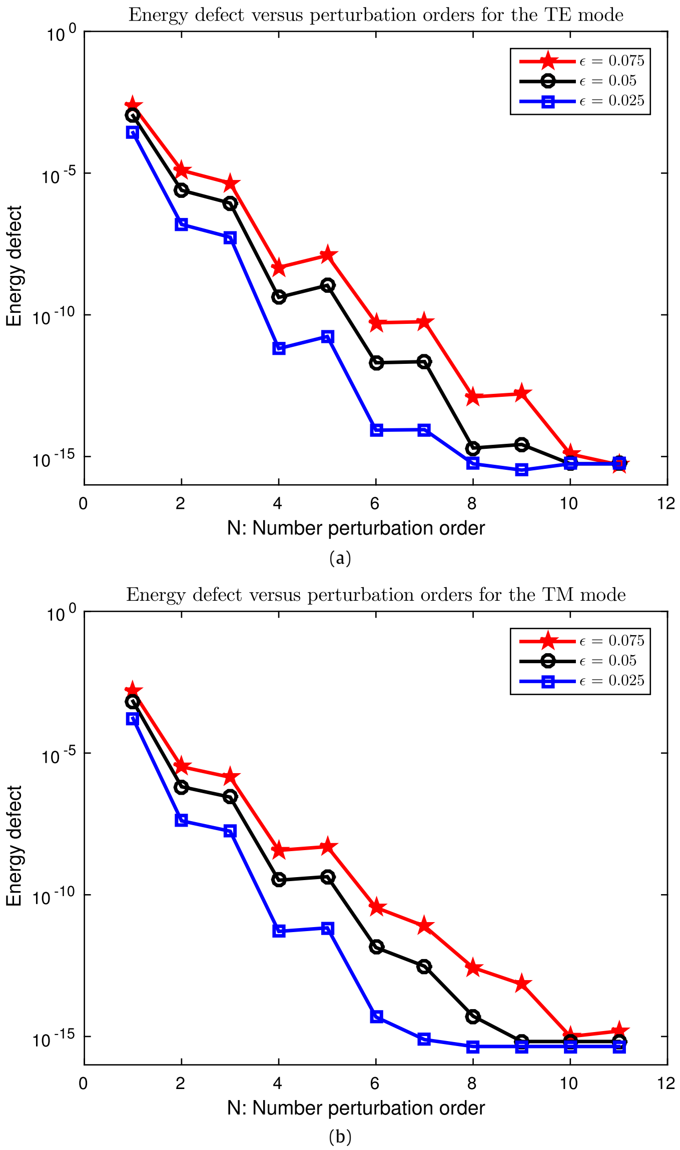

Fig. 13. (a) Energy defect versus perturbation order, N, for smooth interface configuration (27): TE mode with 3 layers in three dimensions. (b) Energy defect versus perturbation order, N, for smooth interface configuration (27): TM mode with 3 layers in three dimensions.

As a final experiment with our algorithm we consider a rough profile configuration in three dimensional media. Recalling the Fourier series representations (24) and (25), we replace g_1 and g_2 in (27) with

$$g_1(x) = \varepsilon f_{r,40}(x_1) f_{L,40}(x_2), \quad g_2(x) = \varepsilon f_{L,40}(x_1) f_{r,40}(x_2). \tag{28}$$

In Fig. 14 we display the results of the experiment (27) with the rough interfaces (28). We see that our new method is applicable to configurations with finite and even Lipschitz smoothness, with sufficient resolution, in three dimensional simulations.

7. Conclusions

We have described a High Order Perturbation of Surfaces (HOPS) algorithm for linear wave scattering by a periodic two or three dimensional medium with an arbitrary number of layers. Using Dirichlet–Neumann Operators at artificial boundaries placed above and below the structure, we equivalently reformulated the governing Helmholtz equations on a bounded domain. Using a suitable change of variables, the governing equations were reformulated on a separable geometry with flat interfaces. Introducing a boundary perturbation method, we described the scattered field in a Taylor series. More precisely, we derived a sequence of linear differential equations to be solved at each perturbation order in the expansion,

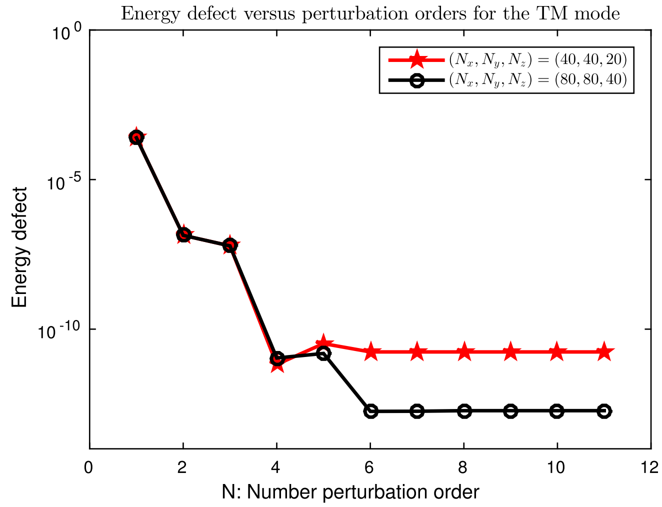

Fig. 14. Energy defect versus perturbation order, N, for rough interface configuration (27): TM mode with 3 layers in three dimensions. Here, the function g_1 and g_2 are replaced in (27) with (28).

resulting in a Transformed Field Expansions (TFE) algorithm. Accurate numerical experiments using these TFE recursions were conducted with a Legendre–Galerkin method based on a novel weak formulation. These experiments include smooth and rough interfaces of small and large size, and high wavenumbers in a multiply layered medium. We also presented three-dimensional simulations in a triply layered medium as a natural extension of our algorithm. The numerical simulations showed the spectral convergence which our new algorithm can achieve.

Our developments clearly point towards several extensions of great importance. In particular, our approach must be generalized to the vector wave scattering problems of electromagnetics (in three dimensions) and linear elastodynamics. These extensions will not be straightforward since more complicated boundary conditions between layers are considered, and hence the algorithmic differences are significant. We will describe them in a future publication. In addition, a rigorous numerical analysis for the proposed Legendre–Galerkin methods will be considered in a forthcoming article.

Acknowledgements

Y.H. gratefully acknowledges support from the Simons Foundation. D.P.N. gratefully acknowledges support from the National Science Foundation through grant No. DMS-1522548.

Appendix A. Derivation of the transformed equations

In this section we provide a full derivation of the transformed equations (3) presented in Section 3.1. For notational convenience we set $g_0 \equiv g_m \equiv 0$. By the chain rule, we find

$$\begin{aligned}
\nabla_x &= \nabla_{x'} + (\nabla_x y_j)\partial_{y_j}, \\
\partial_y &= \partial_{y_j}(\partial_y y_j),
\end{aligned}$$

for $\bar{g}_j + g_j < y < \bar{g}_{j-1} + g_{j-1}$, $1 \le j \le m$. Hence, recalling $a = \bar{g}_0 + g_0$ and $b = \bar{g}_m + g_m$, we deduce that

$$(x', y_j) = \left(x, \bar{g}_j \left(\frac{y - (\bar{g}_{j-1} + g_{j-1})}{(\bar{g}_j + g_j) - (\bar{g}_{j-1} + g_{j-1})} \right) + \bar{g}_{j-1} \left(\frac{(\bar{g}_j + g_j) - y}{(\bar{g}_j + g_j) - (\bar{g}_{j-1} + g_{j-1})} \right) \right),$$

for $\bar{g}_j + g_j < y < \bar{g}_{j-1} + g_{j-1}$, $1 \le j \le m$. Setting $M_j(x) := (\bar{g}_j + g_j) - (\bar{g}_{j-1} + g_{j-1})$ for $1 \le j \le m$, we write

$$\begin{aligned}
M_j \nabla_x &= M_j \nabla_{x'} + M_j (\nabla_x y_j)\partial_{y_j} \\
&= M_j \nabla_{x'} + [\nabla_x M_j y_j - (\nabla_x M_j) y_j]\partial_{y_j} \\
&= M_j \nabla_{x'} + [\nabla_x g_{j-1}(y_j - \bar{g}_j) - \nabla_x g_j (y_j - \bar{g}_{j-1})]\partial_{y_j}, \\
M_j \mathrm{div}_x &= M_j \mathrm{div}_{x'} + [\nabla_x g_{j-1}(y_j - \bar{g}_j) - \nabla_x g_j (y_j - \bar{g}_{j-1})] \cdot \partial_{y_j}, \\
M_j \partial_y &= M_j (\partial_{y_j})(\partial_y y_j) = \partial_y (M_j y_j)\partial_{y_j} \\
&= (\bar{g}_j - \bar{g}_{j-1})\partial_{y_j},
\end{aligned}$$

Hence, defining

$$C_j(x) = M_j(x), \quad E_j = \bar{g}_j - \bar{g}_{j-1},$$
$$D_j(x) = \nabla_x g_{j-1}(y_j - \bar{g}_j) - \nabla_x g_j(y_j - \bar{g}_{j-1}),$$

we derive that

$$C_j \nabla_x = C_j \nabla_{x'} + D_j \partial_{y_j}, \quad C_j \partial_y = E_j \partial_{y_j}, \quad 1 \le j \le m.$$

As in [34] we rewrite the Laplace operator as

$$\begin{aligned} C_j^2 \Delta = {} & \mathrm{div}_{x'}\left[C_j^2 \nabla_{x'}\right] - (\nabla_{x'} C_j) \cdot [C_j \nabla_{x'}] + \partial_{y_j}[C_j D_j \cdot \nabla_{x'}] \\ & - (\partial_{y_j} D_j) \cdot [C_j \nabla_{x'}] + \mathrm{div}_{x'}\left[C_j D_j \partial_{y_j}\right] - (\nabla_{x'} C_j) \cdot [D_j \partial_{y_j}] \\ & + \partial_{y_j}\left[\left|D_j\right|^2 \partial_{y_j}\right] - (\partial_{y_j} D_j) \cdot [D_j \partial_{y_j}] - (\nabla_{x'} C_j) \cdot [C_j \nabla_{x'}] \\ & - (\nabla_{x'} C_j) \cdot [D_j \partial_{y_j}] + E_j^2 \partial_{y_j}^2. \end{aligned}$$

Then, the governing problem becomes

$$\begin{aligned} 0 & = C_j^2 \Delta' u + C_j^2 k_j^2 u \\ & = \mathrm{div}_{x'}\left[C_j^2 \nabla_{x'} u\right] + \partial_{y_j}[C_j D_j \cdot \nabla_{x'} u] + \mathrm{div}_{x'}\left[C_j D_j \partial_{y_j} u\right] \\ & \quad - (\nabla_{x'} C_j) \cdot [D_j \partial_{y_j} u] + \partial_{y_j}\left[\left|D_j\right|^2 \partial_{y_j} u\right] - (\nabla_{x'} C_j) \cdot [C_j \nabla_{x'} u] \\ & \quad + E_j^2 \partial_{y_j}^2 u + C_j^2 k_j^2 u. \end{aligned}$$

Setting $C_j^2(x) = E_j^2 + F_j(x)$, we find

$$\begin{aligned} \Delta' u + k_j^2 u = {} & \frac{1}{E_j^2} \mathrm{div}_{x'}\left[-F_j \nabla_{x'} u - C_j D_j \partial_{y_j} u\right] \\ & + \frac{1}{E_j^2} \partial_{y_j}\left[-C_j D_j \cdot \nabla_{x'} u - \left|D_j\right|^2 \partial_{y_j} u\right] \\ & + \frac{1}{E_j^2}[-k_j^2 F_j u + (\nabla_{x'} C_j) \cdot (D_j \partial_{y_j} u + C_j \nabla_{x'} u)]. \end{aligned} \tag{A.1}$$

Considering our boundary perturbation method, we introduce ε from $g_j = \varepsilon \psi_j$ and rearrange the right hand side of (A.1). For $1 \le j \le m$, we set

$$\begin{aligned} & C_j = (\bar{g}_j - \bar{g}_{j-1}) + \varepsilon(\psi_j - \psi_{j-1}), \\ & D_j = \varepsilon\left[\nabla_{x'} \psi_{j-1}(y_j - \bar{g}_j) - \nabla_{x'} \psi_j (y_j - \bar{g}_{j-1})\right] =: \varepsilon \tilde{D}_j, \\ & E_j = \bar{g}_j - \bar{g}_{j-1}, \\ & F_j = 2\varepsilon(\bar{g}_j - \bar{g}_{j-1})(\psi_j - \psi_{j-1}) + \varepsilon^2(\psi_j - \psi_{j-1})^2. \end{aligned}$$

Hence (A.1) becomes

$$\Delta' u + k_j^2 u = \frac{1}{E_j^2}\left[\mathrm{div}_{x'}\left[R_j^x\right] + \partial_{y_j} R_j^y + R_j^0\right],$$

where

$$\begin{aligned} R_j^x = {} & -2\varepsilon(\bar{g}_j - \bar{g}_{j-1})(\psi_j - \psi_{j-1}) \nabla_{x'} u - \varepsilon^2 (\psi_j - \psi_{j-1})^2 \nabla_{x'} u \\ & - \varepsilon(\bar{g}_j - \bar{g}_{j-1}) \tilde{D}_j \partial_{y_j} u - \varepsilon^2 (\psi_j - \psi_{j-1}) \tilde{D}_j \partial_{y_j} u, \end{aligned} \tag{A.2a}$$

and

$$R_j^y = -\varepsilon(\bar{g}_j - \bar{g}_{j-1}) \tilde{D}_j \cdot \nabla_{x'} u - \varepsilon^2 (\psi_j - \psi_{j-1}) \tilde{D}_j \cdot \nabla_{x'} u - \varepsilon^2 \left|\tilde{D}_j\right|^2 \partial_{y_j} u, \tag{A.2b}$$

and

$$R_j^0 = -k_j^2\varepsilon^2(\psi_j - \psi_{j-1})u - 2\varepsilon k_j^2(\bar{g}_j - \bar{g}_{j-1})(\psi_j - \psi_{j-1})u$$
$$+ \varepsilon^2 \mathrm{div}_{x'}\left[(\psi_j - \psi_{j-1})\tilde{D}_j \partial_{y_j} u\right] + \varepsilon \mathrm{div}_{x'}\left[(\psi_j - \psi_{j-1})(\bar{g}_j - \bar{g}_{j-1})\nabla_{x'} u\right]$$
$$+ \varepsilon^2(\psi_j - \psi_{j-1})\mathrm{div}_{x'}\left[(\psi_j - \psi_{j-1})\nabla_{x'} u\right]. \tag{A.2c}$$

For the boundary terms, we begin by recalling

$$v^{\mathrm{inc}}(x, y) = e^{i\alpha\cdot x - i\beta^1 y},$$

so we can write

$$u_1 - u_2 = -e^{i\alpha\cdot x - i\beta(\bar{g}_1 + g_1(x))}, \qquad y = \bar{g}_1,$$
$$u_j - u_{j+1} = 0, \qquad y = \bar{g}_j, \quad 2 \le j \le m-1.$$

At $y = a$ and $y = b$ we transform the equations

$$\partial_y v_1 - T_1[v_1] = 0, \qquad y = a,$$
$$\partial_y v_m - T_m[v_m] = 0, \qquad y = b,$$

into

$$\partial_{y_j} u_j - \frac{C_j}{E_j} T_j[u_j] = 0, \quad j = 1, m,$$

which implies

$$\partial_{y_1} u_1 - T_1[u_1] = -\frac{g_1}{a - \bar{g}_1} T_1[u_1],$$
$$\partial_{y_m} u_m - T_m[u_m] = -\frac{g_{m-1}}{b - \bar{g}_{m-1}} T_m[u_m].$$

Finally, regarding the Neumann conditions

$$\partial_{N_{g_1}} v_1 - \tau_1^2 \partial_{N_{g_1}} v_2 = -\partial_{N_{g_1}} v^{\mathrm{inc}}$$
$$= (i\alpha\cdot\nabla_x g_1(x) + i\beta)e^{i\alpha\cdot x} - i\beta(\bar{g}_1 + g_1(x)), \qquad y = \bar{g}_1,$$
$$\partial_{N_{g_j}} v_j - \tau_j^2 \partial_{N_{g_j}} v_{j+1} = 0, \qquad y = \bar{g}_j,$$

for $2 \le j \le m-1$. At $y = \bar{g}_1 + g_1(x)$ we find that

$$\partial_{N_{g_1}} u_1 - \tau_1^2 \partial_{N_{g_1}} u_2 = (i\alpha\cdot\nabla_{x'} g_1 + i\beta)e^{i\alpha\cdot x - i\beta(\bar{g}_1 + g_1(x))},$$

which implies that

$$-(\nabla_{x'} g_1)\cdot\nabla_{x'} u_1 + \partial_y u_1 - \tau_1^2((-\nabla_{x'} g_1)\cdot\nabla_{x'} u_2 + \partial_y u_2)$$
$$= (i\alpha\cdot\nabla_{x'} g_1 + i\beta)e^{i\alpha\cdot x - i\beta(\bar{g}_1 + g_1(x))}.$$

Using the change of variables we deduce that

$$-\varepsilon(\nabla_{x'}\psi_1)\cdot[\nabla_{x'} u_1 + \frac{\varepsilon\tilde{D}_1}{C_1}\partial_{y_1} u_1] + \frac{E_1}{C_1}\partial_{y_1} u_1 - \tau_1^2[-\varepsilon(\nabla_{x'}\psi_1)\cdot(\nabla_{x'} u_2 + \frac{\varepsilon\tilde{D}_2}{C_2}\partial_{y_2} u_2) + \frac{E_2}{C_2}\partial_{y_2} u_2]$$
$$= (\varepsilon i\alpha\cdot\nabla_{x'}\psi_1 + i\beta)e^{i\alpha\cdot x - i\beta(g_1 + g_1(x))}. \tag{A.3}$$

Simplifying (A.3), we find that

$$\frac{E_1}{C_1}\partial_{y_1} u_1 - \tau_1^2 \frac{E_2}{C_2}\partial_{y_2} u_2 = J_1',$$

where

$$J_1' := \varepsilon(\nabla_{x'}\psi_1)\cdot(\nabla_{x'} u_1) + \varepsilon^2(\nabla_{x'}\psi_1)\cdot\frac{\tilde{D}_1}{C_1}\partial_{y_1} u_1 - \varepsilon\tau_1^2(\nabla_{x'}\psi_1)\cdot(\nabla_{x'} u_2)$$
$$-\varepsilon^2\tau_1^2(\nabla_{x'}\psi_1)\cdot\frac{\tilde{D}_2}{C_2}(\partial_{y_2} u_2) + (\varepsilon i\alpha\cdot(\nabla_{x'}\psi_1) + i\beta)e^{i\alpha\cdot x - i\beta\bar{g}_1}e^{-i\beta g_1(x)}.$$

Noting that

$$E_1C_2 = (a - \bar{g}_1)(\bar{g}_1 - \bar{g}_2) + (a - \bar{g}_1)(g_1 - g_2),$$
$$E_2C_1 = (a - \bar{g}_1)(\bar{g}_1 - \bar{g}_2) - g_1(\bar{g}_1 - \bar{g}_2),$$

we derive that

$$\partial_{y_1} u_1 - \tau_1^2 \partial_{y_2} u_2 = J_1,$$

where

$$J_1 = \frac{C_1C_2 J_1' - \varepsilon(a - \bar{g}_1)(\psi_1 - \psi_2)\partial_{y_1} u_1 - \tau_1^2 \varepsilon \psi(\bar{g}_1 - \bar{g}_2)\partial_{y_2} u_2}{(a - \bar{g}_1)(\bar{g}_1 - \bar{g}_2)}.$$

The term $C_1C_2J_1'$ in J_1 can be simplified for later computations

$$\begin{aligned} C_1C_2J_1' = &(a - \bar{g}_1)(\bar{g}_1 - \bar{g}_2)J_1' + \varepsilon J_1' \left((a - \bar{g}_1)(\psi_1 - \psi_2) + \psi_1(\bar{g}_2 - \bar{g}_1)\right) \\ &+ \varepsilon^2 \psi_1(\psi_2 - \psi_1)J_1'. \end{aligned}$$

Similarly, for $y = \bar{g}_j + g_j(x)$, $2 \le j \le m - 1$, we recall that

$$\partial_{N_{g_j}} v_j - \tau_j^2 \partial_{N_{g_j}} v_{j+1} = 0,$$

and this implies that

$$(-(\nabla_{x'} g_j) \cdot \nabla_{x'} v_j + \partial_y v_j) - \tau_j^2(-(\nabla_{x'} g_j) \cdot \nabla_{x'} v_{j+1} + \partial_y v_{j+1}) = 0.$$

By the change of variables, for $2 \le j \le m - 1$, we see that

$$\begin{aligned} &- \varepsilon(\nabla_{x'}\psi_j) \cdot [\nabla_{x'} u_j + \frac{\varepsilon \tilde{D}_j}{C_j}\partial_{y_j} u_j] + \frac{E_j}{C_j}\partial_{y_j} u_j \\ &- \tau_j^2 \left[-\varepsilon(\nabla_{x'}\psi_j) \cdot \left(\nabla_{x'} u_{j+1} + \frac{\varepsilon \tilde{D}_{j+1}}{C_{j+1}} \partial_{y_{j+1}} u_{j+1} \right) + \frac{E_{j+1}}{C_{j+1}} \partial_{y_{j+1}} u_{j+1} \right] = 0. \end{aligned}$$

Rearranging the equation, we deduce that

$$\frac{E_j}{C_j}\partial_{y_j} u_j - \tau_j^2 \frac{E_{j+1}}{C_{j+1}} \partial_{y_{j+1}} u_{j+1} = J_j',$$

where

$$\begin{aligned} J_j' = &\varepsilon(\nabla_{x'}\psi_j) \cdot (\nabla_{x'} u_j) + \varepsilon^2(\nabla_{x'}\psi_j) \cdot \frac{\tilde{D}_j}{C_j}\partial_{y_j} u_j \\ &- \varepsilon\tau_j^2(\nabla_{x'}\psi_j) \cdot \nabla_{x'} u_{j+1} - \varepsilon^2\tau_j^2(\nabla_{x'}\psi_j) \cdot \frac{\tilde{D}_{j+1}}{C_{j+1}}\partial_{y_{j+1}} u_{j+1}. \end{aligned}$$

Noting that

$$E_jC_{j+1} = (\bar{g}_j - \bar{g}_{j-1})(\bar{g}_{j+1} - \bar{g}_j) + (\bar{g}_j - \bar{g}_{j-1})(g_{j+1} - g_j),$$
$$E_{j+1}C_j = (\bar{g}_{j+1} - \bar{g}_j)(\bar{g}_j - \bar{g}_{j-1}) + (\bar{g}_{j+1} - \bar{g}_j)(g_j - g_{j-1}),$$

we find that

$$\partial_{y_j} u_j - \tau_j^2 \partial_{y_{j+1}} u_{j+1} = J_j,$$

where

$$J_j = \frac{C_jC_{j+1}J_j' - \varepsilon(\bar{g}_j - \bar{g}_{j-1})(\psi_{j+1} - \psi_j)\partial_{y_j} u_j}{(\bar{g}_j - \bar{g}_{j-1})(\bar{g}_{j+1} - \bar{g}_j)} + \frac{\varepsilon\tau_j^2(\bar{g}_{j+1} - \bar{g}_j)(\psi_j - \psi_{j-1})\partial_{y_{j+1}} u_{j+1}}{(\bar{g}_j - \bar{g}_{j-1})(\bar{g}_{j+1} - \bar{g}_j)}. \tag{A.4}$$

For later computations, we rewrite the term $C_jC_{j+1}J_j'$ as

$$\begin{aligned} C_jC_{j+1}J_j' = &J_j'(\bar{g}_j - \bar{g}_{j-1})(\bar{g}_{j+1} - \bar{g}_j) + \varepsilon J_j'(\psi_j - \psi_{j-1})(\bar{g}_{j+1} - \bar{g}_j) \\ &- \varepsilon J_j'(\psi_{j+1} - \psi_j)(\bar{g}_j - \bar{g}_{j-1}) + \varepsilon^2 J_j'(\psi_j - \psi_{j-1})(\psi_{j+1} - \psi_j). \end{aligned}$$

Hence, we arrive at (3).

For our boundary perturbation approach it is important to expand the inhomogeneities of these equations in power series in ε; see e.g. (5). For $1 \le j \le m$, we find

$$R_{j,n} = \frac{1}{E_j^2}\left[\mathrm{div}_{x'}\left[R^x_{j,n}\right] + \partial_{y_j} R^y_{j,n} + R^0_{j,n}\right], \tag{A.5a}$$

where

$$\begin{aligned} R^x_{j,n} &= -2(\bar{g}_j - \bar{g}_{j-1})(\psi_j - \psi_{j-1})\nabla_{x'} u_{j,n-1} - (\bar{g}_j - \bar{g}_{j-1})\tilde{D}_j \partial_{y_j} u_{j,n-1} \\ &\quad - (\psi_j - \psi_{j-1})^2 \nabla_{x'} u_{j,n-2} - (\psi_j - \psi_{j-1})\tilde{D}_j \partial_{y_j} u_{j,n-2}, \end{aligned} \tag{A.5b}$$

and

$$R^y_{j,n} = -(\bar{g}_j - \bar{g}_{j-1})\tilde{D}_j \cdot \nabla_{x'} u_{j,n-1} - (\psi_j - \psi_{j-1})\tilde{D}_j \cdot \nabla_{x'} u_{j,n-2} - \left|\tilde{D}_j\right|^2 \partial_{y_j} u_{j,n-2}, \tag{A.5c}$$

and

$$\begin{aligned} R^0_{j,n} &= -2k_j^2(\bar{g}_j - \bar{g}_{j-1})(\psi_j - \psi_{j-1}) u_{j,n-1} \\ &\quad + \mathrm{div}_{x'}\left[(\psi_j - \psi_{j-1})(\bar{g}_j - \bar{g}_{j-1})\nabla_{x'} u_{j,n-1}\right] \\ &\quad - k_j^2(\psi_j - \psi_{j-1})^2 u_{j,n-2} + \mathrm{div}_{x'}\left[(\psi_j - \psi_{j-1})\tilde{D}_j \partial_{y_j} u_{j,n-2}\right] \\ &\quad + (\psi_j - \psi_{j-1})\mathrm{div}_{x'}\left[(\psi_j - \psi_{j-1})\nabla_{x'} u_{j,n-2}\right]. \end{aligned} \tag{A.5d}$$

For the boundary terms, we write

$$J_{1,n} = \frac{C_1 C_2 J'_{1,n} - (a - \bar{g}_1)(\psi_1 - \psi_2)\partial_{y_1} u_{1,n-1} - \tau_1^2 \psi_1(\bar{g}_1 - \bar{g}_2)\partial_{y_2} u_{2,n-1}}{(a - \bar{g}_1)(\bar{g}_1 - \bar{g}_2)},$$

and, for $2 \le l \le m-1$,

$$J_{j,n} = \frac{C_j C_{j+1} J'_{j,n} - (\bar{g}_l - \bar{g}_{j+1})(\psi_{j+1} - \psi_j)\partial_{y_j} u_{j,n-1}}{(\bar{g}_j - \bar{g}_{j-1})(\bar{g}_{j+1} - \bar{g}_j)} + \frac{\tau_j^2(\psi_j - \psi_{j-1})(\bar{g}_{j+1} - \bar{g}_j)\partial_{y_{j+1}} u_{j+1,n-1}}{(\bar{g}_j - \bar{g}_{j-1})(\bar{g}_{j+1} - \bar{g}_j)}, \tag{A.6}$$

where

$$\begin{aligned} J'_{1,n} &= (\nabla_{x'}\psi_1)\cdot\nabla_{x'} u_{1,n-1} - \tau_1^2(\nabla_{x'}\psi_1)\cdot(\nabla_{x'} u_{2,n-1}) + (\nabla_{x'}\psi_1)\cdot\frac{\tilde{D}_1}{C_1}\partial_{y_1} u_{1,n-2} \\ &\quad - \tau_1^2(\nabla_{x'}\psi_1)\cdot\frac{\tilde{D}_2}{C_2}\partial_{y_2} u_{2,n-2} + i\alpha\cdot(\nabla_{x'}\psi_1)e^{i\alpha\cdot x'}e^{-i\beta\bar{g}_1}\frac{(-\beta\psi_1)^{n-1}}{(n-1)!} \\ &\quad + i\beta e^{i\alpha\cdot x'}e^{i\beta\bar{g}_1}\frac{(-i\beta\psi_1)^n}{n!}, \end{aligned}$$

and

$$\begin{aligned} J'_{j,n} &= (\nabla_{x'}\psi_j)\cdot(\nabla_{x'} u_{j,n-1}) - \tau_j^2(\nabla_{x'}\psi_j)\cdot\nabla_{x'} u_{j+1,n-1} \\ &\quad + (\nabla_{x'}\psi_j)\cdot\frac{\tilde{D}_j}{C_j}\partial_{y_j} u_{j,n-2} - \tau_j^2(\nabla_{x'}\psi_j)\cdot\frac{\tilde{D}_{j+1}}{C_{j+1}}\partial_{y_{j+1}} u_{j+1,n-2}, \end{aligned}$$

and hence we reformulate

$$\begin{aligned} C_1 C_2 J'_{1,n} &= (a - \bar{g}_1)(\bar{g}_1\bar{g}_2) J'_{1,n} + J'_{1,n-1}\left((a - \bar{g}_1)(\psi_1 - \psi_2) + \psi_1(\bar{g}_2 - \bar{g}_1)\right) \\ &\quad + \psi_1(\psi_2 - \psi_1) J'_{1,n-2}, \\ C_j C_{j+1} J'_{j,n} &= (\bar{g}_j - \bar{g}_{j-1})(\bar{g}_{j+1} - \bar{g}_j) J'_{j,n} + (\bar{g}_{j+1} - \bar{g}_j)(\psi_j - \psi_{j-1}) J'_{j,n-1} \\ &\quad + (\psi_{j+1} - \psi_j)(\bar{g}_j - \bar{g}_{j-1}) J'_{j,n-1} \\ &\quad + (\psi_j - \psi_{j-1})(\psi_{j+1} - \psi_j) J'_{j,n-2}. \end{aligned}$$

References

[1] C. Godrèche (Ed.), Solids Far from Equilibrium, Cambridge University Press, Cambridge, 1992.
[2] J. Virieux, S. Operto, An overview of full-waveform inversion in exploration geophysics, Geophysics 74 (6) (2009) WCC1–WCC26.
[3] F. Bleibinhaus, S. Rondenay, Effects of surface scattering in full-waveform inversion, Geophysics 74 (6) (2009) WCC69–WCC77.
[4] F. Natterer, F. Wübbeling, Mathematical Methods in Image Reconstruction, SIAM Monographs on Mathematical Modeling and Computation, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001.
[5] L.M. Brekhovskikh, Y.P. Lysanov, Fundamentals of Ocean Acoustics, Springer-Verlag, Berlin, 1982.
[6] H. Raether, Surface Plasmons on Smooth and Rough Surfaces and on Gratings, Springer, Berlin, 1988.
[7] S.A. Maier, Plasmonics: Fundamentals and Applications, Springer, New York, 2007.
[8] S. Enoch, N. Bonod, Plasmonics: From Basics to Advanced Topics, Springer Series in Optical Sciences, Springer, New York, 2012.
[9] J.C. Strikwerda, Finite Difference Schemes and Partial Differential Equations, 2nd edition, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2004.
[10] R.J. LeVeque, Finite Difference Methods for Ordinary and Partial Differential Equations: Steady-State and Time-Dependent Problems, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2007.
[11] C. Johnson, Numerical Solution of Partial Differential Equations by the Finite Element Method, Cambridge University Press, Cambridge, 1987.
[12] F. Ihlenburg, Finite Element Analysis of Acoustic Scattering, Springer-Verlag, New York, 1998.
[13] J.S. Hesthaven, T. Warburton, Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications, Texts in Applied Mathematics, vol. 54, Springer, New York, 2008.
[14] M.O. Deville, P.F. Fischer, E.H. Mund, High-Order Methods for Incompressible Fluid Flow, Cambridge Monographs on Applied and Computational Mathematics, vol. 9, Cambridge University Press, Cambridge, 2002.
[15] D. Gottlieb, S.A. Orszag, Numerical Analysis of Spectral Methods: Theory and Applications, CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 26, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1977.
[16] C. Canuto, M.Y. Hussaini, A.M. Quarteroni, T.A. Zang, Spectral Methods in Fluid Dynamics, Springer-Verlag, New York, 1988.
[17] J.P. Boyd, Chebyshev and Fourier Spectral Methods, 2nd edition, Dover Publications Inc., Mineola, NY, 2001.
[18] D.P. Nicholls, A method of field expansions for vector electromagnetic scattering by layered periodic crossed gratings, J. Opt. Soc. Am. A 32 (5) (2015) 701–709.
[19] D.P. Nicholls, S.-H. Oh, T.W. Johnson, F. Reitich, Launching surface plasmon waves via vanishingly small periodic gratings, J. Opt. Soc. Am. A 33 (3) (2016) 276–285.
[20] F. Sanchez-Sesma, L. Perez-Rocha, S. Chavez-Perez, Diffraction of elastic waves by three-dimensional surface irregularities, part II, Bull. Seismol. Soc. Am. 79 (1) (1989) 101–112.
[21] D. Colton, R. Kress, Inverse Acoustic and Electromagnetic Scattering Theory, 2nd edition, Springer-Verlag, Berlin, 1998.
[22] A. Barnett, L. Greengard, A new integral representation for quasi-periodic scattering problems in two dimensions, BIT Numer. Math. 51 (2011) 67–90.
[23] O. Bruno, S. Shipman, C. Turc, S. Venakides, Efficient evaluation of doubly periodic green functions in 3D scattering, including Wood anomaly frequencies, http://dx.doi.org/10.1098/rspa.2016.0255.
[24] J. Lai, M. Kobayashi, A. Barnett, A fast and robust solver for the scattering from a layered periodic structure containing multi-particle inclusions, J. Comput. Phys. 298 (2015) 194–208.
[25] L. Rayleigh, On the dynamical theory of gratings, Proc. R. Soc. Lond. A 79 (1907) 399–416.
[26] S.O. Rice, Reflection of electromagnetic waves from slightly rough surfaces, Commun. Pure Appl. Math. 4 (1951) 351–378.
[27] O. Bruno, F. Reitich, Numerical solution of diffraction problems: a method of variation of boundaries, J. Opt. Soc. Am. A 10 (6) (1993) 1168–1175.
[28] D.P. Nicholls, F. Reitich, Shape deformations in rough surface scattering: cancellations, conditioning, and convergence, J. Opt. Soc. Am. A 21 (4) (2004) 590–605.
[29] A. Malcolm, D.P. Nicholls, A field expansions method for scattering by periodic multilayered media, J. Acoust. Soc. Am. 129 (4) (2011) 1783–1793.
[30] D.P. Nicholls, J. Shen, A stable, high-order method for two-dimensional bounded-obstacle scattering, SIAM J. Sci. Comput. 28 (4) (2006) 1398–1419.
[31] Q. Fang, D.P. Nicholls, J. Shen, A stable, high-order method for three-dimensional bounded-obstacle scattering, J. Comput. Phys. 224 (2) (2007) 1145–1169.
[32] D.P. Nicholls, J. Shen, A rigorous numerical analysis of the transformed field expansion method, SIAM J. Numer. Anal. 47 (4) (2009) 2708–2734.
[33] Y. He, D.P. Nicholls, J. Shen, An efficient and stable spectral method for electromagnetic scattering from a layered periodic structure, J. Comput. Phys. 231 (8) (2012) 3007–3022.
[34] Y. Hong, D.P. Nicholls, A stable high-order perturbation of surfaces method for numerical simulation of diffraction problems in triply layered media, J. Comput. Phys. 330 (2017) 1043–1068.
[35] G.A. Baker Jr., P. Graves-Morris, Padé Approximants, 2nd edition, Cambridge University Press, Cambridge, 1996.
[36] J. Shen, Efficient spectral-Galerkin method. I. Direct solvers of second- and fourth-order equations using Legendre polynomials, SIAM J. Sci. Comput. 15 (6) (1994) 1489–1505.
[37] J. Shen, T. Tang, L.-L. Wang, Spectral Methods: Algorithms, Analysis and Applications, Springer Series in Computational Mathematics, vol. 41, Springer, Heidelberg, 2011.
[38] R. Petit (Ed.), Electromagnetic Theory of Gratings, Springer-Verlag, Berlin, 1980.
[39] N.A. Phillips, A coordinate system having some special advantages for numerical forecasting, J. Atmos. Sci. 14 (2) (1957) 184–185.
[40] J. Chandezon, D. Maystre, G. Raoult, A new theoretical method for diffraction gratings and its numerical application, J. Opt. 11 (7) (1980) 235–241.
[41] H.B. Keller, Accurate difference methods for linear ordinary differential systems subject to linear constraints, SIAM J. Numer. Anal. 6 (1969) 8–30.
[42] G.Q. Zhang, Integrated solutions of ordinary differential equation system and two-point boundary value problems. I. Integrated solution method, Math. Numer. Sin. 3 (3) (1981) 245–254.
[43] G.Q. Zhang, Integrated solutions of ordinary differential equations system and two-point boundary value problems. II. Stability, Math. Numer. Sin. 3 (4) (1981) 329–339.
[44] T. Cheng, P. Li, Y. Wang, Near-field imaging of perfectly conducting grating surfaces, J. Opt. Soc. Am. A 30 (2013) 2473–2481.
[45] D.P. Nicholls, F. Reitich, A new approach to analyticity of Dirichlet–Neumann operators, Proc. R. Soc. Edinb., Sect. A 131 (6) (2001) 1411–1433.
[46] O.A. Ladyzhenskaya, N.N. Ural'tseva, Linear and Quasilinear Elliptic Equations, Academic Press, New York, 1968.
[47] O. Bruno, F. Reitich, Numerical solution of diffraction problems: a method of variation of boundaries. II. Finitely conducting gratings, Padé approximants, and singularities, J. Opt. Soc. Am. A 10 (11) (1993) 2307–2316.
[48] O. Bruno, F. Reitich, Numerical solution of diffraction problems: a method of variation of boundaries. III. Doubly periodic gratings, J. Opt. Soc. Am. A 10 (12) (1993) 2551–2562.
[49] D.P. Nicholls, F. Reitich, T.W. Johnson, S.-H. Oh, Fast high-order perturbation of surfaces (HOPS) methods for simulation of multi-layer plasmonic devices and metamaterials, J. Opt. Soc. Am. A 31 (8) (2014) 1820–1831.
[50] D.P. Nicholls, F. Reitich, Stability of high-order perturbative methods for the computation of Dirichlet–Neumann operators, J. Comput. Phys. 170 (1) (2001) 276–298.
[51] D.P. Nicholls, F. Reitich, Shape deformations in rough surface scattering: improved algorithms, J. Opt. Soc. Am. A 21 (4) (2004) 606–621.