Noname manuscript No.
(will be inserted by the editor)

A Spectral Element Method with Transparent Boundary Condition for Periodic Layered Media Scattering

Ying He · Misun Min · David P. Nicholls

Received: March 3, 2015 / Accepted: date

Abstract We present a high-order spectral element method for solving layered media scattering problems featuring an operator that can be used to transparently enforce the far-field boundary condition. The incorporation of this Dirichlet-to-Neumann (DtN) map into the spectral element framework is a novel aspect of this work, and the resulting method can accommodate plane-wave radiation of arbitrary angle of incidence. In order to achieve this, the governing Helmholtz equations subject to quasi-periodic boundary conditions are rewritten in terms of periodic unknowns. We construct a spectral element operator to approximate the DtN map, thus ensuring nonreflecting outgoing waves on the artificial boundaries introduced to truncate the computational domain. We present an explicit formula that accurately computes the Fourier coefficients of the solution in the spectral element discretization space projected onto the boundary which is required by the DtN map. Our solutions are represented by the tensor product basis of one-dimensional Legendre-Lagrange interpolation polynomials based on the Gauss-Lobatto-Legendre grids. We study the scattered field in singly and doubly layered media with smooth and nonsmooth interfaces. We consider rectangular, triangular, and sawtooth interfaces that are accurately represented by the body-fitted quadrilateral elements. We use GMRES iteration to solve the resulting linear system, and we validate our results by demonstrating spectral convergence in comparison with exact solutions and the results of an alternative computational method.

Keywords Spectral element method · Transparent boundary condition · Dirichlet-to-Neumann map · Periodic layered media · Scattering

Ying He
Department of Mathematics, University of California, Davis, CA 95616
E-mail: yinghe@math.ucdavis.edu

Misun Min
Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL 60439
E-mail: mmin@mcs.anl.gov
David P. Nicholls
Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago, Chicago, IL 60607
E-mail: davidn@uic.edu

1 Introduction

Scattering problems involving layered media arise in many engineering applications in electromagnetics, optics, and acoustics. Over the years, robust and accurate simulation capability has received increased attention as a cost-effective tool for predictive measurement and analysis of modern physical systems. Highly accurate boundary treatment and flexibility to treat complex geometries are essential for solving layered media scattering problems arising in a broad range of applications.

Many competing numerical methods have been developed for these scattering problems, such as the boundary integral and boundary element methods [1,2]. These surface methods require discretization only at the layer interfaces, thus significantly reducing the number of unknowns to compute. With the correct choice of the Green's function, the far-field boundary conditions can be enforced exactly, and these methods can deliver highly accurate solutions with reduced operation counts. Such methods face a number of drawbacks, however, including the fact that inhomogeneities away from the layer interfaces cannot be accommodated and high-order accuracy can be realized only with specially designed quadrature nodes, because of singularities in the Green's function. Moreover, these methods typically give rise to a dense linear system of equations whose solution requires preconditioned iterative methods featuring accelerated matrix-vector products (e.g., fast multipole methods [3]).

As an alternative, boundary perturbation methods have been explored. Bruno and Reitich studied the method of field expansions [4–6], and Milder studied the method of operator expansions [7,8,10–12,9]. These methods also pose surface unknowns, thereby enjoying the favorable operation counts of surface integral methods, while avoiding the subtle quadrature rules, dense linear systems, and algorithms for matrix-vector product acceleration. However, these algorithms depend on strong cancellations that can result in ill-conditioning [17–19]. Nicholls and Reitich proposed an enhanced boundary perturbation algorithm, referred to as the method of transformed field expansions (TFE) [20,13], which does not rely on strong cancellations. In this approach, the resulting recursions can be used for a direct, rigorous demonstration of the strong convergence of the relevant perturbation expansions in an appropriate function space. Furthermore, these formulas were proven to be a stable and accurate numerical scheme for simulating scattering problems defined on layered periodic gratings. This was later generalized to the case of irregularly bounded obstacles [14,15], multiply layered media for vector electromagnetic scattering [21], and a rigorous numerical analysis was provided in [13]. However, this method is limited when complex geometries and nonhomogeneous media are considered.

To address the limitations of these boundary methods, we consider a high-order spectral element method for layered media problems [22]. Of particular interest, in this paper we describe for the first time how a boundary operator, which transparently enforces the far-field boundary condition, can be incorporated into the spectral element framework. This is very much in the spirit of the DtN-FE method of Han and Wu [23] and Keller, Givoli, and Grote [24–29] and Nicholls and Nigam [30–32]. The relevant operator is the Dirichlet-to-Neumann (DtN) map [17–19], which, in our formulation, produces the normal derivative of the truncated Fourier series of the Dirichlet data on an artificial boundary introduced to truncate the computational domain [20]. We present a novel formula for computing the Fourier data in spectral element discretization space, in particular, we consider incident waves at arbitrary angles impinging on various types of periodic gratings, resulting in quasi-periodic solutions of the scalar Helmholtz equation. We rewrite our governing equation in a form that eliminates the quasi-periodicity and solve the reformulated scalar Helmholtz equation with periodic, Dirichlet, and transparent boundary conditions. We solve various example problems and demonstrate our computational results with validation. We note that the resulting linear system is not Hermitian positive definite and thus we resort to the generalized minimum residual (GMRES) method [36] for its solution.

This paper is organized as follows. In Section 2, we define the governing equations for our model problems and provide formulations. Section 3 discusses the spectral element discretization, while Section 4 presents the computational results and their validation. Section 5 summarizes our conclusion.

2 Problem Formulation

A downward-propagating time-harmonic incident plane wave of frequency ω can be expressed in complex form as

$$\bar{U}_{\mathrm{inc}}(x,y,t) = U_{\mathrm{inc}}(x,y)e^{-\mathbf{i}\omega t} = e^{\mathbf{i}\kappa\cdot\mathbf{x}}e^{-\mathbf{i}\omega t} = e^{\mathbf{i}(\alpha x-\beta y)}e^{-\mathbf{i}\omega t},$$

where the wave vector $\kappa = (\alpha, -\beta)$ with $\beta > 0$ defines the propagation direction. This will solve the scalar wave equation in a homogeneous medium with velocity c,

$$\frac{\partial^2\bar{U}}{\partial t^2} - c^2\Delta\bar{U} = 0, \tag{1}$$

if $|\kappa|^2 = \alpha^2+\beta^2 = \omega^2/c^2 =: k^2$. More generally, time-harmonic solutions of (1) can be written as

$$\bar{U}(x,y,t) = U(x,y)e^{-\mathbf{i}\omega t},$$

and the reduced field $U(x,y)$ satisfies a scalar Helmholtz equation at each frequency ω:

$$\Delta U + k^2U = 0. \tag{2}$$

To consider polychromatic waves, one can simply sum over different frequencies:

$$\bar{U}(x,y,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\mathbf{i}\omega t}U(x,y)\,d\omega.$$

Thus it suffices to work in the "frequency domain" by solving the Helmholtz equation as we do here.

2.1 Model Problems

In this paper we focus on singly and doubly layered media in two spatial dimensions as shown in Figure 1. We define the unbounded domains,

$$\Omega_0^+ = \{y > g(x)\} \quad \text{and} \quad \Omega_0^- = \{y < g(x)\}, \tag{3}$$

with wave numbers $k^{\pm} = \omega/c^{\pm}$ in $\Omega_0^{\pm}$, respectively. For the layer interface we consider a bounded, measurable, d–periodic function $g(x+d) = g(x)$ which specifies

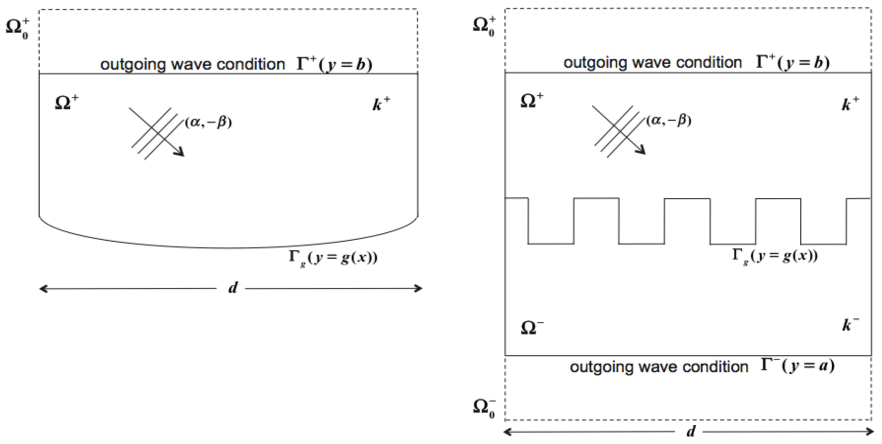
$$\Gamma_g = \{(x, y) \in \mathbb{R}^2 \, | y = g(x)\} \tag{4}$$

that gives the shape of the d-periodic grating structure. Here we note that the total reduced field is quasi-periodic [35]:

$$U(x + d, y) = e^{\mathbf{i}\alpha d} U(x, y).$$

For the single-layer model shown in Figure 1(a), a homogeneous Dirichlet boundary condition is specified on Γ_g, denoted Γ_D, and the scattered waves must be outgoing as $y \to +\infty$. The Dirichlet boundary can be interpreted as an impenetrable lower layer medium while the single-layer would be interpreted as the upper layer having most of scattering phenomena. Thus the single-layer model can be also considered as double-layered medium. For the double-layer model shown in Figure 1(b), the total field is required to be continuous across the scatterer interface Γ_g, and the scattered waves must be outgoing as $y \to \pm\infty$.

These model problems can be described by the Helmholtz equation with proper boundary conditions as follows.



(a) Model 1: single layer

(b) Model 2: double layer

Fig. 1 Geometric illustration of the model problems.

Model 1. The total field U in the single layer $\Omega_0 := \Omega_0^+$ satisfies

$$\Delta U + k^2 U = 0 \qquad \text{on} \quad \Omega_0, \tag{5}$$

$$U(x + d, y) = e^{\mathbf{i}\alpha d} U(x, y) \qquad \text{on} \quad \Omega_0, \tag{6}$$

$$U(x, y) = 0 \qquad \text{on} \quad \Gamma_D. \tag{7}$$

Model 2. The total field U in the double layer $\Omega_0 := \Omega_0^+ \cup \Omega_0^-$ satisfies

$$\Delta U + k^2 U = 0 \qquad \text{on } \Omega_0, \tag{8}$$

$$U(x+d,y) = e^{\mathbf{i}\alpha d} U(x,y) \qquad \text{on } \Omega_0, \tag{9}$$

where $k = k^\pm$ on $\Omega_0^\pm$.

2.2 Periodic Formulation

In normal incidence ($\alpha = 0$) the solution U is x-periodic, however, for oblique incidence ($\alpha \neq 0$) the solution is quasi-periodic in the x-direction. In our algorithm implementation, it is more convenient to handle periodic boundary conditions. Thus we introduce a new variable u by

$$u(x,y) = e^{-\mathbf{i}\alpha x} U(x,y), \tag{10}$$

where u is periodic for any α from the fact that

$$u(x+d,y) = e^{-\mathbf{i}\alpha(x+d)} U(x+d,y) = e^{-\mathbf{i}\alpha(x+d)} e^{\mathbf{i}\alpha d} U(x,y) = u(x,y).$$

Plugging (10) into Eqs. (5) and (8), we find

$$\Delta u + (k^2 - \alpha^2) u + 2\mathbf{i}\alpha \frac{\partial u}{\partial x} = 0. \tag{11}$$

We note that the first-order derivative term in Eq. (11) results from the quasi-periodicity of the solution U with $\alpha \neq 0$. This quasi-periodic term introduces new operators to our formulation in addition to the usual Helmholtz operator.

2.3 Transparent Boundary Conditions

Separation of variables applied to the Helmholtz equations yields the following periodic solutions, which are valid away from the interface (outside the grating grooves):

$$u^\pm(x,y) = \sum_{p=-\infty}^{\infty} \left\{ A_p e^{\mathbf{i}\beta_p^\pm y} + B_p e^{-\mathbf{i}\beta_p^\pm y} \right\} e^{\mathbf{i}(\alpha_p - \alpha)x}, \tag{12}$$

where $\alpha_p = \alpha + \frac{2\pi p}{d}$ and $(\beta_p^\pm)^2 = (k^\pm)^2 - \alpha_p^2$ for integer p. Defining the set of propagating modes

$$\mathbb{K}^\pm := \left\{ p \in \mathbb{Z} \mid (k^\pm)^2 - \alpha_p^2 > 0 \right\},$$

we have

$$\beta_p^\pm = \sqrt{(k^\pm)^2 - \alpha_p^2}, \quad p \in \mathbb{K}^\pm \quad \text{and} \quad \beta_p^\pm = \mathbf{i}\sqrt{\alpha_p^2 - (k^\pm)^2}, \quad p \notin \mathbb{K}^\pm.$$

Assuming incident plane-wave radiation $U_{\text{inc}} = e^{\mathbf{i}(\alpha x - \beta y)}$ in Ω_0^+ and $U_{\text{inc}} \equiv 0$ in Ω_0^-, we have $u_{\text{inc}} = e^{-\mathbf{i}\beta y}$ in Ω_0^+ and $u_{\text{inc}} \equiv 0$ in Ω_0^-. The total field thus u can be expressed as

$$u = \begin{cases} u_{\text{inc}} + u_{\text{scat}}^+ & \text{in } \Omega_0^+ \\ u_{\text{scat}}^- & \text{in } \Omega_0^- \end{cases}. \tag{13}$$

The scattered field is also of the form (12); however, the outgoing wave condition eliminates either A_p or B_p, so that

$$u^+_{\text{scat}} = \sum_{p=-\infty}^{\infty} \hat{u}^+_{\text{scat},p} e^{\mathbf{i}(\alpha_p-\alpha)x} e^{\mathbf{i}\beta^+_p y}, \quad u^-_{\text{scat}} = \sum_{p=-\infty}^{\infty} \hat{u}^-_{\text{scat},p} e^{\mathbf{i}(\alpha_p-\alpha)x} e^{-\mathbf{i}\beta^-_p y}, \quad (14)$$

which are the well-known Rayleigh expansions [33,16].

Now, we discuss a boundary operator that enforces the outgoing wave transparently at a boundary at *finite* distance from the interface. Consider the model problems defined on the finite computational domains, as shown in Figure 1:

$$\Omega^+ = \{g(x) < y < b, \quad 0 \le x \le d\} \quad \text{and} \quad \Omega^- = \{a < y < g(x), \quad 0 \le x \le d\}.$$

Here we define a hyperplane $\Gamma = \{(x, y) \in \mathbb{R}^2, \quad 0 \le x \le d, \quad y = c^*\}$ such that $\Gamma \cap \Gamma_g = \emptyset$, where the constant c^* can represent either a or b for our model problems. Without loss of generality we focus on $c^* = b$ in Ω^+ and drop the $+$ superscript. Taking the normal derivative of u_{scat} on Γ, we define the operator T by

$$T[u_{\text{scat}}] \mid_{y=c^*} := \frac{\partial u_{\text{scat}}}{\partial \mathbf{n}} \mid_{y=c^*} = \mathbf{n} \cdot \nabla u_{\text{scat}} \mid_{y=c^*},$$

where $\mathbf{n} = (n_x, n_y)$ is the outward unit normal vector. From (14), the Dirichlet-to-Neumann (DtN) map T can be expressed as

$$T[u_{\text{scat}}] \mid_{y=c^*} = n_y \frac{\partial u_{\text{scat}}}{\partial y} \mid_{y=c^*} = n_y \sum_{p=-\infty}^{\infty} (\mathbf{i}\beta_p) \hat{u}_{\text{scat},p} e^{\mathbf{i}(\alpha_p-\alpha)x} e^{\mathbf{i}\beta_p c^*}, \quad (15)$$

where $\mathbf{n} = (0, -1)$ is outward to $\{y > b\}$. We note that $\hat{u}_{\text{scat},p} e^{\mathbf{i}\beta_p c^*}$ is related to the one-dimensional Fourier coefficient of u_{scat} on Γ. Solutions of (11) and their normal derivatives are continuous across Γ, and the DtN map enforces this feature exactly by

$$T[u - u_{\text{inc}}] = \frac{\partial(u - u_{\text{inc}})}{\partial \mathbf{n}}$$

or

$$\partial_{\mathbf{n}} u - T[u] = \partial_y(u_{\text{inc}}) - T[u_{\text{inc}}] =: \rho.$$

2.4 Governing Equations

Defining $\Gamma^+ = \Gamma \subset \Omega_0^+$ (if $c^* = b$) and $\Gamma^- = \Gamma \subset \Omega_0^-$ (if $c^* = a$), we can summarize our governing equations for our model problems as follows.

Model 1. For the single-layer case, with $\Omega = \Omega^+$, our governing equations are

$$\Delta u + (k^2 - \alpha^2)u + 2\mathbf{i}\alpha \frac{\partial u}{\partial x} = 0 \qquad \text{on } \Omega, \quad (16)$$

$$u(x + d, y) = u(x, y) \qquad \text{on } \Omega, \quad (17)$$

$$\partial_{\mathbf{n}} u - T^+[u] = \rho \qquad \text{on } \Gamma^+, \quad (18)$$

$$u = 0 \qquad \text{on } \Gamma_D. \quad (19)$$

Model 2. For the double layer case, with $\Omega = \Omega^+ \cup \Omega^- \cup \Gamma_g$, our governing equations are

$$\Delta u + (k^2 - \alpha^2)u + 2\mathbf{i}\alpha \frac{\partial u}{\partial x} = 0 \qquad \text{on } \Omega, \tag{20}$$

$$u(x+d, y) = u(x, y) \qquad \text{on } \Omega, \tag{21}$$

$$\partial_{\mathbf{n}} u - T^+[u] = \rho \qquad \text{on } \Gamma^+, \tag{22}$$

$$\partial_{\mathbf{n}} u - T^-[u] = 0 \qquad \text{on } \Gamma^-. \tag{23}$$

2.5 Variational Formulation

In this section, we derive the variational formulations of our governing equations for the model problems (16)–(19) and (20)–(23). Consider a test function $v \in H^1_{\text{per}}(\Omega)$ where

$$H^1_{\text{per}}(\Omega) := \{\varphi \in H^1(\Omega) \mid \varphi(x+d, y) = \varphi(x, y)\}, \tag{24}$$

and $H^1(\Omega)$ is the classical Hilbert space of $L^2(\Omega)$ functions with one weak derivative in $L^2(\Omega)$. Multiplying (16) and (20) by v and integrating the results over Ω, whose boundary is denoted by $\partial\Omega$, we have

$$\int_\Omega \nabla u \cdot \nabla \overline{v} d\Omega - \int_{\partial\Omega} \mathbf{n} \cdot \nabla u \overline{v} dS - \int_\Omega (k^2 - \alpha^2) u \overline{v} d\Omega - \int_\Omega 2\mathbf{i}\alpha \frac{\partial u}{\partial x} \overline{v} d\Omega = 0. \tag{25}$$

The surface integrations with the boundary conditions applied on the single layer are

$$\int_{\partial\Omega} \mathbf{n} \cdot \nabla u \overline{v} dS = \int_{\Gamma^+} T^+[u] \overline{v} d\Gamma - \int_{\Gamma^+} \rho \overline{v} d\Gamma - \int_{\Gamma_D} \mathbf{n} \cdot \nabla u \overline{v} d\Gamma, \tag{26}$$

and those for the double layer are

$$\int_{\partial\Omega} \mathbf{n} \cdot \nabla u \overline{v} dS = \int_{\Gamma^+} T^+[u] \overline{v} d\Gamma - \int_{\Gamma^+} \rho \overline{v} d\Gamma + \int_{\Gamma^-} T^-[u] \overline{v} d\Gamma. \tag{27}$$

We seek a solution $u \in H^1_{\text{per}}(\Omega)$, shown to exist in [34], such that

$$a(u, v) = \langle \rho, v \rangle \qquad \text{for all } v \in H^1_{\text{per}}(\Omega), \tag{28}$$

where the sesquilinear form for each model is defined as follows.

Model 1. From (25) and (26), we have

$$a(u, v) = \int_\Omega \left(\nabla u \cdot \nabla \overline{v} - (k^2 - \alpha^2) u \overline{v} - 2\mathbf{i}\alpha \frac{\partial u}{\partial x} \overline{v} \right) d\Omega - \int_{\Gamma^+} T^+[u] \overline{v} d\Gamma. \tag{29}$$

Model 2. From (25) and (27), we have

$$a(u, v) = \int_\Omega \left(\nabla u \cdot \nabla \overline{v} - (k^2 - \alpha^2) u \overline{v} - 2\mathbf{i}\alpha \frac{\partial u}{\partial x} \overline{v} \right) d\Omega - \int_{\Gamma^+} T^+[u] \overline{v} d\Gamma - \int_{\Gamma^-} T^-[u] \overline{v} d\Gamma. \tag{30}$$

The linear functional $\langle \cdot, v \rangle$ in (28) is defined for both models as follows:

$$\langle \rho, v \rangle = \int_{\Gamma^+} \rho \overline{v} d\Gamma.$$

In particular, we define the following notation for the volume integrations:

$$A(u,v) = \int_\Omega \nabla u \cdot \nabla \overline{v} d\Omega, \quad B(u,v) = \int_\Omega u\overline{v} d\Omega, \quad C(u,v) = \int_\Omega \frac{\partial u}{\partial x}\overline{v} d\Omega, \tag{31}$$

and for the surface integrations:

$$T(u,v) = \int_\Gamma T[u]\overline{v} d\Gamma, \quad F(\rho, v) = \int_{\Gamma^+} \rho \overline{v} d\Gamma. \tag{32}$$

Here $\Gamma = \Gamma^+$ and $T = T^+$ for the single-layer geometry, and $\Gamma = \Gamma^+ \cup \Gamma^-$ and $T = \{T^+, T^-\}$ for the double-layer case. We note that in the upper layer

$$\begin{aligned} T(u,v) &= n_y \sum_{p=-\infty}^{\infty} \mathbf{i}\beta_p \hat{u}_p \int_\Gamma e^{\mathbf{i}d_p x}\overline{v} dx \\ &= n_y \sum_{p=-\infty}^{\infty} \mathbf{i}\beta_p \hat{u}_p \int_\Gamma \overline{e^{-\mathbf{i}d_p x} v} dx = n_y \sum_{p=-\infty}^{\infty} \mathbf{i}\beta_p \hat{u}_p \overline{\hat{v}_p}, \end{aligned} \tag{33}$$

where $d_p = \frac{2\pi p}{d} = \alpha_p - \alpha$ in Eq. (15). On the other hand, we have

$$\overline{T(v,u)} = n_y \sum_{p=-\infty}^{\infty} \overline{\mathbf{i}\beta_p \hat{v}_p \overline{\hat{u}_p}} = n_y \sum_{p=-\infty}^{\infty} (\overline{\mathbf{i}\beta_p}) \hat{u}_p \overline{\hat{v}_p}, \tag{34}$$

so there is no easily identified symmetry in the operator T.

3 Spectral Element Discretization

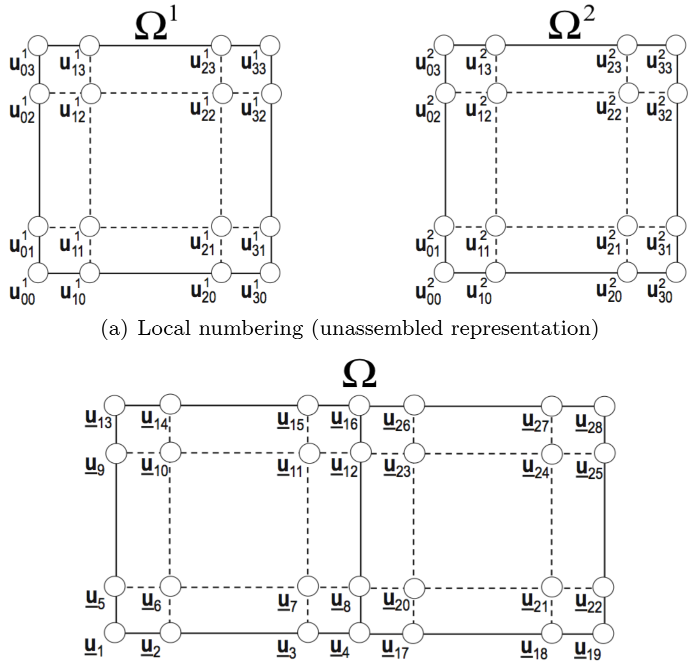
We denote our computational domain as $\Omega = \cup_{e=1}^{E} \Omega^e$, where Ω^e represents nonoverlapping body-conforming quadrilateral elements. Let us define a finite-dimensional approximation space $V_N \subset H^1(\Omega)$ such that $V_N = \text{span}\{\psi_{ij}(\xi,\eta)\}_{i,j=0}^{N}$. With this choice of approximation space, we consider a local approximate solution $u^e(x,y) \in V_N$, or simply u^e, that has the representation

$$u^e(x,y) = \sum_{i,j=0}^{N} u_{ij}^e \psi_{ij}(\xi,\eta). \tag{35}$$

The basis coefficients u_{ij}^e are the nodal values $u^e(x_i, y_j)$ on Ω^e, and the basis $\psi_{ij}(\xi,\eta) = \ell_i(\xi)\ell_j(\eta)$, or simply ψ_{ij}, has a tensor product form of the one-dimensional Nth-order Legendre-Lagrange interpolation polynomials given as

$$\ell_i(\xi) = [N(N+1)^{-1}(1-\xi^2)L_N'(\xi)]/[(\xi - \xi_i)L_N(\xi_i)] \quad \text{for } \xi \in [-1,1], \tag{36}$$

based on the Gauss-Lobatto-Legendre (GLL) quadrature nodes ξ_i with the Nth-order Legendre polynomial L_N and its derivative L_N'. We map each physical coordinate $(x,y) \in \Omega^e$ onto the reference domain $(\xi,\eta) \in I = [-1,1]^2$ through the Gordon-Hall mapping [22] and formulate the computational scheme on the reference domain.


Ω^1
$\mathbf{u}^1_{03}$ $\mathbf{u}^1_{13}$ $\mathbf{u}^1_{23}$ $\mathbf{u}^1_{33}$
$\mathbf{u}^1_{02}$ $\mathbf{u}^1_{12}$ $\mathbf{u}^1_{22}$ $\mathbf{u}^1_{32}$
$\mathbf{u}^1_{01}$ $\mathbf{u}^1_{11}$ $\mathbf{u}^1_{21}$ $\mathbf{u}^1_{31}$
$\mathbf{u}^1_{00}$ $\mathbf{u}^1_{10}$ $\mathbf{u}^1_{20}$ $\mathbf{u}^1_{30}$
Ω^2
$\mathbf{u}^2_{03}$ $\mathbf{u}^2_{13}$ $\mathbf{u}^2_{23}$ $\mathbf{u}^2_{33}$
$\mathbf{u}^2_{02}$ $\mathbf{u}^2_{12}$ $\mathbf{u}^2_{22}$ $\mathbf{u}^2_{32}$
$\mathbf{u}^2_{01}$ $\mathbf{u}^2_{11}$ $\mathbf{u}^2_{21}$ $\mathbf{u}^2_{31}$
$\mathbf{u}^2_{00}$ $\mathbf{u}^2_{10}$ $\mathbf{u}^2_{20}$ $\mathbf{u}^2_{30}$
Ω
$\underline{\mathbf{u}}_{13}$ $\underline{\mathbf{u}}_{14}$ $\underline{\mathbf{u}}_{15}$ $\underline{\mathbf{u}}_{16}$ $\underline{\mathbf{u}}_{26}$ $\underline{\mathbf{u}}_{27}$ $\underline{\mathbf{u}}_{28}$
$\underline{\mathbf{u}}_{9}$ $\underline{\mathbf{u}}_{10}$ $\underline{\mathbf{u}}_{11}$ $\underline{\mathbf{u}}_{12}$ $\underline{\mathbf{u}}_{23}$ $\underline{\mathbf{u}}_{24}$ $\underline{\mathbf{u}}_{25}$
$\underline{\mathbf{u}}_{5}$ $\underline{\mathbf{u}}_{6}$ $\underline{\mathbf{u}}_{7}$ $\underline{\mathbf{u}}_{8}$ $\underline{\mathbf{u}}_{20}$ $\underline{\mathbf{u}}_{21}$ $\underline{\mathbf{u}}_{22}$
$\underline{\mathbf{u}}_{1}$ $\underline{\mathbf{u}}_{2}$ $\underline{\mathbf{u}}_{3}$ $\underline{\mathbf{u}}_{4}$ $\underline{\mathbf{u}}_{17}$ $\underline{\mathbf{u}}_{18}$ $\underline{\mathbf{u}}_{19}$

(a) Local numbering (unassembled representation)

(b) Global numbering (assembled representation)

Fig. 2 Illustration of a solution vector in a local numbering and a global numbering, using an example mesh with $(E, N) = (2, 3)$: $\Omega = \Omega^1 \cup \Omega^2$ and the GLL nodes ($\circ$).

Let us denote our numerical solution $\mathbf{u}$ on Ω by the vector

$$\mathbf{u} := (u_1, u_2, ..., u_{\hat{l}}, ..., u_n) := (u^1, u^2, ..., u^e, ..., u^E)^T, \tag{37}$$

$$u^e := (u_1^e, u_2^e, ..., u_l^e, ..., u_{(N+1)(N+1)}^e)^T := (u_{00}^e, u_{10}^e, ..., u_{ij}^e, ..., u_{NN}^e)^T, \tag{38}$$

where $n = E(N+1)^2$ is the total number of basis coefficients, and $\hat{l} = 1 + i + j(N+1) + (e-1)(N+1)^2$ and $l = 1 + i + j(N+1)$ translate the two-index coefficient representation into a vector form, with the leading index i. In Figure 2, we show a mesh with two elements $E = 2$ including the GLL grids for $N = 3$ on $\Omega = \Omega^1 \cup \Omega^2$. Figure 2(a) illustrates a local ordering of a solution vector $\mathbf{u}$ based on the two-index expression in an unassembled representation for the coincident grids, $u_{3i}^1 = u_{0i}^2$ $(i = 0, ..., 3)$, appearing redundantly. In Figure 2(b), we demonstrate the same solution vector in a global ordering in an assembled representation using only the distinct nodes, denoted by

$$\underline{\mathbf{u}} = (\underline{u}_1, \underline{u}_2, ..., \underline{u}_{\bar{n}})^T. \tag{39}$$

The size $(\bar{n} < n)$ of the solution vector $\underline{\mathbf{u}}$ in the assembled representation is reduced after eliminating the redundancy from the coincident grids. In practice, our implementations are based on elementwise computations using the data structure in the local ordering. The global ordering is used when it is more convenient to describe our method in this paper.

3.1 Stiffness Matrices

To obtain the stiffness matrix, we consider the following inner product in Eq. (31):

$$A(u,v) = \int_{\Omega} \nabla u \cdot \nabla \overline{v} d\Omega = \int_{\Omega} \left(\frac{\partial u}{\partial x}\frac{\partial \overline{v}}{\partial x} + \frac{\partial u}{\partial y}\frac{\partial \overline{v}}{\partial y} \right) d\Omega, \tag{40}$$

where the partial derivatives are expressed by the chain rule for $x = x(\xi,\eta)$ and $y = y(\xi,\eta)$ on Ω^e:

$$\begin{aligned}
\frac{\partial u}{\partial x}\frac{\partial \overline{v}}{\partial x} &= \left(\frac{\partial u}{\partial \xi}\frac{\partial \xi}{\partial x} + \frac{\partial u}{\partial \eta}\frac{\partial \eta}{\partial x}\right)\left(\frac{\partial \overline{v}}{\partial \xi}\frac{\partial \xi}{\partial x} + \frac{\partial \overline{v}}{\partial \eta}\frac{\partial \eta}{\partial x}\right) \\
&= \frac{\partial u}{\partial \xi}\frac{\partial \overline{v}}{\partial \xi}\mathcal{G}^{\xi\xi}_{xx} + \frac{\partial u}{\partial \eta}\frac{\partial \overline{v}}{\partial \eta}\mathcal{G}^{\eta\eta}_{xx} + \frac{\partial u}{\partial \xi}\frac{\partial \overline{v}}{\partial \eta}\mathcal{G}^{\xi\eta}_{xx} + \frac{\partial u}{\partial \eta}\frac{\partial \overline{v}}{\partial \xi}\mathcal{G}^{\xi\eta}_{xx}, &(41)\\
\frac{\partial u}{\partial y}\frac{\partial \overline{v}}{\partial y} &= \left(\frac{\partial u}{\partial \xi}\frac{\partial \xi}{\partial y} + \frac{\partial u}{\partial \eta}\frac{\partial \eta}{\partial y}\right)\left(\frac{\partial \overline{v}}{\partial \xi}\frac{\partial \xi}{\partial y} + \frac{\partial \overline{v}}{\partial \eta}\frac{\partial \eta}{\partial y}\right) \\
&= \frac{\partial u}{\partial \xi}\frac{\partial \overline{v}}{\partial \xi}\mathcal{G}^{\xi\xi}_{yy} + \frac{\partial u}{\partial \eta}\frac{\partial \overline{v}}{\partial \eta}\mathcal{G}^{\eta\eta}_{yy} + \frac{\partial u}{\partial \xi}\frac{\partial \overline{v}}{\partial \eta}\mathcal{G}^{\xi\eta}_{yy} + \frac{\partial u}{\partial \eta}\frac{\partial \overline{v}}{\partial \xi}\mathcal{G}^{\xi\eta}_{yy}, &(42)
\end{aligned}$$

introducing the short notations for the geometric factors as $\mathcal{G}^{\xi\xi}_{xx} = \frac{\partial \xi}{\partial x}\frac{\partial \xi}{\partial x}$, $\mathcal{G}^{\eta\eta}_{xx} = \frac{\partial \eta}{\partial x}\frac{\partial \eta}{\partial x}$, $\mathcal{G}^{\xi\eta}_{xx} = \frac{\partial \xi}{\partial x}\frac{\partial \eta}{\partial x}$, $\mathcal{G}^{\xi\xi}_{yy} = \frac{\partial \xi}{\partial y}\frac{\partial \xi}{\partial y}$, $\mathcal{G}^{\eta\eta}_{yy} = \frac{\partial \eta}{\partial y}\frac{\partial \eta}{\partial y}$, and $\mathcal{G}^{\xi\eta}_{yy} = \frac{\partial \xi}{\partial y}\frac{\partial \eta}{\partial y}$. Using the expansion (35) for $u, v \in V_N$, we derive the discrete operator for (40) including (41)–(42) as

$$\begin{aligned}
\mathcal{A}^N(u,v) = &\sum_{e=1}^{E}\sum_{\hat{\imath},\hat{\jmath}=0}^{N}\sum_{i,j=0}^{N} \bar{v}^e_{\hat{\imath}\hat{\jmath}} \left(\int_I \frac{\partial \psi_{ij}}{\partial \xi}\frac{\partial \psi_{\hat{\imath}\hat{\jmath}}}{\partial \xi}\bar{\mathcal{G}}^{11} J d\mathbf{r} + \frac{\partial \psi_{ij}}{\partial \xi}\frac{\partial \psi_{\hat{\imath}\hat{\jmath}}}{\partial \eta}\bar{\mathcal{G}}^{12} J d\mathbf{r} \right) u^e_{ij} \\
+ &\sum_{e=1}^{E}\sum_{\hat{\imath},\hat{\jmath}=0}^{N}\sum_{i,j=0}^{N} \bar{v}^e_{\hat{\imath}\hat{\jmath}} \left(\int_I \frac{\partial \psi_{ij}}{\partial \eta}\frac{\partial \psi_{\hat{\imath}\hat{\jmath}}}{\partial \xi}\bar{\mathcal{G}}^{21} J d\mathbf{r} + \frac{\partial \psi_{ij}}{\partial \eta}\frac{\partial \psi_{\hat{\imath}\hat{\jmath}}}{\partial \eta}\bar{\mathcal{G}}^{22} J d\mathbf{r} \right) u^e_{ij}
\end{aligned} \tag{43}$$

where $Jd\mathbf{r} = Jd\xi d\eta$. On each local element, the Jacobian J and the geometric factors, defined by

$$\begin{aligned}
\bar{\mathcal{G}}^{11} &= (\mathcal{G}^{\xi\xi}_{xx} + \mathcal{G}^{\xi\xi}_{yy}),\ \ \bar{\mathcal{G}}^{12} = (\mathcal{G}^{\xi\eta}_{xx} + \mathcal{G}^{\xi\eta}_{yy}), &(44)\\
\bar{\mathcal{G}}^{21} &= (\mathcal{G}^{\xi\eta}_{xx} + \mathcal{G}^{\xi\eta}_{yy}),\ \ \bar{\mathcal{G}}^{22} = (\mathcal{G}^{\eta\eta}_{xx} + \mathcal{G}^{\eta\eta}_{yy}), &(45)
\end{aligned}$$

are introduced from the coordinate transformation and computed from the following relation:

$$J = \begin{vmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial x}{\partial \eta} \\ \frac{\partial y}{\partial \xi} & \frac{\partial y}{\partial \eta} \end{vmatrix} \quad \text{from} \quad \begin{pmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial x}{\partial \eta} \\ \frac{\partial y}{\partial \xi} & \frac{\partial y}{\partial \eta} \end{pmatrix} \begin{pmatrix} \frac{\partial \xi}{\partial x} & \frac{\partial \xi}{\partial y} \\ \frac{\partial \eta}{\partial x} & \frac{\partial \eta}{\partial y} \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{46}$$

We apply the numerical quadrature on the GLL nodes for the integrations in Eq. (43) as

$$\int_I \frac{\partial \psi_{ij}}{\partial \xi} \frac{\partial \psi_{\hat{i}\hat{j}}}{\partial \xi} \bar{\mathcal{G}}^{11} d\mathbf{r} = \sum_{k,m=0}^{N} \bar{\mathcal{G}}^{11}_{km} J_{km} w_k w_m l'_i(\xi_k) l_j(\eta_m) l'_{\hat{i}}(\xi_k) l_{\hat{j}}(\eta_m), \tag{47}$$

$$\int_I \frac{\partial \psi_{ij}}{\partial \xi} \frac{\partial \psi_{\hat{i}\hat{j}}}{\partial \eta} \bar{\mathcal{G}}^{12} d\mathbf{r} = \sum_{k,m=0}^{N} \bar{\mathcal{G}}^{12}_{km} J_{km} w_k w_m l'_i(\xi_k) l_j(\eta_m) l_{\hat{i}}(\xi_k) l'_{\hat{j}}(\eta_m), \tag{48}$$

$$\int_I \frac{\partial \psi_{ij}}{\partial \eta} \frac{\partial \psi_{\hat{i}\hat{j}}}{\partial \xi} \bar{\mathcal{G}}^{21} d\mathbf{r} = \sum_{k,m=0}^{N} \bar{\mathcal{G}}^{21}_{km} J_{km} w_k w_m l_i(\xi_k) l'_j(\eta_m) l'_{\hat{i}}(\xi_k) l_{\hat{j}}(\eta_m), \tag{49}$$

$$\int_I \frac{\partial \psi_{ij}}{\partial \eta} \frac{\partial \psi_{\hat{i}\hat{j}}}{\partial \eta} \bar{\mathcal{G}}^{22} d\mathbf{r} = \sum_{k,m=0}^{N} \bar{\mathcal{G}}^{22}_{km} J_{km} w_k w_m l_i(\xi_k) l'_j(\eta_m) l_{\hat{i}}(\xi_k) l'_{\hat{j}}(\eta_m), \tag{50}$$

where $\bar{\mathcal{G}}^{(\cdot)}_{km}$ and J_{km} represent the geometric values and the Jacobian at the nodal points, respectively, and w_k and w_m are the one-dimensional GLL quadrature weights. Note that $\bar{\mathcal{G}}^{12}_{km} = \bar{\mathcal{G}}^{21}_{km}$. We now have (40) in a discrete form as the following:

$$\mathcal{A}^N(u,v) = \sum_{e=1}^{E} (\bar{v}^e)^T \begin{bmatrix} \mathbf{D}_\xi \\ \mathbf{D}_\eta \end{bmatrix}^T \begin{bmatrix} \mathbf{G}^{11} & \mathbf{G}^{12} \\ \mathbf{G}^{21} & \mathbf{G}^{22} \end{bmatrix}^e \begin{bmatrix} \mathbf{D}_\xi \\ \mathbf{D}_\eta \end{bmatrix} u^e \tag{51}$$

$$= \sum_{e=1}^{E} (\bar{v}^e)^T \mathbf{D}^T \mathbf{G}^e \mathbf{D} u^e = \sum_{e=1}^{E} (\bar{v}^e)^T \mathbf{A}^e u^e, \tag{52}$$

where the differentiation matrices with respect to ξ and η, $\mathbf{D}_\xi$ and $\mathbf{D}_\eta$, respectively, are written as

$$\mathbf{D}_\xi = \mathbf{I} \otimes \hat{\mathbf{D}} \quad \text{and} \quad \mathbf{D}_\eta = \hat{\mathbf{D}} \otimes \mathbf{I}$$

in a tensor product form of the one-dimensional differentiation matrix $\hat{\mathbf{D}}_{ki} = l'_i(\xi_k)$ and the identity matrix $\mathbf{I}$ in $R^{(N+1)\times(N+1)}$. The entries of the one-dimensional differentiation matrix are $\hat{\mathbf{D}}_{ij} = \frac{L_N(\xi_i)}{L_N(\xi_j)(\xi_i - \xi_j)}$ $(i \neq j)$; $\hat{\mathbf{D}}_{00} = -\frac{(N+1)N}{4}$; $\hat{\mathbf{D}}_{NN} = \frac{(N+1)N}{4}$; $\hat{\mathbf{D}}_{ii} = 0$ $(0 < i < N)$, which is skew-centrosymmetric $\hat{\mathbf{D}}_{ij} = -\hat{\mathbf{D}}_{N-i,N-j}$. Equation (51) involves the pointwise multiplication of the nodal values $u^e = [u^e_l]$ by each diagonal component of $\mathbf{G}^{(\cdot)} = [\mathbf{G}^{(\cdot)}_l] = \text{diag}\{\bar{\mathcal{G}}^{(\cdot)}_{km} J_{km} w_k w_m\}$ for $l = k + (N+1)(m-1)$ on the nodal points on each local element Ω^e. Let us denote the stiffness matrix on Ω as $\mathbf{A}$, using the local stiffness matrices $\mathbf{A}^e$, represented by

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}^1 & & & & \\ & \ddots & & & \\ & & \mathbf{A}^e & & \\ & & & \ddots & \\ & & & & \mathbf{A}^E \end{bmatrix} \quad \text{with} \quad \mathbf{A}^e = \mathbf{D}^T \mathbf{G}^e \mathbf{D}. \tag{53}$$

Then we can write Eq. (51) simply as

$$\mathcal{A}^N(u,v) = \bar{\mathbf{v}}^T \mathbf{A} \mathbf{u}. \tag{54}$$

Here we note that the matrix $\mathbf{A}$ is symmetric from the fact that

$$(\mathbf{A}^e)^T = (\mathbf{D}^T\mathbf{G}^e\mathbf{D})^T = \mathbf{D}^T(\mathbf{G}^e)^T\mathbf{D} = \mathbf{D}^T(\mathbf{G}^e)\mathbf{D} = \mathbf{A}^e. \tag{55}$$

Arithmetic Operations: The matrix $\mathbf{A}$ is never explicitly formed. We perform matrix-matrix multiplication acting only on the block diagonal matrices $\mathbf{A}^e$. We begin with the tensor product–based derivative evaluations (51) that can be recast as matrix-matrix products on each element:

$$\mathrm{u}_\xi := (\mathbf{I} \otimes \hat{\mathbf{D}})u^e := \hat{\mathbf{D}}[u]^e, \tag{56}$$

$$\mathrm{u}_\eta := (\hat{\mathbf{D}} \otimes \mathbf{I})u^e := [u]^e\hat{\mathbf{D}}^T, \tag{57}$$

where u^e is a vector arranged in columnwise consecutive entries of u^e_{ij}, advancing with the leading index (i) as shown in (38). In (56), u^e is treated as an $(N+1) \times (N+1)$ matrix, denoted by $[u]^e$ as

$$[u]^e = \begin{bmatrix} u^e_{00} & u^e_{01} & \dots & u^e_{0N} \\ \vdots & \vdots & \ddots & \vdots \\ u^e_{N0} & u^e_{N1} & \dots & u^e_{NN} \end{bmatrix}. \tag{58}$$

This requires $2E(N+1)^3$ operations on Ω. The pointwise multiplications with the geometric factors $\mathrm{u}_x = \mathbf{G}^{11}\mathrm{u}_\xi + \mathbf{G}^{12}\mathrm{u}_\eta$ and $\mathrm{u}_y = \mathbf{G}^{21}\mathrm{u}_\xi + \mathbf{G}^{22}\mathrm{u}_\eta$ require $6E(N+1)^2$ operations. Then we compute the summation of transposed derivative operators, $\mathbf{D}_\xi\mathrm{u}_x + \mathbf{D}_\eta\mathrm{u}_y$, involving $4E(N+1)^3 + E(N+1)^2$ operations. Thus the total operation for $\mathbf{Au}$ is $6E(N+1)^3 + 7E(N+1)^2$. The leading-order storage requirement for the factored stiffness matrix is $3E(N+1)^2$, because of the relation $\mathbf{G}_{12} = \mathbf{G}_{21}$ on Ω^e.

Direct Stiffness Summation: The solution vector in (54) is based on the unassembled representation, recalling Figure 2(a), without applying the continuity at the element interface between neighboring elements. To construct the solution vector continuous across element interfaces on the coincident nodal values,

$$(x_{ij}, y_{ij})^e = (x_{\hat{i}\hat{j}}, y_{\hat{i}\hat{j}})^{\hat{e}} \to u^e_{ij} = u^{\hat{e}}_{\hat{i}\hat{j}} \quad \text{for} \quad e \neq \hat{e}, \tag{59}$$

we introduce a Boolean connectivity matrix $\mathbf{Q}$ [22] that maps the global representation $\underline{\mathbf{u}}$ to the local representation $\mathbf{u}$, and its transpose $\mathbf{Q}^T$ that maps the local representation $\mathbf{u}$ to the global representation $\underline{\mathbf{u}}$. Then we can define the following:

$$\mathbf{u} = \mathbf{Q}\underline{\mathbf{u}} \quad \text{and} \quad \mathbf{u}^* = \mathbf{Q}^T\mathbf{u}. \tag{60}$$

The action of $\mathbf{Q}$ on $\underline{\mathbf{u}}$ returns the copy entries of $\underline{\mathbf{u}}$ on the coincident nodes, referred to as the scatter operation. The action of $\mathbf{Q}^T$ on $\mathbf{u}$ returns $\mathbf{u}^*$ with the sum entries of $\mathbf{u}$ on the coincident nodes, referred to as the gather operation. The interior nodes are unchanged from both of the actions. Using these matrices, we can rewrite Eq. (54) for the continuous solution $\underline{\mathbf{u}}$ as

$$\mathcal{A}^N(u,v) = \underline{\bar{\mathbf{v}}}^T\mathbf{Q}^T\mathbf{A}\mathbf{Q}\underline{\mathbf{u}} = \underline{\bar{\mathbf{v}}}^T\bar{\mathbf{A}}\underline{\mathbf{u}}. \tag{61}$$

For a continuous solution $\mathbf{u}$ in the local ordering representation, the following equivalence holds:

$$\mathbf{Q}^T\mathbf{A}\mathbf{Q}\underline{\mathbf{u}} \iff \left(\mathbf{Q}\mathbf{Q}^T\right)\mathbf{A}\mathbf{u}. \tag{62}$$

We note that the gather-scatter operation $\mathbf{Q}\mathbf{Q}^T$ can be viewed as a single operation, involving summation of the variables on the shared interface nodes and redistribution of them to their original locations within one communication. The operation is referred to as *direct stiffness summation*, or simply *dssum*. In this paper, we use the following notation for the gather-scatter operation:

$$dssum := \mathbf{Q}\mathbf{Q}^T. \tag{63}$$

In practical implementations, we write our algorithms in an element-based format by utilizing matrix-vector products evaluated independently on each local element. Thus it is natural to consider the *dssum* approach and perform the local-to-local transformation as in the right-hand side of (62), that is, $dssum(\mathbf{Au})$. We build a local-to-global mapping array to handle the actions of $\mathbf{Q}$ and $\mathbf{Q}^T$ without constructing $\mathbf{Q}$ and $\mathbf{Q}^T$ explicitly. A detailed description of the algorithms and parallel implementations can be found in Chapter 4 and Chapter 8 of [22].

3.2 Mass Matrices

To obtain the mass matrix, we consider the following inner product:

$$\mathcal{B}(u,v) = \int_{\Omega} u\bar{v}\, d\Omega, \tag{64}$$

which can be discretized as

$$\begin{aligned}
\mathcal{B}^N(u,v) &= \sum_{e=1}^{E} \sum_{\hat{i},\hat{j}=0}^{N} \sum_{i,j=0}^{N} \bar{v}^e_{ij} \left(\int_{\Omega^e} \psi_{ij} \psi_{\hat{i}\hat{j}} d\Omega \right) u^e_{ij} \\
&= \sum_{e=1}^{E} \sum_{\hat{i},\hat{j}=0}^{N} \sum_{i,j=0}^{N} \bar{v}^e_{ij} \left(\int_{I} \psi_{ij} \psi_{\hat{i}\hat{j}} J d\mathbf{r} \right) u^e_{ij} \\
&= \sum_{e=1}^{E} \sum_{\hat{i},\hat{j}=0}^{N} \sum_{i,j=0}^{N} \bar{v}^e_{ij} \left(\sum_{k,m=0}^{N} J_{km} w_k w_m l_i(\xi_k) l_j(\eta_m) l_{\hat{i}}(\xi_k) l_{\hat{j}}(\eta_m) \right) u^e_{ij} \\
&= \sum_{e=1}^{E} (\bar{v}^e)^T \mathbf{J}^e \left(\hat{\mathbf{M}} \otimes \hat{\mathbf{M}} \right) u^e = \sum_{e=1}^{E} (\bar{v}^e)^T \mathbf{B}^e u^e,
\end{aligned} \tag{65}$$

where $\hat{\mathbf{M}} = \text{diag}\{w_k\}$ is the one-dimensional mass matrix and $\mathbf{J}^e = [\mathbf{J}^e_{ll}] = \text{diag}\{J_{km}\}$ for $l = k + (N+1)(m-1)$. We can denote the mass matrix $\mathbf{B}$, using the local mass matrices $\mathbf{B}^e$, as

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}^1 & & & & \\ & \ddots & & & \\ & & \mathbf{B}^e & & \\ & & & \ddots & \\ & & & & \mathbf{B}^E \end{bmatrix} \quad \text{with} \quad \mathbf{B}^e = \mathbf{J}^e (\hat{\mathbf{M}} \otimes \hat{\mathbf{M}}), \tag{66}$$

which is fully diagonal. Then we can write Eq. (65) simply as

$$\mathcal{B}^N(u,v) = \bar{\mathbf{v}}^T\mathbf{B}\mathbf{u}. \tag{67}$$

For a continuous solution, Eq. (67) in the assembled representation can be expressed as

$$\mathcal{B}^N(u,v) = \underline{\bar{\mathbf{v}}}^T\mathbf{Q}^T\mathbf{B}\mathbf{Q}\underline{\mathbf{u}} = \underline{\bar{\mathbf{v}}}^T\bar{\mathbf{B}}\underline{\mathbf{u}}. \tag{68}$$

3.3 The Quasi-Periodic Matrix

We consider the following inner product for the quasi-periodic operator in Eq. (25):

$$C(u,v) = \int_{\Omega}\frac{\partial u}{\partial x}\bar{v}\,d\Omega, \tag{69}$$

which can be discretized as

$$\begin{aligned}
\mathcal{C}^N(u,v) &= \sum_{e=1}^{E}\sum_{\hat{\imath},\hat{\jmath}=0}^{N}\sum_{i,j=0}^{N}\bar{v}^e_{\hat{\imath}\hat{\jmath}}\left(\int_{\Omega^e}\frac{\partial\psi_{ij}}{\partial x}\psi_{\hat{\imath}\hat{\jmath}}\,d\Omega\right)u^e_{ij}\\
&= \sum_{e=1}^{E}\sum_{\hat{\imath},\hat{\jmath}=0}^{N}\sum_{i,j=0}^{N}\bar{v}^e_{\hat{\imath}\hat{\jmath}}\left(\int_{I}\frac{\partial\psi_{ij}}{\partial x}\psi_{\hat{\imath}\hat{\jmath}}J\,d\mathbf{r}\right)u^e_{ij}\\
&= \sum_{e=1}^{E}\sum_{\hat{\imath},\hat{\jmath}=0}^{N}\sum_{i,j=0}^{N}\bar{v}^e_{\hat{\imath}\hat{\jmath}}\left(\sum_{k,m=0}^{N}J_{km}w_k w_m l'_i(\xi_k)l_j(\eta_m)l_{\hat{\imath}}(\xi_k)l_{\hat{\jmath}}(\eta_m)\right)u^e_{ij}\\
&= \sum_{e=1}^{E}(\bar{v}^e)^T\mathbf{J}^e\left(\hat{\mathbf{M}}\otimes\hat{\mathbf{M}}\hat{\mathbf{D}}\right)u^e = \sum_{e=1}^{E}(\bar{v}^e)^T\mathbf{J}^e\left(\hat{\mathbf{M}}\otimes\hat{\mathbf{C}}\right)u^e\\
&= \sum_{e=1}^{E}(\bar{v}^e)^T\mathbf{C}^e u^e.
\end{aligned} \tag{70}$$

By convention, (69) could be referred as a convective operator in computational fluids. In this context, this relates to the quasi-periodic term in (11) for oblique incidence ($\alpha \neq 0$). Thus we refer to it as a quasi-periodic operator, because the operator is a derivative, resulting from imposing the periodicity for the quasi-periodic solution in (10). Then, we define the quasi-periodic matrix $\mathbf{C}$ on Ω using the local quasi-periodic matrices as

$$\mathbf{C} = \begin{bmatrix}\mathbf{C}^1 & & & & \\ & \ddots & & & \\ & & \mathbf{C}^e & & \\ & & & \ddots & \\ & & & & \mathbf{C}^E\end{bmatrix} \quad \text{with} \quad \mathbf{C}^e = \mathbf{J}^e(\hat{\mathbf{M}}\otimes\hat{\mathbf{C}}). \tag{71}$$

We can write Eq. (70) simply as

$$\mathcal{C}^N(u,v) = \bar{\mathbf{v}}^T\mathbf{C}\mathbf{u}. \tag{72}$$

For the continuous solution, Eq. (72) in the assembled representation can be expressed as

$$\mathcal{C}^N(u,v) = \underline{\bar{\mathbf{v}}}^T\mathbf{Q}^T\mathbf{C}\mathbf{Q}\underline{\mathbf{u}} = \underline{\bar{\mathbf{v}}}^T\bar{\mathbf{C}}\underline{\mathbf{u}}. \tag{73}$$

3.4 Spectral Element Dirichlet-to-Neumann Operator

In this section, we formulate a spectral element discretization for the Dirichlet-to-Neumann (DtN) map T. For simplicity we consider the operator T in the upper layer, Ω_0^+ and assume that the (ξ,η) coordinates of the element are aligned with (x,y). Let us denote $\Gamma = \cup_{\hat{e}=1}^{\hat{E}}\Gamma^{\hat{e}}$, where $\Gamma^{\hat{e}} = \Omega^{\hat{e}} \cap \partial\Omega$ are nonoverlapping boundary surfaces on the local elements $\Omega^{\hat{e}}$. We define a DtN-to-local mapping array that contains the indices of the transparent boundary surface nodes $(i,j,\hat{e})$ to the local index $(i,j,e) :=$ DtN-to-local$(i,j,\hat{e})$. We note that these nodes in y fall on the index either with $j=0$ or with $j=N$, which will be represented simply by a fixed index as $j=j_b$.

DtN Matrix T: We can represent our approximate solution on $\Gamma^{\hat{e}}$ in the form of (35) as

$$u^{\hat{e}}(x,b) = \sum_{i,j=0}^{N} u_{ij}^{\hat{e}} l_i(\xi) l_j(\eta(b)) = \sum_{i=0}^{N} u_{ij_b}^{\hat{e}} l_i(\xi). \tag{74}$$

From Eqs. (32)–(33), we have

$$T(u,v) = \int_\Gamma T[u]\bar{v}d\Gamma = \sum_{p=-\infty}^{\infty} \mathbf{i}\beta_p \hat{u}_p \int_\Gamma e^{\mathbf{i}d_p x}\bar{v}dx, \tag{75}$$

where $d_p = \frac{2\pi p}{d}$ and $\hat{u}_p$ are the one-dimensional Fourier coefficients of $u(x,b)$ on Γ given as

$$\hat{u}_p = \frac{1}{d}\int_0^d u(x',b)e^{-\mathbf{i}d_p x'}dx' \approx \frac{1}{d}\sum_{\hat{e}=1}^{\hat{E}}\int_{\Gamma^{\hat{e}}} u^{\hat{e}}(x',b)e^{-\mathbf{i}d_p x'}dx'. \tag{76}$$

Plugging (76) into (75) with a finite expansion of $T[u]$ ($|p| \le P$) and applying (74), we have

$$\begin{aligned} T^N(u,v) &= \sum_{p=-P}^{P} \mathbf{i}\beta_p \left(\frac{1}{d}\sum_{\hat{e}=1}^{\hat{E}}\int_{\Gamma^{\hat{e}}} u^{\hat{e}}(x',b)e^{-\mathbf{i}d_p x'}dx'\right)\left(\sum_{\bar{e}=1}^{\hat{E}}\int_{\Gamma^{\bar{e}}} e^{\mathbf{i}d_p x}\bar{v}dx\right) \\ &= \sum_{p=-P}^{P} \mathbf{i}\beta_p \left(\frac{1}{d}\sum_{\hat{e}=1}^{\hat{E}}\sum_{i=0}^{N} u_{ij_b}^{\hat{e}} \int_{\Gamma^{\hat{e}}} l_i(\xi)e^{-\mathbf{i}d_p x'}dx'\right)\left(\sum_{\bar{e}=1}^{\hat{E}}\int_{\Gamma^{\bar{e}}} e^{\mathbf{i}d_p x}\bar{v}dx\right). \end{aligned}$$

Choosing $\bar{v} = l_{\hat{i}}(\xi)$ with a different index set of $\hat{i}$ on each $\Omega^{\hat{e}}$ and defining the following,

$$s_i^{\hat{e},p} = \frac{1}{\sqrt{d}}\int_{\Gamma^{\hat{e}}} l_i(\xi)e^{-\mathbf{i}d_p x'}dx' \quad \text{and} \quad s_{\hat{i}}^{\bar{e},-p} = \frac{1}{\sqrt{d}}\int_{\Gamma^{\bar{e}}} l_{\hat{i}}(\xi)e^{\mathbf{i}d_p x}dx, \tag{77}$$

we can express (75) in a simplified form as

$$T^N(u,v) = \sum_{\hat{e}=1}^{\hat{E}}\sum_{i=0}^{N} u^{\hat{e}}_{ij_b}\left[\sum_{p=-P}^{P} \mathbf{i}\beta_p s_i^{\hat{e},p}\left(\sum_{\bar{e}=1}^{\hat{E}} s_{\hat{i}}^{\bar{e},-p}\right)\right] = \sum_{\hat{e}=1}^{\hat{E}}\sum_{i=0}^{N} u^{\hat{e}}_{ij_b}T^{\hat{e}}_{\hat{i}i}. \tag{78}$$

Here we note that $s_i^{\hat{e},-p}$ is the complex conjugate of $s_i^{\hat{e},p}$ from the following:

$$\overline{s_i^{\hat{e},p}} = \overline{\frac{1}{\sqrt{d}}\int_{\Gamma^{\hat{e}}} l_i(\xi)e^{-\mathbf{i}d_p x}dx} = \frac{1}{\sqrt{d}}\int_{\Gamma^{\hat{e}}} l_i(\xi)e^{\mathbf{i}d_p x}dx = s_i^{\hat{e},-p}.$$

Thus we need only to compute $s_i^{\hat{e},p}$ for $p \geq 0$ to obtain

$$T^{\hat{e}}_{\hat{i}i} = \mathbf{i}\left(\beta_0 s_i^{\hat{e},0}\sum_{\bar{e}=1}^{\hat{E}} s_{\hat{i}}^{\bar{e},0} + \sum_{p=1}^{P}\left[\beta_p s_i^{\hat{e},p} + \beta_{-p}\overline{s_i^{\hat{e},p}}\right]\sum_{\bar{e}=1}^{\hat{E}}\overline{s_{\hat{i}}^{\bar{e},p}}\right), \tag{79}$$

where $\beta_p = \beta_{-p}$ only if $\alpha = 0$; $\beta_p \neq \beta_{-p}$ for $\alpha \neq 0$. Therefore, no particular relation can be found between β_p and β_{-p} in general. Here $T^{\hat{e}}_{\hat{i}i}$ is a complex number, so we can alternatively write (78) as

$$T^N(u,v) = \sum_{\hat{e}=1}^{\hat{E}}\sum_{i=0}^{N} u^{\hat{e}}_{ij_b}T^{\hat{e}}_{\hat{i}i} = \sum_{\hat{e}=1}^{\hat{E}}\sum_{i=0}^{N} u^{\hat{e}}_{ij_b}\left[(T^{\hat{e}}_{\hat{i}i})_{\mathrm{real}} + \mathbf{i}(T^{\hat{e}}_{\hat{i}i})_{\mathrm{imag}}\right].$$

Now, we can map the values of $T^{\hat{e}}_{\hat{i}i}$ into a matrix $\mathbf{T}^e = \left[\mathbf{T}^e_{\hat{l}l}\right]$ for $\hat{l} = \hat{i} + (N+1)j$ and $l = i + (N+1)j$ from the DtN-to-local mapping $(\hat{i}, j, e) :=$ DtN-to-local$(\hat{i}, j_b, \hat{e})$ and $(i, j, e) :=$ DtN-to-local$(i, j_b, \hat{e})$. Similarly, $\{u^{\hat{e}}_{ij_b}\}$ can be mapped to the local data $\{u^e_{ij}\}$. Note that the entries of $\mathbf{T}^e$ are zeros if the indices are not indicating the DtN boundary nodes. We now have Eq. (78) in the local representation form as

$$\mathcal{T}^N(u,v) = \sum_{e=1}^{E}(v^e)^T\mathbf{T}^e u^e = \mathbf{v}^T\mathbf{T}\mathbf{u} = \mathbf{v}^T(\mathbf{T}_{\mathrm{r}} + \mathbf{i}\mathbf{T}_{\mathrm{i}})\mathbf{u}, \tag{80}$$

where $\mathbf{T}_{\mathrm{r}}$ and $\mathbf{T}_{\mathrm{i}}$ represent the real and imaginary part of the complex matrix $\mathbf{T}$. Thus we have the assembled representation of (80) as

$$\mathcal{T}^N(u,v) = \underline{\mathbf{v}}^T\mathbf{Q}^T\mathbf{T}\mathbf{Q}\underline{\mathbf{u}} = \underline{\mathbf{v}}^T\bar{\mathbf{T}}\underline{\mathbf{u}} = \underline{\mathbf{v}}^T(\bar{\mathbf{T}}_{\mathrm{r}} + \mathbf{i}\bar{\mathbf{T}}_{\mathrm{i}})\underline{\mathbf{u}}.$$

For ρ in (32), we apply notation similar to that used for u. Then we have the following:

$$\mathcal{F}^N(\rho,v) = \sum_{e=1}^{E}(v^e)^T\mathbf{B}^e\rho^e = \mathbf{v}^T\mathbf{B}\boldsymbol{\rho} = \mathbf{v}^T\mathbf{F}\boldsymbol{\rho},$$

with the assembled representation as

$$\mathcal{F}^N(\rho,v) = \underline{\mathbf{v}}^T\mathbf{Q}^T\mathbf{B}\mathbf{Q}\underline{\boldsymbol{\rho}} = \underline{\mathbf{v}}^T\bar{\mathbf{F}}\underline{\boldsymbol{\rho}}.$$

Matrix T: We next discuss how to compute $s_i^{\hat{e},p}$ in Eq. (79). Note that the data is precomputed only once. One might apply the GLL quadrature for the integrations when d_p is small. For large d_p, however, the GLL quadrature is not accurate enough to capture the high-frequency modes, leading to loss of accuracy in the solution.

One can consider the discrete FFT algorithm since it is the pth component of the inverse DFFT of function $l_i(\xi)$. However, since $l_i(\xi)$ has only a very small portion of compact support on Γ, we can compute it directly on its local compact support using refined GLL quadrature points on each $\Gamma^{\hat{e}}$. Another approach is to use the relation to the Bessel function, which can be more efficient than the other approach.

In this paper, we discuss the computation of $s_i^{\hat{e},p}$ based on the Bessel function representation. We have written $l_i(\xi)$ in the finite expansion of the mth-order Legendre polynomials given as

$$l_i(\xi) = \sum_{m=0}^{N} (\hat{l}_i)_m L_m(\xi), \tag{81}$$

where $(\hat{l}_i)_m$ are the Legendre expansion coefficients defined by

$$(\hat{l}_i)_m = \frac{2m+1}{2} \int_{-1}^{1} l_i(\xi) L_m(\xi) d\xi. \tag{82}$$

Then, substituting (81) in (77) and using simply the notation x, instead of x', we have

$$s_i^{\hat{e},p} = \frac{1}{\sqrt{d}} \int_{\Gamma^{\hat{e}}} l_i(\xi(x)) e^{-\mathbf{i} d_p x} dx = \frac{1}{\sqrt{d}} \sum_{m=0}^{N} (\hat{l}_i)_m \left(\int_{-1}^{1} L_m(\xi) e^{-\mathbf{i} d_p x(\xi)} J_s^{\hat{e}} d\xi \right), \tag{83}$$

where $J_s^{\hat{e}}$ is the surface Jacobian on $\Gamma^{\hat{e}}$. In fact, each $\Gamma^{\hat{e}}$ is represented by an interval $[x_{\min}^{\hat{e}}, x_{\max}^{\hat{e}}]$ with the coordinate transformation by $x(\xi) = \hat{a}_e \xi + \hat{b}_e$ with $\hat{a}_e = (x_{\max}^{\hat{e}} - x_{\min}^{\hat{e}})/2$ and $\hat{b}_e = (x_{\max}^{\hat{e}} + x_{\min}^{\hat{e}})/2$, so that $J_s^{\hat{e}} \equiv \hat{a}_e$ is constant on $\Gamma^{\hat{e}}$. Then, Eq. (83) becomes

$$s_i^{\hat{e},p} = \frac{\hat{a}_e}{\sqrt{d}} \sum_{m=0}^{N} (\hat{l}_i)_m d_m^{p,\hat{e}} \quad \text{with} \quad q_m^{p,\hat{e}} = \int_{-1}^{1} L_m(\xi) e^{-\mathbf{i} d_p (\hat{a}_e \xi + \hat{b}_e)} d\xi. \tag{84}$$

Now we need to compute the two terms $(\hat{l}_i)_m$ and $q_m^{p,\hat{e}}$, in (84). To compute $(\hat{l}_i)_m$, one might apply the GLL quadrature for the integration term in (82) as follows:

$$(\hat{l}_i)_m = \frac{2m+1}{2} \sum_{k=0}^{N} l_i(\xi_k) L_m(\xi_k) w_k = \frac{2m+1}{2} L_m(\xi_i) w_i. \tag{85}$$

An alternative approach is to evaluate (81) on the GLL grids in $[-1,1]$, resulting in the form

$$\mathbf{L}\hat{\mathbf{L}} = \begin{bmatrix} L_0(\xi_0) & L_1(\xi_0) & \cdots & L_m(\xi_0) \\ \vdots & \vdots & \vdots & \vdots \\ L_0(\xi_N) & L_1(\xi_N) & \cdots & L_m(\xi_N) \end{bmatrix} \begin{bmatrix} (\hat{l}_0)_0 & (\hat{l}_1)_0 & \cdots & (\hat{l}_N)_0 \\ \vdots & \vdots & \vdots & \vdots \\ (\hat{l}_0)_N & (\hat{l}_1)_N & \cdots & (\hat{l}_N)_N \end{bmatrix} \equiv \mathbf{I},$$

and compute the inverse of the matrix $\mathbf{L} = [\mathbf{L}_{ji}] = [L_i(\xi_j)]$ to obtain $\hat{\mathbf{L}} = [\hat{\mathbf{L}}_{mi}] = [(\hat{l}_i)_m] = \mathbf{L}^{-1}$. To compute $q_m^{p,\hat{e}}$, we recall that the Legendre polynomials are related to the Bessel functions as

$$\int_{-1}^{1} L_m(\xi) e^{-\mathbf{i} x \xi} d\xi = \frac{1}{\mathbf{i}^m} \sqrt{\frac{2\pi}{x}} J_{m+1/2}(x) = \frac{2}{\mathbf{i}^m} j_m(x) \quad \text{for} \quad x \in R,$$

where j_m is the spherical Bessel function and J_m is the ordinary Bessel function with the relation

$$j_m(x) = \sqrt{\frac{\pi}{2x}} J_{m+1/2}(x).$$

Then, we can write

$$q_m^{p,\hat{e}} = \int_{-1}^{1} L_m(\xi) e^{-\mathbf{i} d_p(\hat{a}_e \xi + \hat{b}_e)} d\xi = e^{-\mathbf{i} d_p \hat{b}_e} \left(\frac{2}{\mathbf{i}^m} j_m(d_p \hat{a}_e) \right). \tag{86}$$

From (85) and (86), we have the final form of $s_i^{\hat{e},p}$ by

$$s_i^{\hat{e},p} = \frac{\hat{a}_e e^{-\mathbf{i} d_p \hat{b}_e}}{\sqrt{d}} \sum_{m=0}^{N} (\hat{l}_i)_m \left(\frac{2}{\mathbf{i}^m} j_m(d_p \hat{a}_e) \right).$$

3.5 Matrix Structure and Eigenvalues

We arrange our solution as a real vector of length $2n$ expressed by $\mathbf{u}^N = [u_{\mathrm{r}}^N, u_{\mathrm{i}}^N]^T$, where u_{r}^N and u_{i}^N represent real and imaginary parts of the solution, respectively. The spectral element discretization leads to a linear system:

$$\mathcal{H} \mathbf{u}^N = \mathcal{F}, \tag{87}$$

where

$$\mathcal{H} := \begin{bmatrix} \mathbf{A} - (k^2 - \alpha^2)\mathbf{B} + \mathbf{T}_{\mathrm{r}} & -\mathbf{T}_{\mathrm{i}} - 2\alpha\mathbf{C} \\ \mathbf{T}_{\mathrm{i}} + 2\alpha\mathbf{C} & \mathbf{A} - (k^2 - \alpha^2)\mathbf{B} + \mathbf{T}_{\mathrm{r}} \end{bmatrix} \text{ and } \mathcal{F} := \begin{bmatrix} \mathbf{F}\boldsymbol{\rho}_{\mathrm{r}} \\ \mathbf{F}\boldsymbol{\rho}_{\mathrm{i}} \end{bmatrix}.$$

Equation (87) in assembled representation can be expressed as

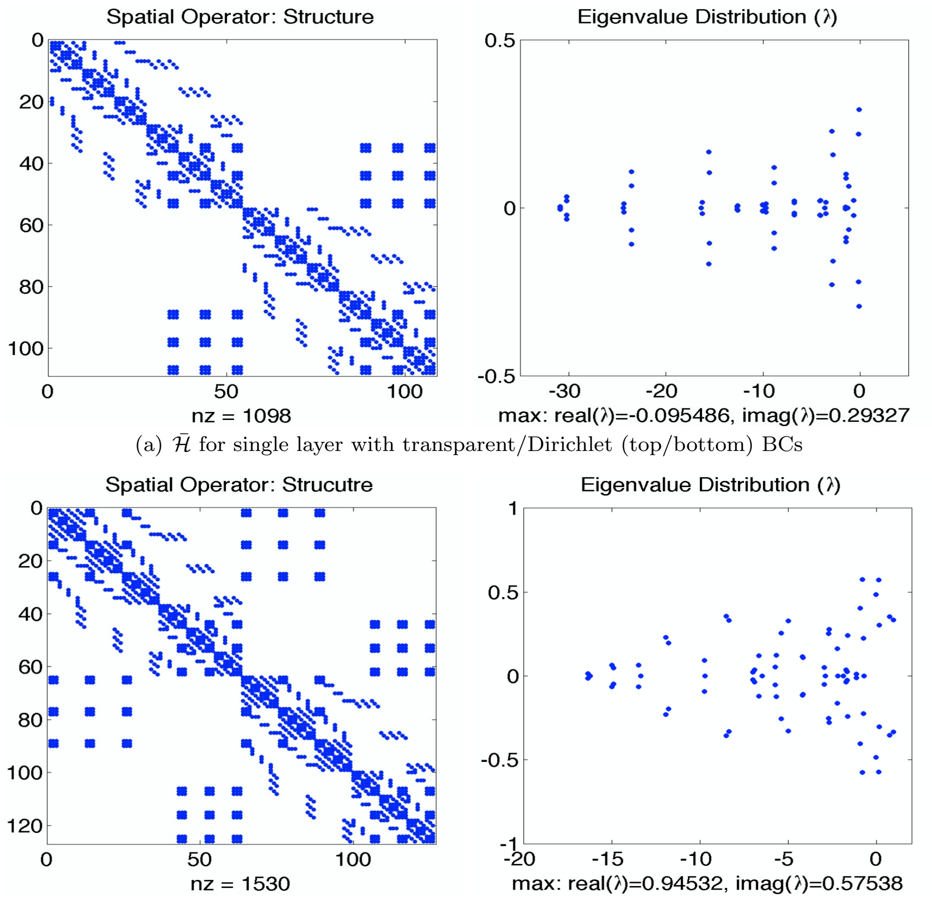
$$\bar{\mathcal{H}} \underline{\mathbf{u}}^N = \bar{\mathcal{F}}, \tag{88}$$

where

$$\bar{\mathcal{H}} := \begin{bmatrix} \bar{\mathbf{A}} - (k^2 - \alpha^2)\bar{\mathbf{B}} + \bar{\mathbf{T}}_{\mathrm{r}} & -\bar{\mathbf{T}}_{\mathrm{i}} - 2\alpha\bar{\mathbf{C}} \\ \bar{\mathbf{T}}_{\mathrm{i}} + 2\alpha\bar{\mathbf{C}} & \bar{\mathbf{A}} - (k^2 - \alpha^2)\bar{\mathbf{B}} + \bar{\mathbf{T}}_{\mathrm{r}} \end{bmatrix} \text{ and } \bar{\mathcal{F}} := \begin{bmatrix} \bar{\mathbf{F}}\underline{\boldsymbol{\rho}}_{\mathrm{r}} \\ \bar{\mathbf{F}}\underline{\boldsymbol{\rho}}_{\mathrm{i}} \end{bmatrix}.$$

In Figure 3, we demonstrate the structure of matrix and its eigenvalue distribution for our spectral element operator. For simplicity, we chose a simple box geometry for the domain $[0, 2\pi] \times [-1, 1]$ with equi-sized non-deformed rectangular elements (3 elements in x and 2 elements in y directions) and a relatively small $N = 3$. Figure 3(a) demonstrates the case of single layer with DtN boundary on the top and Dirichlet boundary at the bottom and the wave number $k = 1.5$, and Figure3(b) demonstrates the case of double layer, defining Γ_g at $y = 0$, with DtN boundaries at the top and bottom and the wave numbers $k = 1.5$ on the top layer and $k = 2.5$ on the bottom layer.

In Table 1, we list the condition numbers for these operators. The resulting linear system (88) is not Hermitian positive definite and thus it was natural choice to consider the GMRES method [36] for its solution.



(a) $\bar{\mathcal{H}}$ for single layer with transparent/Dirichlet (top/bottom) BCs

(b) $\bar{\mathcal{H}}$ for double layer with transparent/transparent (top/bottom) BCs

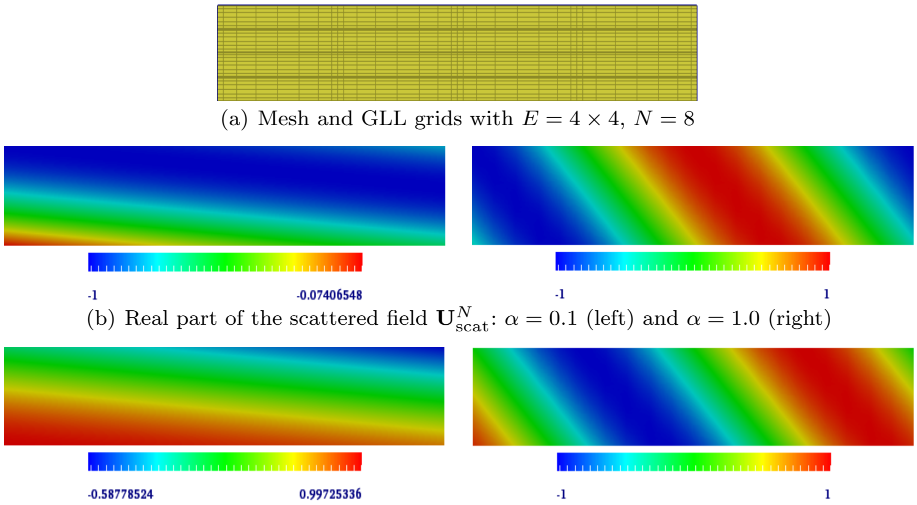
Fig. 3 Spatial operator in matrix structure (assembled) and its eigenvalue distribution

Table 1 Condition numbers for $\bar{\boldsymbol{\mathcal{H}}}$

Transparent (top)			Transparent (top/bottom)		
E	N	Condition #	E	N	Condition #
3×2	3	1.1873947E+02	3×2	3	5.9542135E+01
	5	4.6002926E+02		5	2.3470212E+02
	7	1.1704056E+03		7	5.9022047E+02
	9	2.4000224E+03		9	1.2031346E+03
	11	4.2909436E+03		11	2.1437669E+03
	13	6.9854623E+03		13	3.4825776E+03

4 Computational Results

In this section, we consider scattering returns by three types of periodic grating surfaces: flat, smooth curved, and nonsmooth. We consider different angles of incidence impinging on the scattering surface in singly and doubly layered media.



(a) Mesh and GLL grids with $E = 4 \times 4$, $N = 8$

(b) Real part of the scattered field $\mathbf{U}^N_{\text{scat}}$: $\alpha = 0.1$ (left) and $\alpha = 1.0$ (right)

(c) Imaginary part of the scattered field $\mathbf{U}^N_{\text{scat}}$: with $\alpha = 0.1$ (left) and $\alpha = 1.0$ (right)

Fig. 4 Single layer: $k = 1.5$ (yellow); transparent (top) and Dirichlet (bottom) boundary conditions.

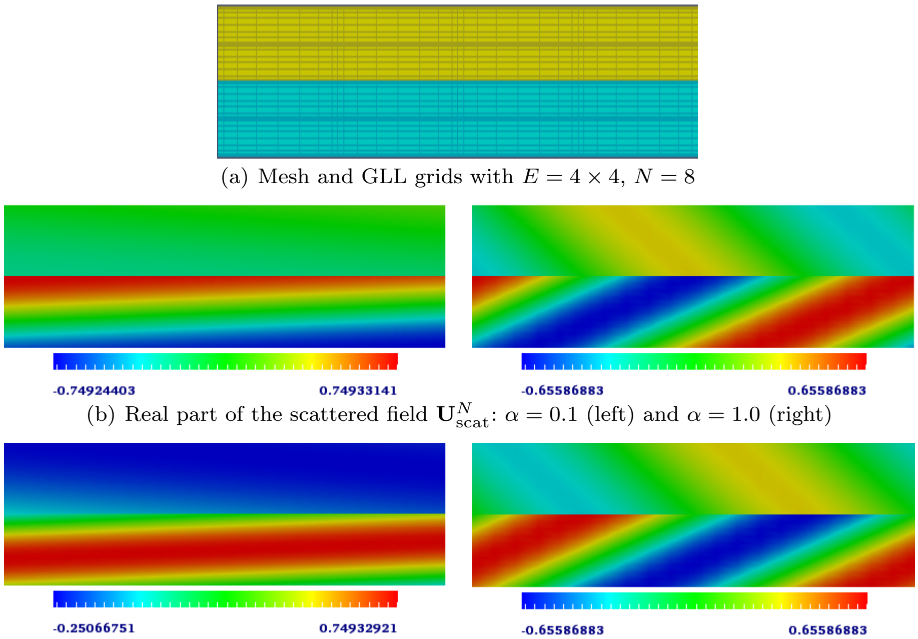
We solve the scalar Helmholtz equation and compute the total field in a finite computational domain with transparent boundary conditions enforced at artificial boundaries based on the spectral element discretization. For validation of our computational approach, in the case of a flat grating we compare our results with analytic solutions and provide convergence studies. For smooth curved periodic surface gratings, we consider sinusoidal grooves and compare our results with those from the transformed field expansion (TFE) method [20,13]. For nonsmooth periodic surface gratings (rectangular, triangular, and sawtooth) separating doubly layered media we demonstrate the accuracy of our computational solutions by studying the energy defect [4–6,16].

4.1 Flat Scattering Surface

To begin, we consider singly and doubly layered media with flat interface in the x-direction. For these configurations, there exist analytic solutions for incident waves at arbitrary angles of incidence $\kappa = (\alpha, -\beta)$. Here we consider downward propagating incidence with $\beta > 0$.

Single Layer: Consider a finite computational domain $\Omega = [0, 2\pi] \times [0, 1]$ with the scattering surface defined by $\Gamma_g = \{(x, y) \in \Omega \mid y = 0\}$ and the artificial boundary defined at $\Gamma = \{(x, y) \in \Omega \mid y = 1\}$. We apply homogeneous Dirichlet boundary conditions on the scatterer Γ_g and a transparent boundary condition via the DtN operator on Γ. Figure 4(a) shows our quadrilateral element mesh with $E = 4 \times 4$ and the GLL grids for $N = 8$. Considering the incident field

$$U_{\text{inc}}(x, y) = e^{\mathbf{i}(\alpha x - \beta(y+1))},$$



(a) Mesh and GLL grids with $E = 4 \times 4$, $N = 8$

(b) Real part of the scattered field $\mathbf{U}^N_{\text{scat}}$: $\alpha = 0.1$ (left) and $\alpha = 1.0$ (right)

(c) Imaginary part of the scattered field $\mathbf{U}^N_{\text{scat}}$: $\alpha = 0.1$ (left) and $\alpha = 1.0$ (right)

Fig. 5 Double layer: $k^+ = 1.5$ (yellow) and $k^- = 2.5$ (blue); transparent (top/bottom) boundary conditions.

impinging on Γ_g, we can show the total field solution to be

$$U^{\text{exact}}(x, y) = e^{\mathbf{i}(\alpha x - \beta(y+1))} - e^{\mathbf{i}(\alpha x + \beta(y-1))}.$$

For a fixed wavenumber $k = 1.5$ in the single layer medium, we consider incident waves for $\alpha = 0.1$ and $\alpha = 1.0$. Figures 4(b)–4(c) show the numerical solutions of the scattered fields that are obtained by subtracting the incident field from the total field: $\mathbf{U}^N_{\text{scat}} = \mathbf{U}^N - \mathbf{U}^N_{\text{inc}}$, where $\mathbf{U}^N_{\text{inc}}$ denotes the incident field $U^{\text{exact}}_{\text{inc}}$ evaluated on the GLL grid.

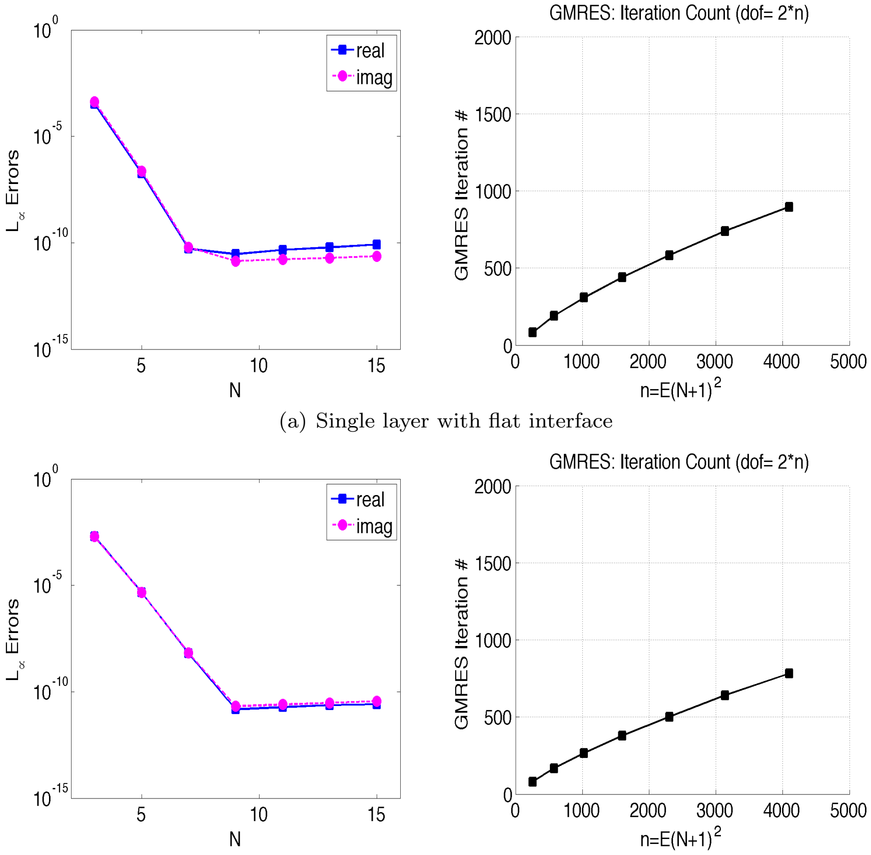
Double Layer: We now consider a computational domain $\Omega = [0, 2\pi] \times [-1, 1]$ with flat scattering surface $\Gamma_g = \{(x, y) \in \Omega \mid y = 0\}$ and artificial boundaries at $\Gamma = \Gamma^+ \cup \Gamma^-$, where $\Gamma^+ = \{(x, y) \in \Omega \mid y = 1\}$ and $\Gamma^- = \{(x, y) \in \Omega \mid y = -1\}$. We apply transparent boundary conditions here using the DtN operator on the GLL points on Γ. Figure 5(a) shows our mesh with $E = 4 \times 4$ and the GLL grids for $N = 8$. The incident field and analytic solution are given as follows:

- On $\Omega^+ = [0, 2\pi] \times [0, 1]$ with $k^+ = 1.5$ and $\beta^+ > 0$:

$$U_{\text{inc}}(x, y) = e^{\mathbf{i}(\alpha x - \beta^+ y)},$$
$$U^{\text{exact}}(x, y) = e^{\mathbf{i}(\alpha x - \beta^+ y)} + c^+ e^{\mathbf{i}(\alpha x + \beta^+ y)}.$$

- On $\Omega^- = [0, 2\pi] \times [-1, 0]$ with $k^- = 2.5$ and $\beta^- > 0$:

$$U^{\text{exact}}(x, y) = c^- e^{\mathbf{i}(\alpha x - \beta^- y)}.$$



(a) Single layer with flat interface

(b) Double layer with flat interface

Fig. 6 Convergence and GMRES iteration counts versus mesh refinement with E=4×4 and N=3,5,7,9,11,13,15. The approximation order for the Fourier expansion in the DtN operator is $P = 5$.

Here the (Fresnel) constants are

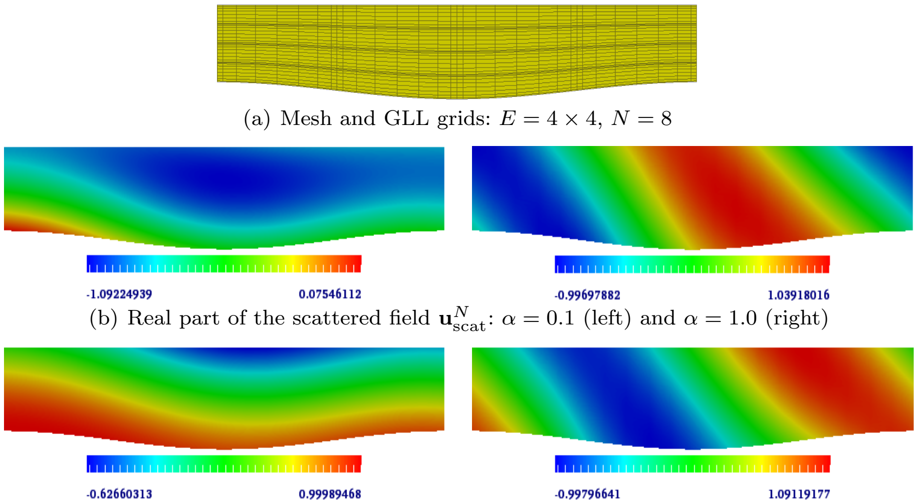
$$c^- = \frac{2\beta^+}{\beta^+ + \beta^-}, \quad c^+ = \frac{\beta^+ - \beta^-}{\beta^+ + \beta^-}.$$

Again, we consider incoming incident waves on Ω^+ for $\alpha = 0.1$ and $\alpha = 1.0$, and we simulate the total field. In Figures 5(b)–5(c) we display the scattered field $\mathbf{U}^N_{\text{scat}}$.

Convergence: Figure 6 depicts the outcomes of our convergence studies, measured in the maximum error, for scattered fields in singly and doubly layered media:

$$\text{error} = \|U^{\text{exact}}_{\text{scat}} - \mathbf{U}^N_{\text{scat}}\|_\infty,$$

where $U^{\text{exact}}_{\text{scat}} = U^{\text{exact}} - U_{\text{inc}}$ is the exact solution for the scattered field. The errors show spectral convergence as N increases. The approximation order for the Fourier data used in the DtN operator is $P = 5$. Table 1 shows that the condition



(a) Mesh and GLL grids: $E = 4 \times 4$, $N = 8$

(b) Real part of the scattered field $\mathbf{u}^N_{\text{scat}}$: $\alpha = 0.1$ (left) and $\alpha = 1.0$ (right)

(c) Imaginary part of the scattered field $\mathbf{u}^N_{\text{scat}}$: $\alpha = 0.1$ (left) and $\alpha = 1.0$ (right)

Fig. 7 Single layer: $k = 1.5$ (yellow); transparent (top) and Dirichlet (bottom) boundary conditions.

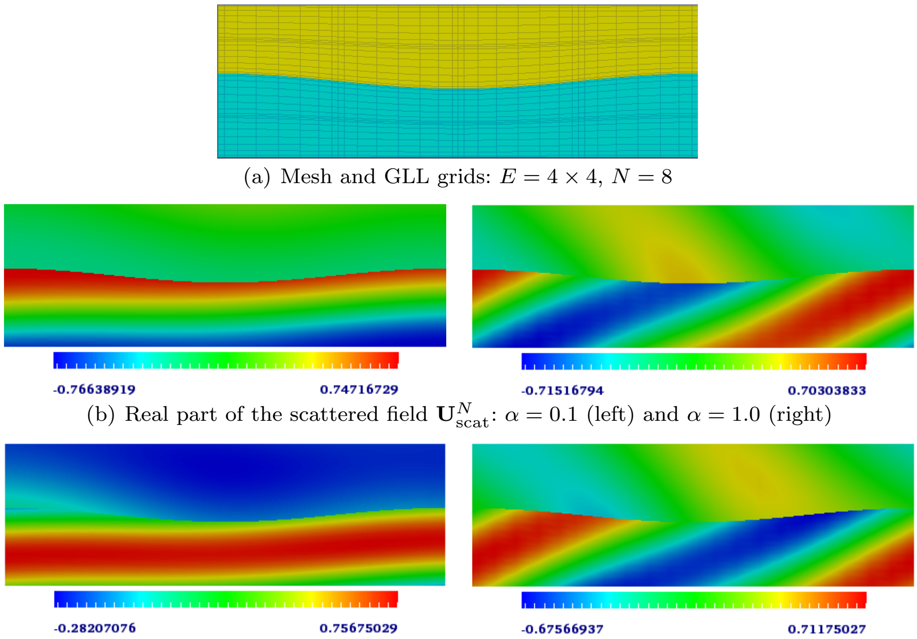
numbers increase as N increases, thus explaining why the errors do not improve beyond 10^{-10}. Figures 6(a)–6(b) demonstrate the iteration count increasing up to ~ 900 for $N = 15$.

Computation: In practice, we transform $\mathbf{U}^N_{\text{inc}}$ into $\mathbf{u}^N_{\text{inc}} = e^{-\mathbf{i}\alpha x}\mathbf{U}^N_{\text{inc}}$ and compute the solution of Eq. (11) $\mathbf{u}^N$ with periodic boundary treatment in x. Then, we transform back to $\mathbf{U}^N$ through the relation $\mathbf{U}^N = e^{\mathbf{i}\alpha x}\mathbf{u}^N$. This approach makes our algorithm much simpler by eliminating additional boundary treatments in the x-direction. The same idea is applied for solving all other example problems presented in the remaining sections.

4.2 Smooth Curved Scattering Surfaces

In this section, we examine singly and doubly layered media with smooth, curved interfaces. Dirichlet and transparent boundary conditions are once again applied in the y-direction. For these configurations no analytic solutions are available, so we validate our results in comparison with results provided by the TFE method [20, 13].

Single Layer: Consider a computational domain $\Omega = [0, 2\pi] \times [g(x), 1]$ with the scattering surface defined by $\Gamma_g = \{(x, y) \in \Omega \mid y = g(x)\}$ and an artificial boundary defined on $\Gamma = \{(x, y) \in \Omega \mid y = 1\}$. We choose a sinusoidal interface $g(x) = \epsilon \cos(x)$ with the grating depth varying with ϵ. We apply homogeneous Dirichlet boundary conditions on Γ_g and the DtN operator on Γ. Figure 7(a) displays the mesh with $E = 4 \times 4$ and the GLL grids for $N = 8$, representing $g(x)$ with surface fitted



(a) Mesh and GLL grids: $E = 4 \times 4$, $N = 8$

(b) Real part of the scattered field $\mathbf{U}^N_{\text{scat}}$: $\alpha = 0.1$ (left) and $\alpha = 1.0$ (right)

(c) Imaginary part of the scattered field $\mathbf{U}^N_{\text{scat}}$: $\alpha = 0.1$ (left) and $\alpha = 1.0$ (right)

Fig. 8 Double layer: $k^+ = 1.5$ (yellow) and $k^- = 2.5$ (blue); transparent (top/bottom) boundary conditions.

elements for the case of $\epsilon = 0.1$. We consider the incident field

$$U_{\text{inc}}(x, y) = e^{\mathbf{i}(\alpha x - \beta y)}$$

with varying incident angles $\alpha = 0.1$ and $\alpha = 1.0$ for a fixed wavenumber $k = 1.5$ with $\beta > 0$. The scattered fields are shown in Figures 7(b)–7(c).

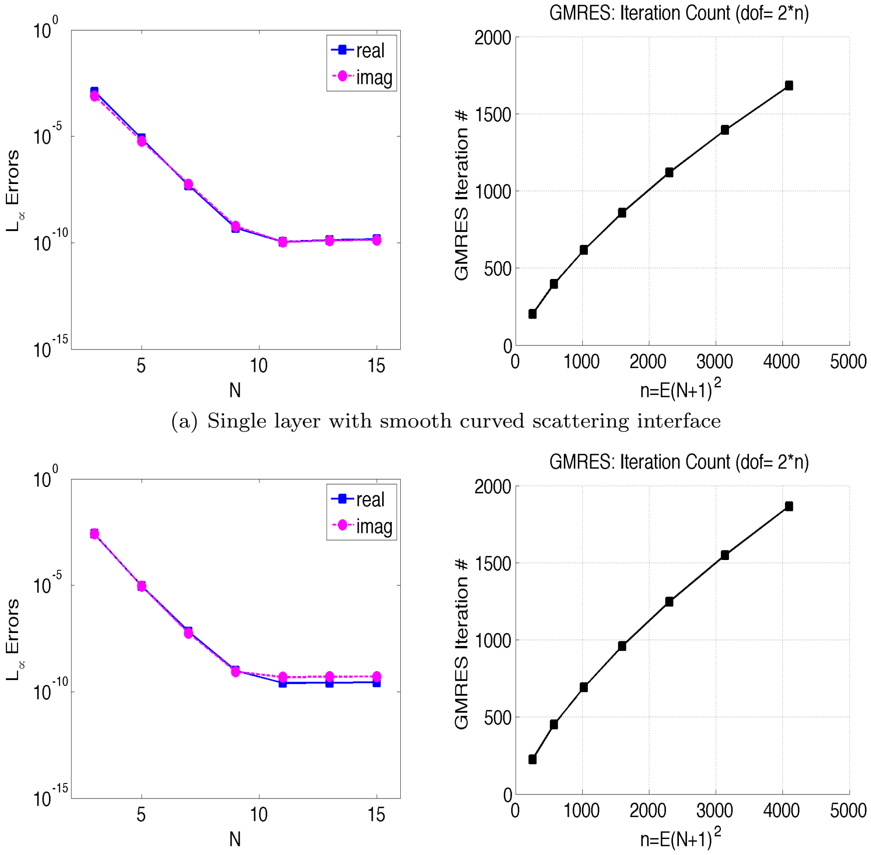
Double Layer: Consider a computational domain $\Omega = \Omega^- \cup \Omega^+$, consisting of two different media $\Omega^+ = [0, 2\pi] \times [g(x), 1]$ and $\Omega^- = [0, 2\pi] \times [-1, g(x)]$ with a sinusoidal interface shaped by $g(x) = \epsilon \cos(x)$. We define the artificial boundaries at $\Gamma = \Gamma^+ \cup \Gamma^-$, where $\Gamma^+ = \{(x, y) \in \Omega \mid y = 1\}$ and $\Gamma^- = \{(x, y) \in \Omega \mid y = -1\}$. Figure 8(a) shows the mesh with $E = 4 \times 4$ and the GLL grids for $N = 8$, representing $g(x)$ with surface-fitted elements for the case of $\epsilon = 0.1$. We consider incoming incident waves

$$U_{\text{inc}}(x, y) = e^{\mathbf{i}(\alpha x - \beta^+ y)},$$

in Ω^+ with $k^+ = 1.5$ and $\beta^+ > 0$ and varying incidence angles $\alpha = 0.1$ and $\alpha = 1.0$. We choose the wavenumber $k^- = 2.5$ on Ω^-, and in Figures 8(b)–8(c) we show the scattered fields.

Convergence: Figure 9 displays the convergence of our numerical solutions, measured in the maximum error, for the scattered field in singly and doubly layered media in comparison with results given by the TFE method:

$$\text{error} = \|U^{\text{TFE}}_{\text{scat}} - \mathbf{U}^N_{\text{scat}}\|_\infty.$$



(a) Single layer with smooth curved scattering interface

(b) Double layer with smooth curved scattering interface

Fig. 9 Convergence and GMRES iteration counts versus mesh refinement with E=4×4 and N=3,5,7,9,11,13,15. The approximation order for the Fourier expansion in the DtN operator is $P = 5$.

Here $U^{\mathrm{TFE}}_{\mathrm{scat}}$ is the scattered field approximation given by the TFE method. Our solution $\mathbf{U}^N_{\mathrm{scat}}$ on the GLL grids is interpolated to the TFE grid in order to compute the difference of the solutions on the same grids. The approximation order for the Fourier data used in the DtN operator is fixed with $P = 5$. In Figure 9, the errors show spectral convergence as N increases with the GMRES iteration count increasing up to 1700 $\sim$ 1900 for $N = 15$, as demonstrated in Figures 9(a)–9(b). We note that computational results demonstrated throughout the paper, we used *tolerance* of 1E-10 for the GMRES solution.

In Tables 2–4, we demonstrate the convergence of varying wave numbers and P using the same mesh configuration as in Figure 8(a). In Table 2, we show convergence for varying wave numbers, ranging $k^- = 2.5 \sim 32.5$, with a fixed $k^+ = 1.5$ and a relatively fine resolution with $N = 17$ and $P = 9$. The errors increase because the resolution in terms of N and P is not enough to capture the higher frequency waves well compared to the lower frequency waves as k^- increases. In Tables 3–4, we used fine grid resolution with a higher N and observe

Table 2 Convergence of $\mathrm{error_r} = \|real(U_{\mathrm{scat}}^{\mathrm{TFE}} - \mathbf{U}_{\mathrm{scat}}^{N})\|_\infty$, $\mathrm{error_i} = \|imag(U_{\mathrm{scat}}^{\mathrm{TFE}} - \mathbf{U}_{\mathrm{scat}}^{N})\|_\infty$ and GMRES iteration count for varying wave numbers with $P = 9$ and $N = 17$.

$g(x) = \cos(x)$, Oblique Incidence $\alpha = 0.1$				
E	(k^+, k^-)	error$_r$	error$_i$	iter #
16	(1.5,2.5)	1.7524E-10	1.6718E-10	2199
	(1.5,4.5)	1.2329E-10	1.3031E-10	2578
	(1.5,8.5)	9.9584E-11	9.0917E-10	3373
	(1.5,16.5)	6.2485E-09	5.3194E-09	4846
	(1.5,22.5)	2.0038E-07	1.2528E-07	5784
	(1.5,32.5)	1.2439E-05	1.2840E-05	6742

Table 3 Convergence of $\mathrm{error_r} = \|real(U_{\mathrm{scat}}^{\mathrm{TFE}} - \mathbf{U}_{\mathrm{scat}}^{N})\|_\infty$, $\mathrm{error_i} = \|imag(U_{\mathrm{scat}}^{\mathrm{TFE}} - \mathbf{U}_{\mathrm{scat}}^{N})\|_\infty$ and GMRES iteration count for varying P with $N = 11$ and $(k^+, k^-) = (1.5, 2.5)$.

$g(x) = \cos(x)$, Oblique Incidence $\alpha = 0.1$				
E	P	error$_r$	error$_i$	iter #
16	0	4.0639E-01	8.5534E-01	1391
	1	4.5017E-03	7.3669E-03	1304
	2	2.8951E-05	2.0889E-05	1276
	3	3.7341E-07	4.9301E-07	1248
	4	9.1279E-09	1.4023E-08	1247
	5	2.5614E-10	4.8626E-10	1247
	6	1.1452E-10	1.1020E-10	1247
	7	1.0750E-10	1.0557E-10	1247
	8	1.0368E-10	1.0009E-10	1247
	9	1.0414E-10	1.0169E-10	1247

Table 4 Convergence of $\mathrm{error_r} = \|real(U_{\mathrm{scat}}^{\mathrm{TFE}} - \mathbf{U}_{\mathrm{scat}}^{N})\|_\infty$, $\mathrm{error_i} = \|imag(U_{\mathrm{scat}}^{\mathrm{TFE}} - \mathbf{U}_{\mathrm{scat}}^{N})\|_\infty$ and GMRES iteration count for varying P with $N = 13$ and $(k^+, k^-) = (1.5, 8.5)$.

$g(x) = \cos(x)$, Oblique Incidence $\alpha = 0.1$				
E	P	error$_r$	error$_i$	iter #
16	0	1.6319E+00	1.3417E+00	2601
	1	5.9357E-01	5.6400E-01	2571
	2	2.4980E-01	2.4801E-01	2550
	3	4.8806E-03	5.9880E-03	2490
	4	2.0239E-04	1.5874E-04	2432
	5	2.0005E-06	1.4451E-06	2351
	6	7.0879E-08	3.6509E-08	2381
	7	3.0704E-10	2.8059E-10	2361
	8	1.0474E-10	1.1638E-10	2386
	9	9.7235E-11	9.9168E-11	2384
	10	9.0224E-11	9.6817E-11	2384
	11	9.3851E-11	9.8553E-11	2385

the convergence as we increase P. Table 3 demonstrates the convergence with varying P with a fixed $N = 11$ and $(k^+, k^-) = (1.5, 2.5)$, showing that the relatively small $P = 5$ is a good choice as k^- is relatively small. Table 4 demonstrates the convergence with varying P with a fixed $N = 13$ and $(k^+, k^-) = (1.5, 8.5)$, showing that the relatively higher $P = 9$ is a good choice as k^- is relatively large. As for the convergence depending on the distance of the artificial boundary from the grating interface, the detailed study can be found in [37].

4.3 Nonsmooth Scattering Surfaces

To begin this section, we recall the scattering efficiencies and the energy defect measure of convergence. With these, we examine the behavior of our algorithm in a doubly layered media with rectangular, triangular, and sawtooth scattering interfaces, which severely challenge the capabilities of the TFE approach. We demonstrate the convergence of our method using this energy defect measure.

Energy Defect: We recall the Rayleigh expansions (14) for the reflected and transmitted fields,

$$U^{+}(x,y) = \sum_{p=-\infty}^{\infty} \hat{U}_p^{+} e^{i\alpha_p x + i\beta_p^{+} y}, \quad U^{-}(x,y) = \sum_{p=-\infty}^{\infty} \hat{U}_p^{-} e^{i\alpha_p x - i\beta_p^{-} y}, \tag{89}$$

and the efficiencies

$$e_p^{+} = \frac{\beta_p^{+}}{\beta} \left|\hat{U}_p^{+}\right|^2, \quad \text{and} \quad e_p^{-} = \frac{\beta_p^{-}}{\beta} \left|\hat{U}_p^{-}\right|^2, \tag{90}$$

which measure the energy at wave modes p propagated away from the grating interface. It is a classical calculation to show that for lossless media, a principle of conservation of energy [16] holds:

$$\sum_{p\in\mathbb{K}^{+}} e_p^{+} + \sum_{p\in\mathbb{K}^{-}} e_p^{-} = 1. \tag{91}$$

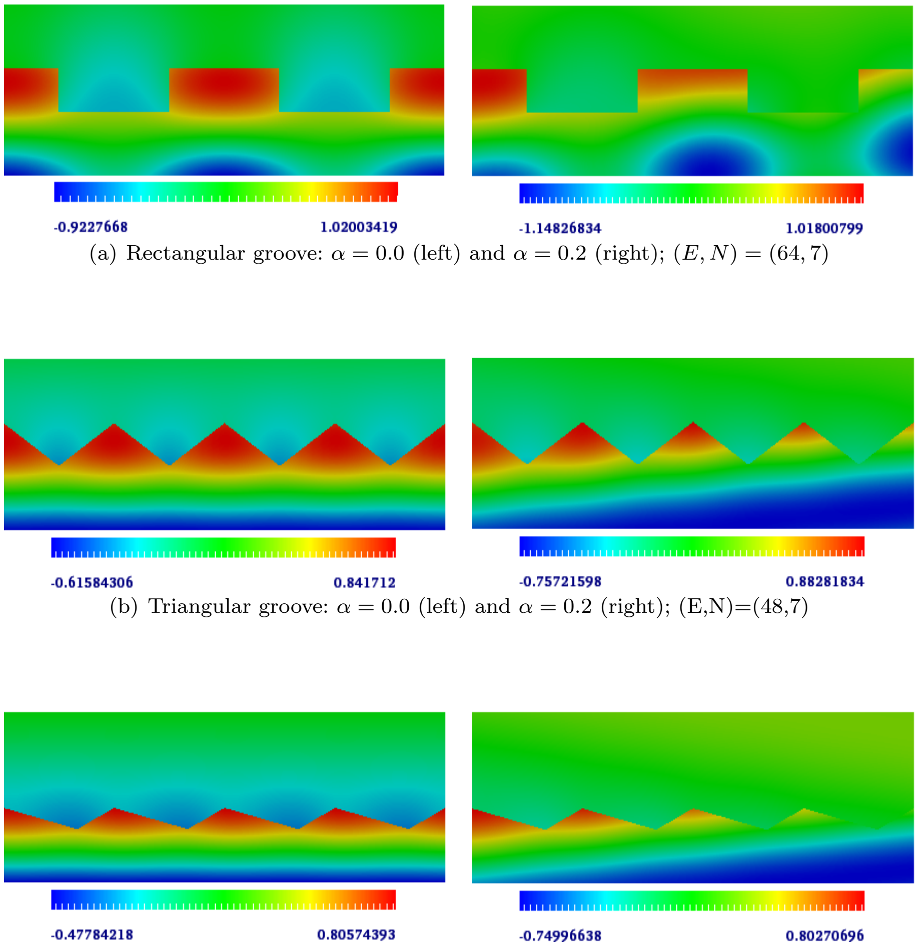
For an explicit demonstration we refer the interested reader to [20].

One measure of the fidelity of a numerical scheme for the approximation of scattering returns from a grating structure is to test the validity of this principle, for example, via the "energy defect" [16]:

$$\varepsilon_{\text{defect}} = \left| 1 - \left(\sum_{p\in\mathbb{K}^{+},\ |p|\le P} e_p^{+} + \sum_{p\in\mathbb{K}^{-},\ |p|\le P} e_p^{-} \right) \right|. \tag{92}$$

While it is not definitive, since the evanescent modes play no role in the energy defect, it is certainly indicative of a convergent scheme.

Double Layer: We consider a computational domain $\Omega = \Omega^{-} \cup \Omega^{+}$ with $\Omega^{+} = [0,2\pi] \times [g(x),1]$ and $\Omega^{-} = [0,2\pi] \times [-1,g(x)]$, including rectangular, triangular, and sawtooth grooves for the scattering surface $g(x)$, as shown in Figures 10–11. Artificial boundaries are set at $\Gamma = \Gamma^{+} \cup \Gamma^{-}$ for $\Gamma^{+} = \{(x,y) \in \Omega \mid y = 1\}$ and $\Gamma^{-} = \{(x,y) \in \Omega \mid y = -1\}$. We consider incoming incident waves $U^{\text{inc}}(x,y) = e^{\mathbf{i}(\alpha x - \beta^{+} y)}$ on Ω^{+} for varying incident angles of $\alpha = 0$ and $\alpha = 0.2$ with $k^{+} = 1.5$ and $\beta^{+} > 0$. The wavenumber $k^{-} = 2.5$ is defined on Ω^{-}, and Figures 10–11 show the computed scattered field. In Table 5, we demonstrate the convergence of our numerical solutions measured in the energy defect, showing spectral convergence as N and the number of GMRES iterations are increased to $700 \sim 1400$ for $N = 9$. The approximation order for the Fourier data used in the DtN operator is fixed with $P = 5$.



(a) Rectangular groove: $\alpha = 0.0$ (left) and $\alpha = 0.2$ (right); $(E, N) = (64, 7)$

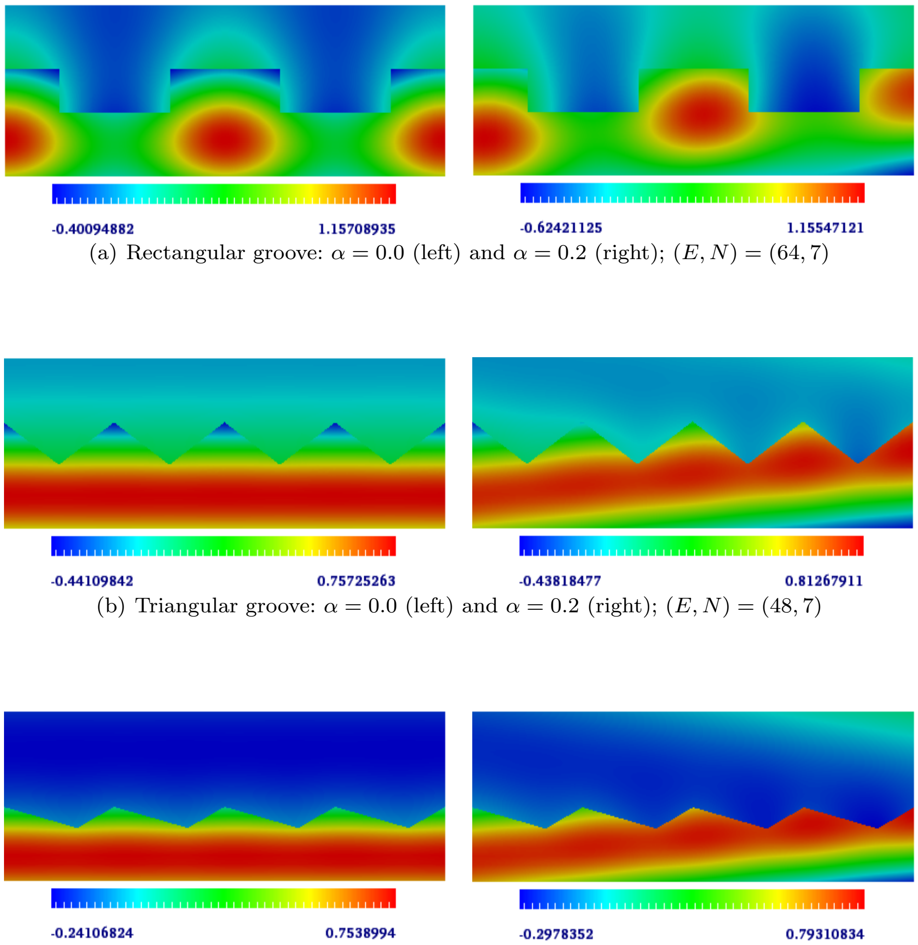
(b) Triangular groove: $\alpha = 0.0$ (left) and $\alpha = 0.2$ (right); (E,N)=(48,7)

(c) Sawtooth groove: $\alpha = 0.0$ (left) and $\alpha = 0.2$ (right); $(E, N) = (48, 7)$

Fig. 10 Real part of the scattered field.

5 Conclusions

In this contribution we have studied quasi-periodic solutions of the scalar Helmholtz equation in two dimensions in the context of layered media scattering problems. We have considered singly and doubly layered media with periodic surface interfaces. We used body-fitted quadrilateral element meshes with spectral element discretization based on the GLL grids. We imposed nonreflecting, outgoing boundary conditions at artificial boundaries which form a truncated computational domain. We introduced an accurate formulation of the spectral element DtN operator by representing the Fourier data in relation to the Bessel function, rather than computing the Fourier coefficients using the GLL quadrature integration, which can cause loss of accuracy depending on the grid resolution. Because of the quasi-periodicity of the solutions and the appearance of the DtN operator, the resulting



(a) Rectangular groove: $\alpha = 0.0$ (left) and $\alpha = 0.2$ (right); $(E, N) = (64, 7)$

(b) Triangular groove: $\alpha = 0.0$ (left) and $\alpha = 0.2$ (right); $(E, N) = (48, 7)$

(c) Sawtooth groove: $\alpha = 0.0$ (left) and $\alpha = 0.2$ (right); $(E, N) = (48, 7)$

Fig. 11 Imaginary part of the scattered field.

linear system is not Hermitian positive definite. Therefore, we applied the GMRES algorithm for solving the resulting linear system. We demonstrated our computational results for the scattered field and validated them with convergence studies showing spectral convergence.

Acknowledgements This material is based upon work supported in part by the U.S. Department of Energy Office of Science, under contract DE-ACO2-O6CH11357, and in part by an NSF grant no. DMS–1115333.

References

1. Prasanta Kumar Banerjee and Roy Butterfield, Boundary element methods in engineering science, McGraw-Hill, UK, 1981. Author, Article title, Journal, Volume, page numbers (year)
2. Marc Bonnet, Boundary integral equation methods for solids and fluids, Wiley, 1999.

Table 5 Convergence of the energy defect $\varepsilon_{\text{defect}}$ and GMRES iteration count ($P = 5$).

Rectangular Groove							
Normal Incidence $\alpha = 0$				Oblique Incidence $\alpha = 0.2$			
E	N	$\varepsilon_{\text{defect}}$	iter #	E	N	$\varepsilon_{\text{defect}}$	iter #
64	3	0.4352422E-03	226	64	3	0.4297330E-03	309
	5	0.7139313E-06	447		5	0.7013786E-06	638
	7	0.4966379E-09	704		7	0.4880223E-09	1001
	9	0.7938834E-12	998		9	0.1396099E-11	1412
Triangular Groove							
Normal Incidence $\alpha = 0$				Oblique Incidence $\alpha = 0.2$			
E	N	$\varepsilon_{\text{defect}}$	iter #	E	N	$\varepsilon_{\text{defect}}$	iter #
48	3	0.6962538E-02	160	48	3	0.6795656E-02	177
	5	0.4811381E-04	321		5	0.4707435E-04	349
	7	0.1354668E-06	515		7	0.1318199E-06	556
	9	0.2083588E-09	728		9	0.1873164E-09	782
Sawtooth Groove							
Normal Incidence $\alpha = 0$				Oblique Incidence $\alpha = 0.2$			
E	N	$\varepsilon_{\text{defect}}$	iter #	E	N	$\varepsilon_{\text{defect}}$	iter #
48	3	0.1407709E-01	182	48	3	0.1366347E-01	186
	5	0.4745677E-04	359		5	0.4642258E-04	368
	7	0.1337871E-06	563		7	0.1302132E-06	574
	9	0.1946178E-09	803		9	0.1829619E-09	813

3. L. Greengard and V. Rokhlin, A fast algorithm for particle simulations, *J. Comput. Phys.*, 73, 2, pp. 325–348 (1987).
4. Oscar P. Bruno and Fernando Reitich, Numerical solution of diffraction problems: A method of variation of boundaries, *J. Opt. Soc. Am. A*, 10, 6, pp. 1168–1175 (1993).
5. Oscar P. Bruno and Fernando 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, pp. 2397–2316 (1993).
6. Oscar P. Bruno and Fernando Reitich, Numerical solution of diffraction problems: A method of variation of boundaries III. Doubly periodic gratings, *J. Opt. Soc. Am. A*, 10, 12, pp. 2551–2562 (1993).
7. D. Michael Milder, An improved formalism for rough-surface scattering of acoustic and electromagnetic waves, Proceedings of SPIE - The International Society for Optical Engineering (San Diego, 1991), 1558, pp. 213–221, *Int. Soc. for Optical Engineering*, Bellingham, WA, 1991.
8. D. Michael Milder, An improved formalism for wave scattering from rough surfaces, *J. Acoust. Soc. Am.*, 89, 2, pp. 529–541 (1991).
9. D. Michael Milder, Role of the admittance operator in rough-surface scattering, *J. Acoust. Soc. Am.*, 100, 2, pp. 759–768 (1996).
10. D. Michael Milder, An improved formalism for electromagnetic scattering from a perfectly conducting rough surface, *Radio Science*, 31, 6, pp. 1369–1376 (1996).
11. D. Michael Milder and H. Thomas Sharp, Efficient computation of rough surface scattering, *Mathematical and numerical aspects of wave propagation phenomena (Strasbourg, 1991)*, pp. 314–322, SIAM, Philadelphia, PA, 1991.
12. D. Michael Milder and H. Thomas Sharp, An improved formalism for rough surface scattering. II: Numerical trials in three dimensions, *J. Acoust. Soc. Am.*, 91, 5, pp. 2620–2626 (1992).
13. David P. Nicholls and Jie Shen, A Rigorous Numerical Analysis of the Transformed Field Expansion Method, *SIAM Journal on Numerical Analysis*, 47, 4, pp. 2708–2734 (2009).
14. David P. Nicholls and Jie Shen, A Stable, High–Order Method for Two–Dimensional Bounded–Obstacle Scattering, *SIAM J. Sci. Comput.*, SIAM Journal on Scientific Computing, 28, 4, pp. 1398–1419 (2006).
15. Qirong Fang and David P. Nicholls and Jie Shen, A Stable, High–Order Method for Three–Dimensional Bounded–Obstacle Scattering, *J. Comput. Phys.*, Journal of Computational Physics, 224, 2, pp. 1145–1169 (2007).

16. R. Petit, Electromagnetic theory of gratings, *Springer-Verlag*, Berlin, 1980.
17. David P. Nicholls and Fernando Reitich, A new approach to analyticity of Dirichlet-Neumann operators, *Proc. Roy. Soc. Edinburgh Sect. A*, Proceedings of the Royal Society of Edinburgh. Section A. Mathematics, 131, 6, pp. 1411–1433 (2001).
18. David P. Nicholls and Fernando Reitich, Stability of High-Order Perturbative Methods for the Computation of Dirichlet-Neumann Operators, *J. Comput. Phys.*, Journal of Computational Physics, 170, 1, pp. 276–298 (2001).
19. David P. Nicholls and Fernando Reitich, Analytic Continuation of Dirichlet-Neumann Operators, *Numer. Math.*, Numerische Mathematik, 94, 1, pp. 107–146 (2003).
20. Ying He and David P. Nicholls and Jie Shen, An Efficient and Stable Spectral Method for Electromagnetic Scattering from a Layered Periodic Structure, *J. Comput. Phys.*, 231, 8, pp. 3007–3022 (2012).
21. David P. Nicholls, A Method of Field Expansions for Vector Electromagnetic Scattering by Layered Periodic Crossed Gratings, *J. Opt. Soc. Amer., A*, 32, 5, 701–709, (2015).
22. M. O. Deville and P.F. Fischer and E.H. Mund, High-order methods for incompressible fluid flow, Cambridge University Press, Cambridge, 2002.
23. Hou De Han and Xiao Nan Wu, Approximation of infinite boundary condition and its application to finite element methods, *J. Comput. Math.*, 3, 2, pp. 179–192 (1985).
24. Joseph Keller and Dan Givoli, Exact nonreflecting boundary conditions, *J. Comput. Phys.*, 82, 1, pp. 172–192 (1989).
25. Dan Givoli, Nonreflecting boundary conditions, *J. Comput. Phys.*, 94, 1, pp. 1–29 (1991).
26. Dan Givoli and Joseph Keller, Special finite elements for use with high-order boundary conditions, *Comput. Methods Appl. Mech. Engrg.*, 119, 3–4, pp. 199–213 (1994).
27. Dan Givoli, Numerical methods for problems in infinite domains, Elsevier Scientific Publishing Co., Amsterdam, 1992.
28. Marcus Grote and Joseph Keller, On nonreflecting boundary conditions, *J. Comput. Phys.*, 122, 2, pp. 231–243 (1995).
29. Dan Givoli, Recent advances in the DtN FE method, *Arch. Comput. Methods Engrg.*, 6, 2, pp. 71–116 (1999).
30. David P. Nicholls and Nilima Nigam, Exact Non-Reflecting Boundary Conditions on General Domains, *J. Comput. Phys.*, 194, 1, pp. 278–303 (2004).
31. David P. Nicholls and Nilima Nigam, Error Analysis of a Coupled Finite Element/DtN Map Algorithm on General Domains, *Numer. Math.*, 105, 2, pp. 267–298 (2006).
32. Tommy L. Binford and David P. Nicholls and Nilima Nigam and Timothy Warburton, Exact Non–Reflecting Boundary Conditions on General Domains and hp-Finite Elements, *J. Sci. Comput.*, 39, 2, pp. 265–292 (2009).
33. J. W. Strutt and Lord Rayleigh, On the manufacture and theory of diffraction gratings, *Philos. Mag.*, 47, 10, pp. 193–205 (1874).
34. G. Bao, Finite Element Approximation of Time Harmonic Waves in Periodic Structures, *SIAM J. Num. Anal.*, 32, 4, pp. 1155–1169 (1995).
35. Alex Barnett and Leslie Greengard, A new integral representation for quasi-periodic fields and its application to two-dimensional band structure calculations, *J. Comp. Phys.*, 229, pp. 6898–6914 (2010).
36. Youcef Saad and Martin H. Schultz, A generalized minimal residual algorithm for solving nonsymmetric linear systems, *SIAM J. Sci. Stat. Comput.*, 7, 3, pp. 856–869 (1986).
37. David P. Nicholls, Efficient enforcement of far-field boundary conditions in the transformed field expansions method, *Journal of Computational Physics*, 230, 22,pp. 8290–8303, 2011.

The following paragraph should be deleted before the paper is published: The submitted manuscript has been created by UChicago Argonne, LLC, Operator of Argonne National Laboratory (“Argonne”). Argonne, a U.S. Department of Energy Office of Science laboratory, is operated under Contract No. DE-AC02-06CH11357. The U.S. Government retains for itself, and others acting on its behalf, a paid-up nonexclusive, irrevocable worldwide license in said article to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government.